Paper 172

PILOT SENSITIVITY TO FLIGHT MODEL DYNAMICS IN ROTORCRAFT SIMULATION[1]

Emma Timson*, Philip Perfect, Mark White, Gareth D Padfield
School of Engineering, University of Liverpool

Robert Erdos, Arthur W. Gubbels
Flight Research Laboratory, Nation Research Council, Canada

ABSTRACT

The use of Synthetic Training Devices (STDs) for rotary wing pilot skill acquisition and development is an integral part of both military and civil flight crew training. The confidence in the capability of an STD to replicate flight rests on quantitative and qualitative standards and regulations that deem them fit for purpose. However, the engineering science base behind these standards is lacking at present. To amend this, the Lifting Standards research project is underway at the University of Liverpool addressing the deficiencies in the standards through the development of predicted and perceptual fidelity metrics. In this study, the application of these metrics to rotorcraft training simulators is of primary interest. A Simulator Fidelity Rating (SFR) scale has been developed, and is used to examine pilot sensitivity to variations in the training environment. The work detailed in this paper focuses on variations in flight model dynamics. In order to utilise the SFR scale, simulator trials at UoL and collaborative flight trials at the National Research Council's (NRC), Flight Research Laboratory (FRL) in Ottawa have been conducted. To ensure only effects of flight model fidelity were being measured and assessed, the flight model variations were implemented into a 'baseline' simulation and then compared against one another (simulator vs. simulator). A selection of test variations were then repeated in a variable stability helicopter at the FRL. Perceptual fidelity was rated for these model variants against the baseline model using the SFR scale and perceptual fidelity metrics used to quantify pilot control activity and performance. Results from these simulation and flight trials are presented in this paper. It was found that subjective ratings and pilot strategy adaptation are affected by both the nature of the task being flown and the level of pilot aggressiveness. It has been demonstrated that this methodology can be used to establish simulator functional fidelity tolerances.

ACRONYMS

AC	Advisory Circular
AD	Acceleration/Deceleration
ADS	Aeronautical Design Standard
ASRA	Advanced System Research Aircraft
CT&M	Correct Trend and Magnitude
EPSRC	Engineering and Physical Science Research Council
FAA	Federal Aviation Authority
FBW	Fly By Wire
FFS	Full Flight Simulator
FoV	Field of View
GARTEUR	Group for Aeronautical Research and Technology in EURope
HQ(R)	Handling Qualities (Rating)
ICAO	International Civil Aviation Organisation
ICATEE	International Committee for Aviation Training in Extended Envelopes
IWG	International Working Group
JAR	Joint Aviation Requirements
MTE	Mission Task Element
NRC	National Research Council
PH	Precision Hover
PTT	Part Task Trainer
RAeS	Royal Aeronautical Society
RCAH	Rate Command Attitude Hold
RPM	Revolutions Per Minute
SFR	Simulator Fidelity Rating
STD	Synthetic Training Device
UoL	University of Liverpool

INTRODUCTION

Major technological advances in flight simulation over the last 20 years, including developments in motion and visual systems and computing power, have led to a point where flight simulation devices can provide highly accurate replications of many flight regimes. This has led to increased utilisation of simulation for aircraft design, development and qualification and also as an integrated part of both military and civil flight crew training programs. The use of synthetic training devices provides cost, safety, environmental and efficiency benefits compared to in-flight training.

The utilisation of synthetic devices in training is not limited to full flight simulators (FFS) but also includes

[1]Submitted to the 37th European Rotorcraft Forum, 2011

*Corresponding author, Email: etimson@liv.ac.uk; Tel: +44(0) 151 7948501

the use of part task trainers (PTT), procedural trainers and desktop PCs to deliver and enhance pilot training [1]. The Royal Aeronautical Society (RAeS) reported that for an airline with 1000 pilots, recurrent training and checking using aircraft could cost some $60M annually, whereas the same procedures would cost one tenth of that if carried out using simulators [2]. Therefore a considerable amount of running-cost can be saved by replacing in-flight training hours with training in synthetic devices. Training in simulators also relieves the aircraft for operational duties which reduces lost revenue for commercial operators. For schools that utilise relatively low cost aircraft (for example the Grob Tutor or Robinson R22) the procurement of an FFS may not be cost effective due to the much lower procurement and running costs of the aircraft. In this case, the use of lower order devices could yield a more economically viable training alternative to flight.

Another benefit of flight simulation is that test conditions can be controlled and repeated and hence flying time is not weather dependent, allowing more convenient and efficient training to be undertaken. Environmental conditions such as wind and time of day can also be readily altered in a simulator, thereby allowing the pilot to be exposed to a wider range of conditions, so long as there are supporting data to validate the flight model in the chosen condition.

To ensure that the use of synthetic devices does not result in negative transfer of training to the aircraft, national and international standards are implemented to regulate the certification and use of, synthetic training facilities and training benefit derived from them. JAR-FSTD H [3] is the European standard for rotary wing simulators; the Federal Aviation Authority (FAA) Advisory Circular (AC) document FAA AC 120-63 [4] is the US equivalent and both are widely adopted to aid the design and assessment of simulation facilities. While these documents provide a framework around which a simulator can be certified, there are limitations. GARTEUR Action Group HC-AG12, “Validation Criteria for Helicopter Real-Time Simulation Models” [5] revealed a number of shortcomings in JAR-STD 1H (the earlier issue of JAR-FSTD H). One limitation identified was that the response tolerances in the rotorcraft standard were often read across from their fixed wing counterparts without clear justification of their application to rotorcraft simulators. Another limitation identified was that the sensitivity of the prescribed tolerances to task duration and task aggression had been neglected. The GARTEUR work highlighted the need for the evaluation of overall fidelity of the integrated system of pilot and machine. Supporting data and analysis techniques are required to verify that adhering to the current criteria guarantees that a simulator is of sufficient quality for the required purpose. Current initiatives from the Royal Aeronautical Society (RAeS) International Working Group (IWG), are striving to improve simulator training through the standardisation of simulation technology and development of a training-centred framework for simulator fidelity assessment [6].

The challenge for the operators and flight training organisations is to find the right balance between in-flight, and simulator training to deliver the most effective outcome and also run an economically viable business – simulator technology can be upgraded to produce very realistic scenarios, but if the cost of this becomes more than “live” training then there is no financial justification for the use of simulation. In addition, the benefit of being able to train for emergency situations, which cannot be trained for in the aircraft, may be compromised by lack of validation data for these situations. Although commentary from the Flight Data Recorder (FDR) will provide some insight into an accident, operational aircraft are not instrumented with equipment to supply validation data for simulators. Therefore pilot subjective opinion on how the training they receive in synthetic devices prepares them for dealing with a “real” emergency (and how the synthetic training might be improved) is crucial.

The use of simulation for training is not limited to the aviation industry but spreads across many fields including manufacturing and healthcare education [7]. Simulation is also extensively used for research and development over a wide range of industries. It is envisaged that the lessons learned in aviation can be transferred across into other disciplines and vice versa.

Transfer of training can be determined through both subjective and objective measures. A study by the US Army calculated transfer of training through a metric known as CTER – the Cumulative Training Effectiveness Ratio for two scenarios: firstly for the transfer of training between a simulator and an aircraft in an instructional environment and secondly, to determine the simulators effectiveness in maintaining/improving combat flying skills in an operational environment [8]. CTER is a measure of the time savings realised in learning to operate an aircraft by first training in a training device, where CTER=1 indicates that training in the synthetic device is as effective as training in the real aircraft. Another method for analysing simulator fidelity is through pilot sensitivity analysis. Mitchell et al, conducted an investigation of Maximum Unnoticeable Added Dynamics (MUAD) and have developed pilot sensitivity curves that can be used as fidelity acceptance boundaries [9]. Burki-Cohen et al. at the Volpe National Transportation Systems Centre have conducted numerous studies into the effect of motion on transfer of training and found there to be no operationally relevant difference between training in a FFS with motion and without motion, independent of the vehicle being simulated [10, 11]. The majority of the work into the effects of simulator fidelity on transfer of training focuses on motion and visual system fidelity. The current paper focuses on the effect of flight model variations on transfer of training.

The aim of this paper is to report an assessment of the effect of flight model fidelity on pilot control strategy in a training scenario, by determining how much transfer of training occurs in a simulation which differs from the aircraft in a known and quantifiable way. The current study is undertaken using both simulation, in Liverpool's HELIFLIGHT–R simulator [12], and flight trials at the NRC in their Bell 412 Advanced System Research Aircraft (ASRA) airborne simulator [13]. The study is rotorcraft based but with the intention that the methods can be applied to any flight vehicle. The following two sections describe the current approaches to assessing simulator fidelity, the need for a more training focused methodology and the work being done on this at UoL. The facilities used in the study and the methodology are then briefly described. Selected results from the trials into the effects of flight model variations on SFRs and pilot activity are then presented. This is followed by conclusions from the outcomes of the research to date, along with recommendations for further work.

ASSESSING SIMULATOR FIDELITY

Fidelity of simulators is traditionally thought of in terms of the physical and functional similarity between the simulator and the air vehicle. A complementary view of simulator fidelity, and the one adopted in this work, is focused on the training benefit derived from its use; the degree to which the simulation imparts the correct behaviours on the pilot in flight. The difference between these views is that a traditionally low-level simulation device, such as a part-task trainer or a desktop trainer, may be considered in the new approach as high fidelity for achieving a particular training requirement. This shift in perspective aims to ensure that operators are not paying for a sophisticated, high-tech device when a lower-tech, more cost effective device will deliver the required training adequately. In addition, this change in focus aims to ensure that 'Level D' full flight simulators (FFS) are not used for training in areas where they are not fit for purpose, under false confidence of their capabilities. ICAO 9625 ed. 3 Volume II [6] has been developed by the IWG and addresses the need for a training-focused fidelity assessment of synthetic training devices and the need to assess fidelity on a task-by-task basis. Secondary IWG groups such as the Motion Task Team (MTT) are being set up to examine the validity of fidelity tolerances, particularly pertaining to the motion and visual systems [14].

Rotorcraft simulation fidelity metrics in JAR-FSTD H are derived from the preceding fixed-wing document (JAR-FSTD A [15]) without an examination of the engineering validity of the tolerances contained within it for use in rotorcraft applications. This lack of clear tolerance guidance, as reported by GARTEUR [5], is one of the drivers for UoL's work into development of new methods to define simulator fidelity in terms of fitness for purpose and transfer of training through the development of perceptual metrics. This work is summarised in the next section; further details can be found in references [16] and [17].

PREDICTED AND PERCEPTUAL FIDELITY

"Lifting Standards: A Novel Approach to the Development of Fidelity Criteria for Rotorcraft Flight Simulators " is an Engineering and Physical Science Research Council (EPSRC) funded project at UoL. The work undertaken in Lifting Standards aims to define a fidelity assessment process that is split into two distinct phases. Firstly, a quantitative analysis of the predicted fidelity of a simulator (the physical and functional fidelity of the individual component systems and flight model) is carried out. Secondly, a simulator fidelity rating scale and new engineering quantitative metrics are being developed to determine the perceptual fidelity of the simulator-pilot system.

ADS-33E PRF Handling Qualities (HQ) metrics[18], such as quickness and bandwidth, have been used to determine the predicted simulator fidelity, with the rationale that a good match in handling qualities has to be part of a good match between flight and simulation. Previous work at UoL [16] has shown that similar handling qualities ratings are necessary but not sufficient to provide high fidelity. Therefore, perceptual metrics such as control attack and cut-off frequency [19] have been used to capture and quantify pilot task strategy variations between flight and simulation. In parallel to this work, the US Army International Technology Center (UK) is funding a project at UoL that includes a comprehensive multi-pilot trial to allow the development of a perceptual Simulation Fidelity Rating (SFR) scale [17] and correlation of ratings with predicted fidelity. The SFR scale developed under this framework has been utilised in the trials detailed within this paper. The handling qualities methodology utilises the Cooper-Harper Handling Qualities Rating (HQR) Scale [20] to rate the handling qualities of an aircraft for a particular task, with Levels 1, 2 and 3 denoting decreasing levels of quality of the aircraft (Level 1= highest quality). In a similar manner, the SFR scale (Figure 1) uses such levels to denote decreasing levels of utility of the simulation (1= Highest Fidelity, Fit for Purpose). In the SFR scale, the pilot rates their task strategy adaptation and relative task performance compared to the flight vehicle or nominal 'baseline' simulation in this study. Further details of the development of the SFR scale are reported in [17].

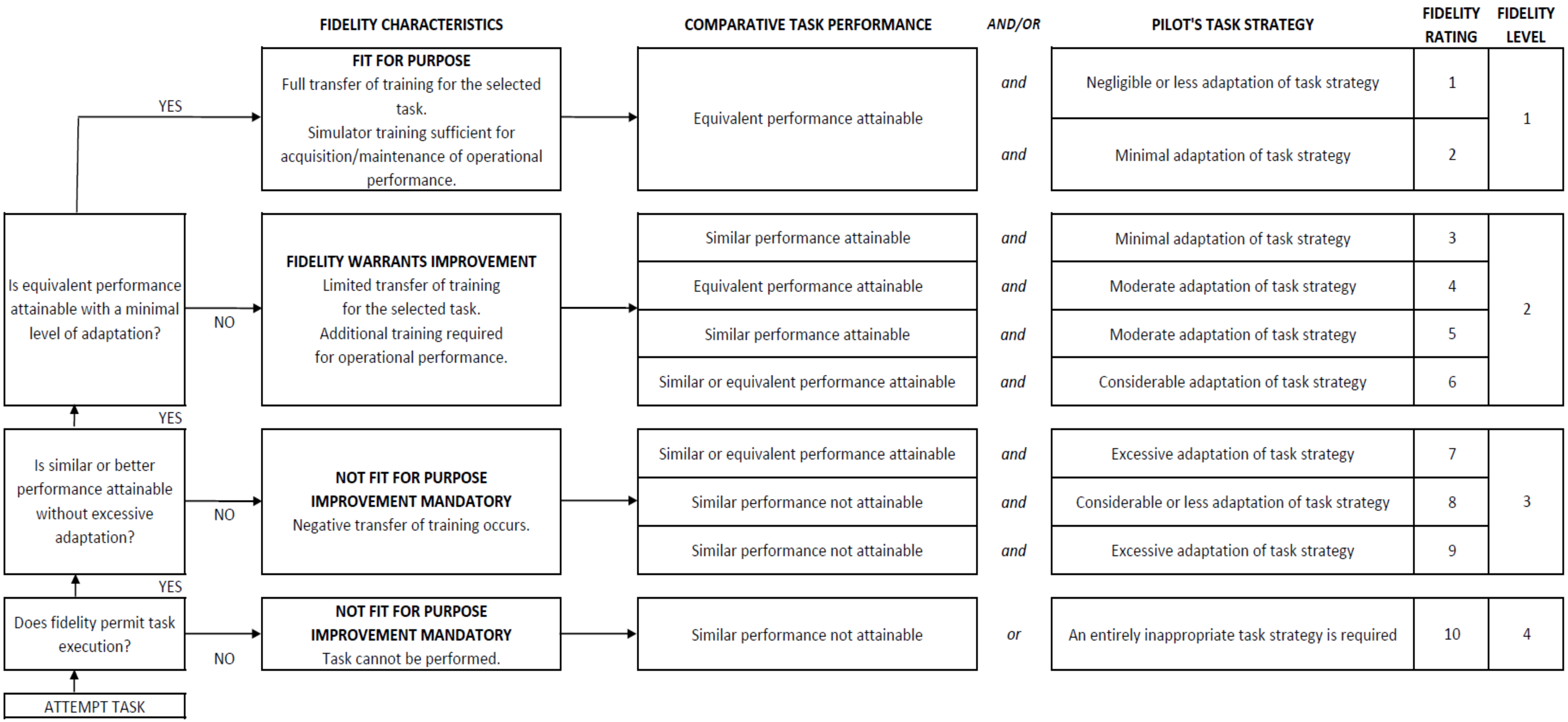

Figure 1 - SFR scale

FACILITIES

The simulator trials were conducted using the UoL's in-house full motion simulator, HELIFLIGHT-R, commissioned in 2008 [12]. The flight trials were conducted in Ottawa, Canada using the NRC's Bell 412 ASRA airborne simulator [13]. This section describes these facilities.

HELIFLIGHT-R

Figure 2 - HELIFLIGHT-R at the University of Liverpool

HELIFLIGHT-R is a reconfigurable, full motion simulator at UoL. Flight models are developed using Advanced Rotorcraft Technology's multi-body flight dynamics modelling environment, FLIGHTLAB [21], and are then loaded into the simulator to be run in real-time.

Features of HELIFLIGHT-R (Figure 2) include a 12 ft visual dome with 210x70 degree Field of View (FoV) and a Moog 6 degree of freedom electric motion system.

The cockpit has an interchangeable crew station with front pilot and co-pilot seats and a rear engineer seat and a reconfigurable instrument panel displays (left and right primary flight displays, backup analogue displays and Head Up Display).

Bell 412 ASRA

Figure 3 - NRC Bell 412 ASRA

The NRC's Bell 412 ASRA was utilised for gathering data to validate the UoL simulation model and to repeat test points from the simulator trials to validate the methodology. The ASRA is configured with on-board research equipment for the development and testing of advanced flight control systems. Its full authority fly-by-wire (FBW) control system provides the aircraft with a variable stability and control capability. Modifications to the FBW system can be quickly and effectively implemented into the ASRA and the aircraft states and control activity data can be recorded for post-flight analysis.

The FBW system contains a number of safety trip points [13] that cause the experimental system used by the evaluation pilot to disengage, at which point control reverts to the safety pilot. The trip limits therefore limit the aggression with which a manoeuvre can be flown. The safety pilot flies the helicopter using the standard mechanical control system, and is responsible for taking control in the event of disengagement, or if a potentially dangerous situation arises [19].

PILOTED TRIALS

The aim of the piloted trials was to determine the effect of flight model variations on transfer of training. To ensure that the effects of flight model variations were isolated from effects from any other simulator characteristics, a nominal baseline simulation was chosen and then modifications were made to the flight model. The results from the modified simulation were then compared to the baseline simulation. Similarly in flight, the Bell 412 ASRA was used to provide the nominal baseline and modifications were made via the FBW system. Removing the influence of the cueing environment on the fidelity allows the method to be repeated in a number of STDs and in flight in order to validate the robustness of the SFR scale. Four test pilots participated in the simulator trials, two of these in the complementary flight trials.

The baseline simulation represents a training simulator used to train a pilot to fly a manoeuvre. Each pilot completed at least four runs of the manoeuvre in the baseline configuration to ensure that they were able to achieve consistent task performance. At this point a HQR for the baseline configuration was awarded by the pilot. Immediately following this training period, a modification was made to the flight model and the pilot repeated the manoeuvre. The pilot was then informed that a 'modification had been made to the simulation' but was not told the nature of the change, whether it be to the flight model, motion/visual system, control feel etc. The intent was to determine to what extent the training in the baseline (representing the STD) prepared the pilot for performing the manoeuvre in the modified simulation (representing the flight vehicle). The pilots used the SFR scale (Figure 1) to rate the fitness for purpose of the baseline configuration, as a training device for the modified configuration, following the first run in the modified simulation. The pilot was then asked to complete three more runs in the modified configuration to allow assessment of adaptation and training. An HQR rating of the modified simulation was awarded by the pilot on the fourth run. Data from the simulation trials were analysed to compare performance and pilot strategy between the simulations. The modifications made are described below and the test matrix is shown in Table 1. This paper focuses on the results from the inter-axis couplings experiments.

Table 1 – Number of completed test points

	Flight		Simulation	
	Precision Hover	Accel-Decel	Precision Hover	Accel-Decel
Baseline	2	2	4	4
10% pitch/roll coupling	0	0	1	0
20% pitch/roll coupling	0	1	2	3
40% pitch/roll coupling	0	2	0	1
50% pitch/roll coupling	2	0	1	3
100% pitch/roll coupling	2	0	4	2
100ms transport delay	2	2	4	4
200ms transport delay	0	0	2	3
300ms transport delay	0	0	3	3
Bare airframe	0	0	4	3

Two ADS-33E-PRF [18] mission task elements (MTEs) were chosen for the study; the Acceleration-Deceleration (Accel-Decel) and the Precision Hover.

The Precision Hover (PH) course is shown in Figure 4, and the performance criteria are detailed in Table 2. The manoeuvre is initiated with the aircraft travelling at a ground speed of between 6 and 10 knots, at an altitude of less than 20 feet. The target hover point is to be oriented at 45° relative to the heading of the rotorcraft. The ground track should be such that the rotorcraft will arrive over the target hover point. The hover should be captured in one smooth manoeuvre following the initiation of deceleration – "it is not acceptable to accomplish most of the deceleration well before the hover point and then to 'creep up' to the final position" [18].

The relationship between the heights of the pole and the hover board is such that, when over the target hover point and aligned with both the marker on the pole and the hover board, the rotorcraft will be at the reference height of 10 feet.

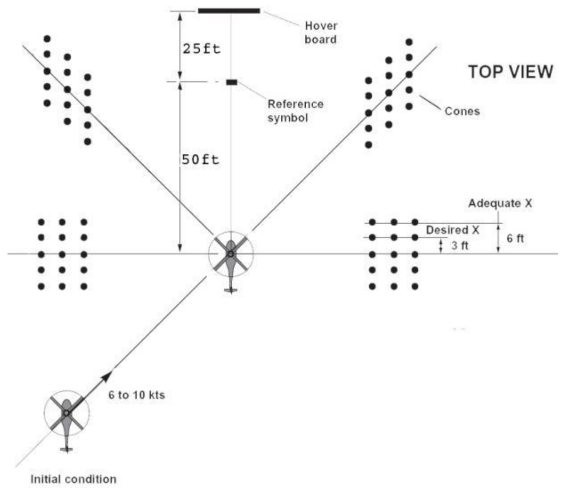

Figure 4 - Precision Hover Course

Table 2 – Performance Criteria for the Precision Hover MTE

Criteria	Desired Performance	Adequate Performance
Attain stabilised hover within X seconds of initiation of deceleration	5	8
Maintain a stabilised hover for at least X seconds	30	30
Maintain the longitudinal and lateral position within ±X feet on the ground	3	6
Maintain altitude within ±X feet	2	4
Maintain heading within ±X °	5	10
There shall be no objectionable oscillations during the transition to hover or during the stabilised hover	Applies	Does not Apply

The Accel-Decel (AD) manoeuvre profile is shown in Figure 5 and the performance criteria are detailed in Table 3. The manoeuvre is started from a stabilised hover. To initiate the MTE, the pilot should rapidly increase power to approximately 95% of the maximum continuous power, maintaining altitude constant using pitch attitude, and hold collective constant during the acceleration to an air speed of 40 knots (Relaxed from the ADS-33E-PRF requirement of 50knots due to space constraints on the NRC flight test course). Upon reaching the target airspeed, the pilot should initiate a deceleration by aggressively reducing the power and holding altitude constant. The peak nose-up attitude should occur just before reaching the final stabilised hover. The test course is shown in Figure 5. The desired track is indicated by a series of cones along the centre-line of the course, with markers to the left and right indicating the boundaries of desired and adequate lateral tracking performance.

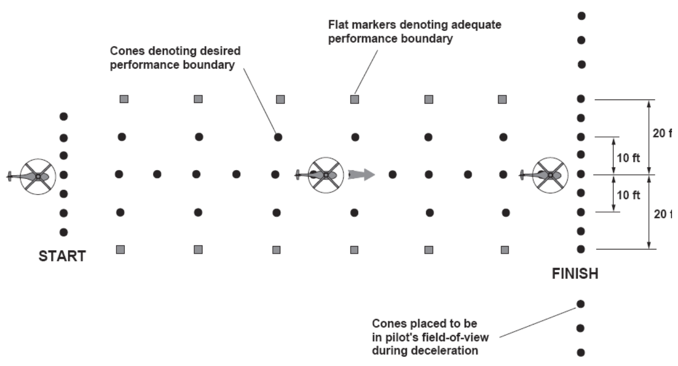

Figure 5 – Accel-Decel Course

Table 3 - Performance Criteria for the Accel-Decel MTE

	Desired Performance	Adequate Performance
*Within X seconds from initiation of the manoeuvre, achieve at least the greater of 95% maximum continuous power or 95% maximum transient limit that can be sustained for the required acceleration, whichever is greater.	1.5	3
Maintain altitude below X feet	50	70
Maintain lateral track within ±X feet	10	20
Maintain heading within ±X °	10	20
Significant increases in power are not allowed until just before the final stabilised hover	Applies	Applies
*Achieve a nose-up pitch attitude during the deceleration of at least X ° above the hover attitude. The maximum nose-up attitude should occur shortly before the hover	30	10
Longitudinal tolerance on the final hover point is ±X feet	21	42
Rotor RPM shall remain within the limits of X without undue pilot compensation	OFE	SFE

* May be relaxed due to ASRA FBW safety trip limits

The trials were undertaken in three phases; first, an exploratory trial with a single pilot to aid calibration and initial validation of the SFR scale and determination of sensitivity to flight model modifications [17]. The second phase was a simulation trial with multiple pilots of varying degrees of experience. The third phase was a flight trial with two pilots in the NRC Bell 412 ASRA airborne simulator to support validation of the fidelity methodology. For the simulator trials, the baseline or 'truth' model was the FLIGHTLAB Bell 412 ASRA model [16] configured with a Rate-Command-Attitude-Hold (RCAH) system. For the flight trial, the baseline configuration was the Bell 412 ASRA with the same Rate Command Attitude Hold (RCAH) configuration as implemented in the FLIGHTLAB model (Although slightly different gearings were implemented in flight).

The flight model modifications were chosen such that they should be quantifiable, predictable and repeatable and create, as far as possible, a single handling qualities variation. The modifications were an incremental increase in cross coupling between the pitch and roll axes (to degrade the inter-axis coupling HQ only) and an incremental increase in simulator transport delay [3], to degrade the bandwidth/phase delay HQs only. There is a need to understand how well training pilots in an augmented aircraft prepares them for an emergency situation where some, or all, of the augmentation may be lost. A further scenario that was investigated was the transfer of training from the augmented RCAH configuration to the un-augmented, or 'bare airframe' configuration. According to the ADS-33E-PRF predicted handling qualities, the augmented RCAH configuration was Level 1, while the 'bare airframe' was Level 3 [22].

The modifications to the simulation model were implemented via the Control System Graphical Editor (CSGE) user interface in FLIGHTLAB. To implement the cross coupling, the on-axis input was fed into the off-axis input as a percentage of the on-axis input in both lateral and longitudinal cyclic (XA and XB). For example, a 2 inch lateral stick input (on-axis) resulting in a 1 inch longitudinal stick input, was referred to as a '50% cross coupling' in the current study. The transport delay was implemented using a delay block in the control system, adding a defined number of time steps onto the pilot's control inputs. A single time step is 7.8ms, hence 13 time steps will be 101.4ms ~ 100ms transport delay. The implementations of these modifications into CSGE are shown in Figure 6 and Figure 7.

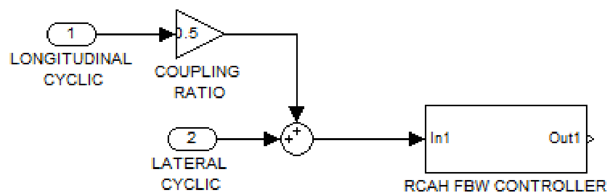

Figure 6 - Implementation of inter-axis coupling into CSGE

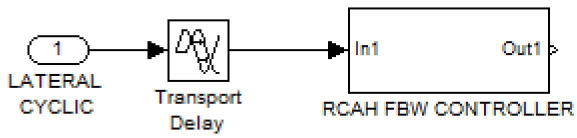

Figure 7 - Implementation of transport delay into CSGE

PILOT VARIABILITY

In a simulation trial there are often pilot related influencing factors that must be considered during analysis of the results. Pilot variability from run to run provides a challenge in identifying effects of the flight model modifications. Also, with experience, pilots become more familiar and hence more proficient with the configuration flown. Clearly, the manner in which each pilot flies, e.g. the level of aggressiveness, will impact their subjective ratings. Figure 8 illustrates aspects of these pilot variability issues; the manoeuvre is the Accel-Decel and the configuration is the baseline, generally flown in groups of 4 runs. The control metric is the average control displacement from trim derived from the control attack computations and variations in this parameter reflect changes in pilot strategy. Pilot A begins to converge on a strategy for the AD MTE after the first 4 runs, with decreasing displacements as he adapts to the simulator. Pilot B uses larger control displacements (and generally more rapid) and appears to be much more variable in his control strategy. Pilot C is very consistent across 24 runs using smaller control inputs. Pilot D shows a similar trend to pilot A, with generally decreasing movements of the longitudinal stick with increased exposure to the baseline.

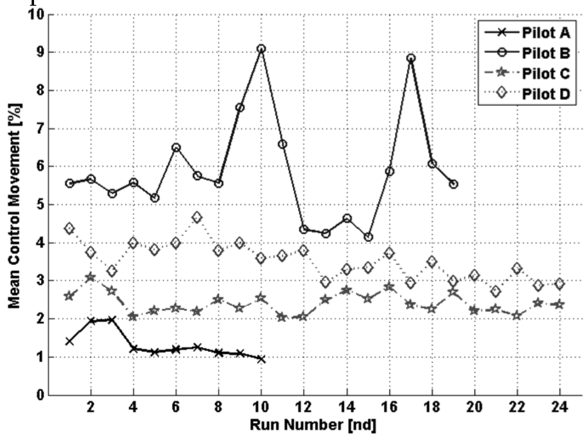

Figure 8 - Longitudinal cyclic (XB, %)) average 'attack' displacements - baseline simulation, AD MTE

The intra-pilot variability should be small for a consistent subjective impression and associated perceptual metrics. Three of the four pilots above fit into this category. The extent of inter-pilot variability shown in Figure 8 is larger than the intra-pilot variability. This is also not unusual in HQ studies, reflecting the different piloting strategies possible in terms of the response to the visual and vestibular motion cueing. So-called 'high-gain' pilots generally use larger and more rapid control inputs to achieve a similar performance to a pilot making slower and smaller inputs. In this study we have some evidence that higher-gain pilots are more sensitive to (flight model) fidelity changes, perhaps as a consequence of the increased exercising of the flight model. This is picked up later in the paper.

THE EFFECTS OF CROSS COUPLING

Although additional flight model modifications have been studied, this paper focuses on the inter/cross coupling test results.

Increasing the pitch/roll cross coupling primarily affects the off-axis response following longitudinal and lateral cyclic inputs. Figure 9 shows the modified configurations on the ADS-33E-PRF inter-axis coupling HQ requirement chart [18] the coupling modifications are seen to affect the handling qualities of the aircraft in a fairly uniform manner. Throughout this paper the 'cross coupling' is the couple introduced and the inter-axis couple is the coupling observed in the output as a result. A 12.5% modification to the model in roll from pitch cross coupling degrades the handling qualities into Level 2, suggesting that, beyond this, the fidelity of the simulation may be compromised.

JAR-FSTD-H requires that, following a longitudinal step input, the on-axis response - the pitch rate (and pitch attitude) - should be within the tolerances shown as the broken lines in Figure 10, and the off-axis response should be of 'correct trend and magnitude' or CT&M. The on-axis response tolerances in JAR-FSTD H are either ±10% of the achieved peak, or ±3 deg/sec, whichever is less restrictive. This choice is seen in Figure 9 to favour the absolute metric, and to a significant degree. This particular metric tolerance is under investigation in a parallel study and results will be reported in a future publication. The off-axis responses all exhibit the correct trend so it becomes a matter of interpretation whether the magnitudes are 'correct'. The ambiguity of the quantitative metrics illustrated by these results is considered to be one of the shortcomings of JAR-FSTD H and is under investigation in the current research.

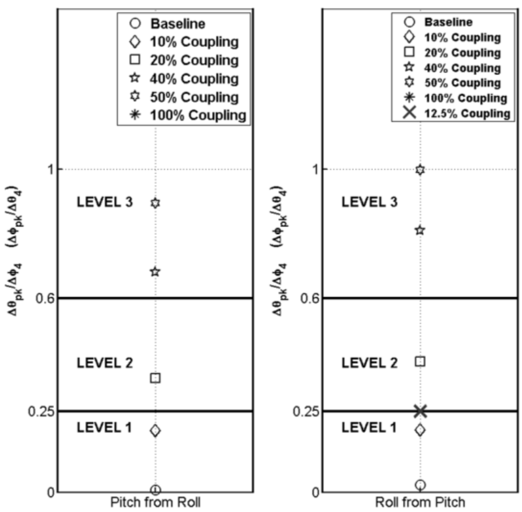

Figure 9 - ADS-33E-PRFPitch/Roll Coupling Requirements for Aggressive Agility

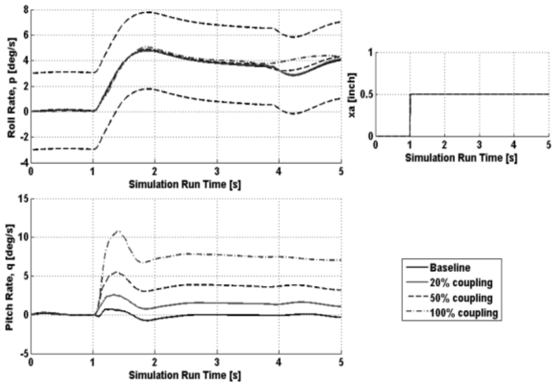

Figure 10 – On and off-axis rate responses to a 0.5 inch, 4s lateral cyclic input for cross coupling variations

Figure 11 and Figure 12 show the HQRs awarded during the trials, as a function of model cross coupling strength; mean and range of ratings are shown. Figure 11 shows the spread of baseline HQRs for the precision hover manoeuvre to be between HQR=2 and HQR=5. This includes nine rated baseline runs (each corresponding to a subsequent coupling test point). Seven of the nine baseline HQRs were either HQR=2 or HQR=3 and the remaining two results were in Level 2 at HQR=4 (Pilot B baseline for the 100% cross couple) and 5 (Pilot D baseline for 50% couple). Performance analysis confirms that Pilot D was unable to achieve desired performance in the baseline for the run rated HQR=5, but was able to achieve desired performance in the baseline on other occasions (HQRs 2 and 3). This may be attributed to pilot fatigue as this test point was the last of the day for pilot D. In general the baseline was considered to have Level 1 HQs for the PH MTE.

On the SFR scale (Fig. 1), the boundary between Level 1 and Level 2 fidelity lies between SFR 2 and 3. The boundary between Level 2 and Level 3 fidelity lies between SFR 6 and 7 [17]. The results in Figure 13 and Figure 14 show that the SFRs degrade as coupling strength increases, as expected.

Figure 9 shows that for Level 1 HQs the roll inter-axis coupling metric for moderate amplitude manoeuvres should be less than 0.25; the baseline model is predicted to be Level 1. However, three of the four pilots awarded Level 2 HQRs for the baseline in the Accel-Decel. Pilots noted that the fly-by-wire pitch rate trip limits of the aircraft were often breached during the deceleration phase of the MTE, and they found it difficult to achieve consistent performance due to lack of cueing, particularly in the vertical and lateral axes, at hover capture. This may be the cause of the discrepancies between the predicted and assigned HQRs. Figure 9 also shows that for additional cross couples of 12.5% and above, there would be a change in predicted handling qualities, suggesting a compromise in fidelity. This might then mean that differences in HQs at the 20% cross couple test points would be so noticeable to a pilot that a Level 2 SFR would be returned. However, two of the three test pilots who flew this test point for the AD MTE awarded Level 1 fidelity (Figure 14) and all three pilots awarded SFR 1 for the PH MTE (Figure 13). It is concluded that while the handling qualities metrics approach are suitable for fidelity assessment, further investigation into the positioning of the SFR boundaries is required.

At the heart of the present research is the motivation to provide evidence for defining new metrics and tolerances for simulation fidelity. If a trend can be identified between incremental parameter variation (e.g. cross coupling) and SFR variation then the point where these trend lines cross the SFR Level boundaries can be used as a basis for quantitative fidelity tolerances. From the results presented, it is expected that the tolerance on pitch/roll cross coupling for Level 1 fidelity for the Precision Hover task may lie between 30% and 40% cross coupling (an interaxis coupling of approximately 0.7 - see Figure 9), but this value would be smaller for the Accel-Decel at approximately 20% cross coupling (inter-axis coupling of approximately 0.4). This result illustrates that fidelity is task dependant, which is not currently considered in the JAR-FSTD H standards.

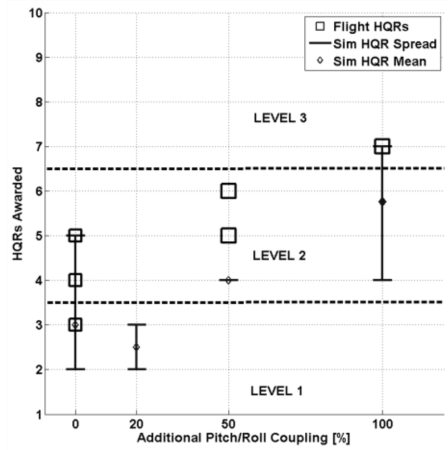

Figure 11 – HQRs awarded for cross coupling tests HELIFLIGHT-R and ASRA –Precision Hover MTE

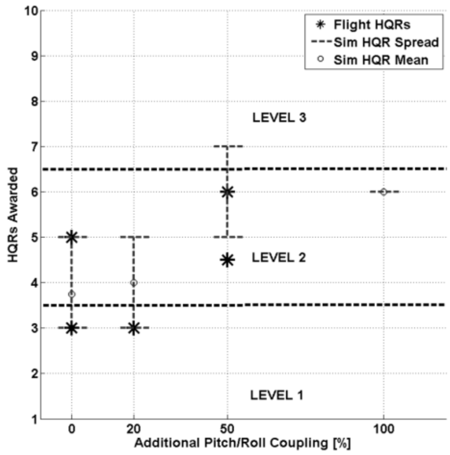

Figure 12– HQRs awarded for cross coupling tests HELIFLIGHT-R and ASRA – Accel-Decel MTE

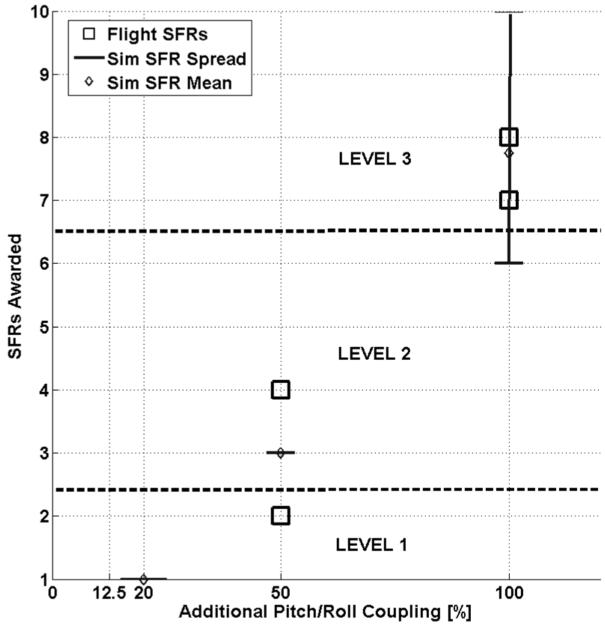

Figure 13 - SFRs awarded for cross coupling tests HELIFLIGHT-R and ASRA –Precision Hover MTE

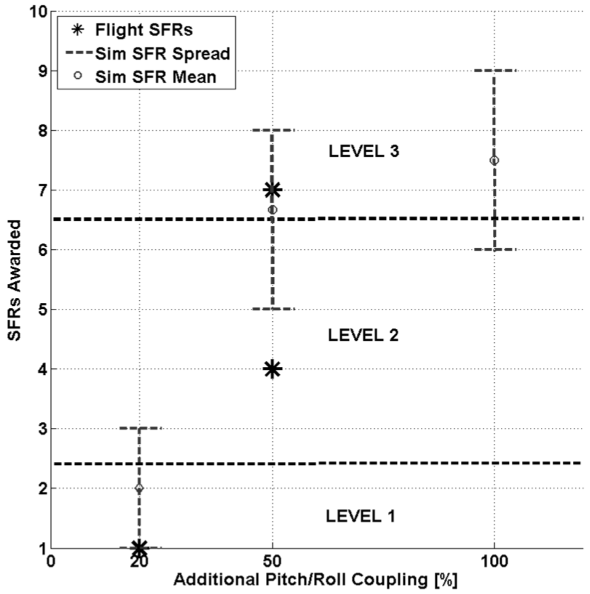

Figure 14 - SFRs awarded for cross coupling tests HELIFLIGHT-R and ASRA – Accel-Decel MTE

Analysis of pilot control strategy through perceptual metrics such as control attack and stick displacement allow for a quantitative substantiation of the subjective ratings in the assessment of transfer of training. A number of metrics have been explored at UoL to capture pilot control activity. However a single control activity metric does not necessarily describe all aspects of task adaptation. When the pilot rates adaptation he is accumulating information reflected by several metrics. For example, Figure 15 shows the accumulation of the % change of the control activity metrics between baseline and modifications against the level of reported adaptation for all test points. The PH results do not show as clear a picture as the AD results. It is hypothesised that because the AD is largely a single axis manoeuvre it is easier for the pilot to deduce their level of adaptation due to inter-axis coupling. However, research into this aspect is ongoing. The Pilot D 20% cross coupling PH case appears anomalous. The control activity (Figure 16) for the 20% cross coupling case is significantly lower than for the baseline. It is hypothesised that the pilot is moving the stick in the direction of translation making a more intuitive control strategy; the inter-axis coupling actually being an advantage. A reduction in workload is reflected in the HQR reducing from a 3 to a 2

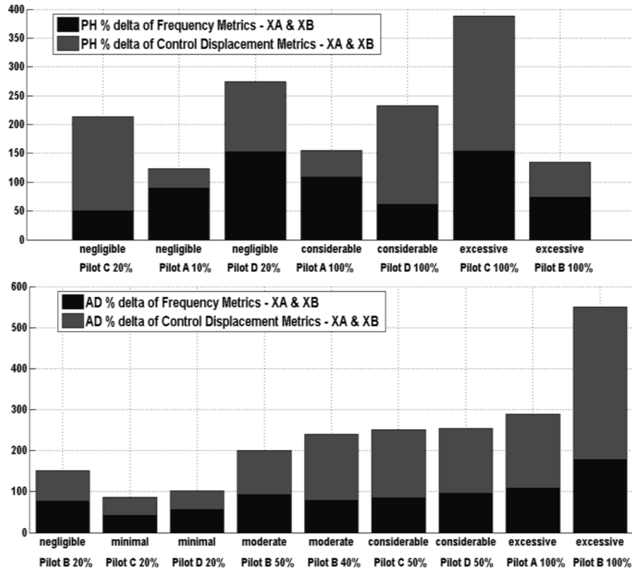

Figure 15 - Relationship between level of adaptation awarded and change in pilot control activity metrics

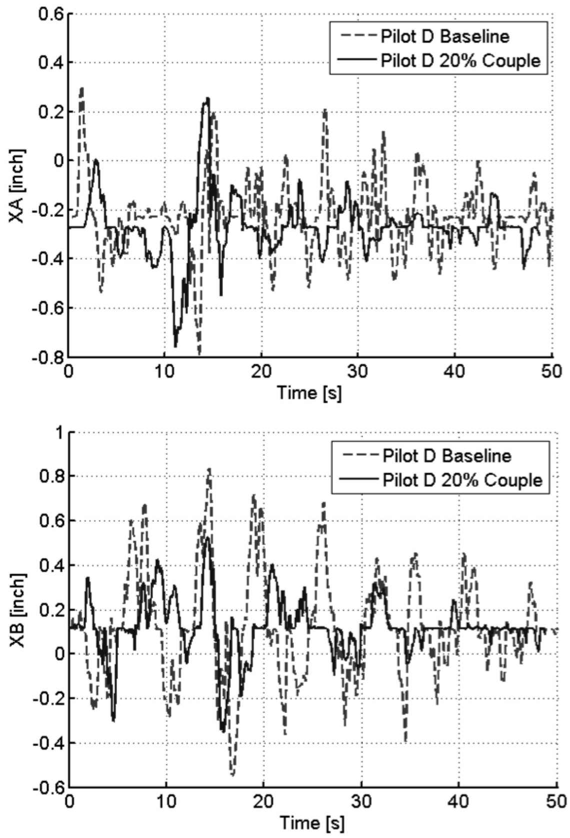

Figure 16 - Comparative control strategy between baseline and 20% cross couple, Pilot D, PH MTE

The methodology has only been used to rate against a nominally Level 1 handling qualities baseline aircraft, with modifications that degrade the handling qualities of the aircraft. It is also intended to carry out studies where the handling qualities of the baseline simulation ('synthetic trainer') are worse than the modified simulation ('aircraft'), an equally realistic scenario. It could be argued that in this case, the transfer of training would be higher; as if a pilot learns to fly in a more difficult vehicle they should have less difficulty flying the real aircraft. This will also include examination of whether the pilot's sensitivity to flight model variations is altered with increased baseline workload.

To support validation of the methodology used in the current study, several test points were repeated in the ASRA airborne simulator by two of the test pilots. The results are overlaid on the ground-based simulator results in Figure 11 to Figure 14. Generally, the SFR awarded in the ASRA was in the same fidelity level as the SFR awarded in HELIFLIGHT-R. The differences are more pronounced in the Accel-Decel compared with the Precision Hover.

SFR sensitivity to the nature of the task being flown is accompanied by pilot sensitivity to fidelity. Different pilots fly tasks in subtly, or in some cases significantly, different ways, depending on proficiency and flying experience. This will impact the pilot’s fidelity ratings as more active, or higher gain, pilots are likely to excite more of the aircraft's dynamics, thereby exposing more fidelity issues. Generally, pilots that were more active on the controls gave a poorer SFR. For the 100% cross coupling case in the PH MTE the spread in SFRs for the four pilots is ranged from SFR=6 to SFR=10. Looking particularly at the two extremes of the range, pilot A (SFR=6) and pilot B (SFR=10), a number of deductions can be made. An SFR 6 corresponds to considerable adaptation and an SFR 10 corresponds to an inability to perform the task; however on the second run a full set of data was recorded and the adaptation was noted to be excessive. Inspection of Figure 15 shows that although Pilot B awarded a higher level of adaptation, the cumulative percentage change in metrics compared to the baseline run was less than that of pilot A, who reported only considerable adaptation. This may be because pilot B is more active in terms of cyclic displacement in the baseline (see Figure 17) and therefore has a higher workload to begin with; his adaptation then pushes him toward his workload capacity, causing the adaptation to feel more severe. For this reason, the guidance notes that accompany the scale must help the evaluating pilots understanding of adaptation; clear briefing and de-briefing protocols (supported by perceptual metrics) should also aid the assessment.

Considering all four pilots once more, if the SFRs for the 100% cross coupling PH test point are normalised by either the on-axis or off-axis control attack rate (the mean rate at which the pilot is moving the control), the normalised results tend to a similar value (Figure 18). As previously suggested, it is hypothesised that pilots who use a more active control strategy excite the dynamic modes of the aircraft more readily and are therefore more likely to be sensitive to flight model variations. This highlights a significant question - is a particular type of pilot (high/low gain) more suitable for fidelity assessment? A low gain pilot may not excite the aircraft dynamics sufficiently to expose deficiencies and a high gain pilot may be too highly variable. Instead of restricting the pilots that are suitable for fidelity assessment, could a calibration test allow the engineers to ‘normalise’ for the pilot’s control strategy? The development and practicality of such a test is part of ongoing research.

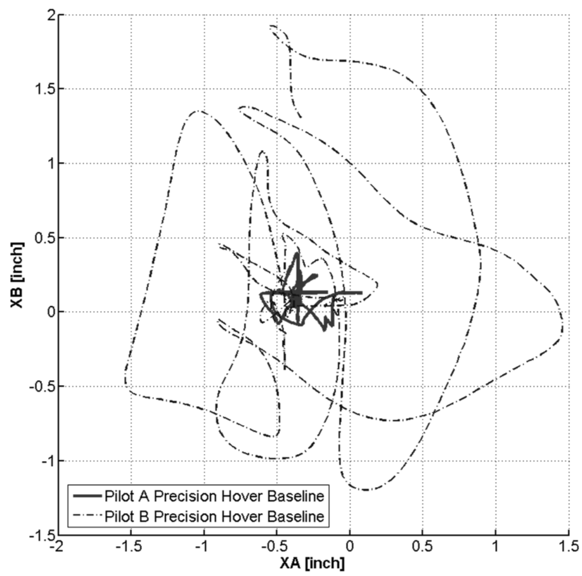

Figure 17 - Inspection of differing baseline strategies - Pilots A and B, PH MTE

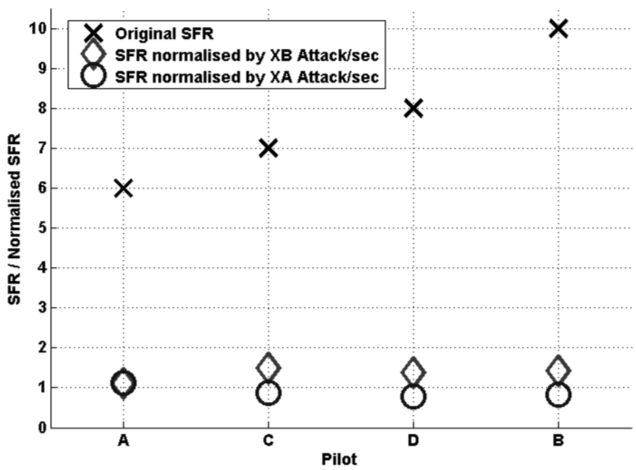

Figure 18 - Normalising SFRs – 100% cross coupling, Precision Hover MTE

DISCUSSION

The data presented in the results section above allow an initial inspection of how subjective fidelity assessment correlates with predictive metrics to facilitate exploration of fidelity tolerance boundaries.. However, a large influencing factor in this work has been the variation from one pilot to another. For the sake of describing the way in which the methodology utilised in this study could work to justify quantitative fidelity tolerances, one pilot (Pilot B) is considered: Figure 19 shows that the HQRs awarded by pilot B for the baseline, 20% and 50% cross couplings for the AD MTE, agree with the HQ coupling criteria for moderate amplitude aggressive manoeuvring. These results suggest that the Level 1 HQ border will be crossed as predicted. The SFR plot however shows a much higher tolerance to coupling suggesting the boundary for Level 1 fidelity may be crossed around 20-30% cross coupling (in agreement with the 4 pilot study). As for the Level 2-3 fidelity boundary, both the SFR and HQR plots show a cross into level 3 between 40% and 50% cross coupling suggesting there may be a ‘cliff-edge’ in pilot tolerance around this point. Note that between 20% and 40% additional cross coupling there is no change in HQR but the SFR degrades from Level 1 to Level 2. This result further highlights that similar HQRs are insufficient for high fidelity.

The lateral cyclic control activity metrics (Figure 20), particularly the % change from baseline of the average displacements, increase both with SFR and level of adaptation awarded. This suggests that the change in off-axis control activity may be a suitable metric for assessing inter-axis coupling fidelity. This would be expected given that larger off-axis control inputs must be required to suppress increased couplings.

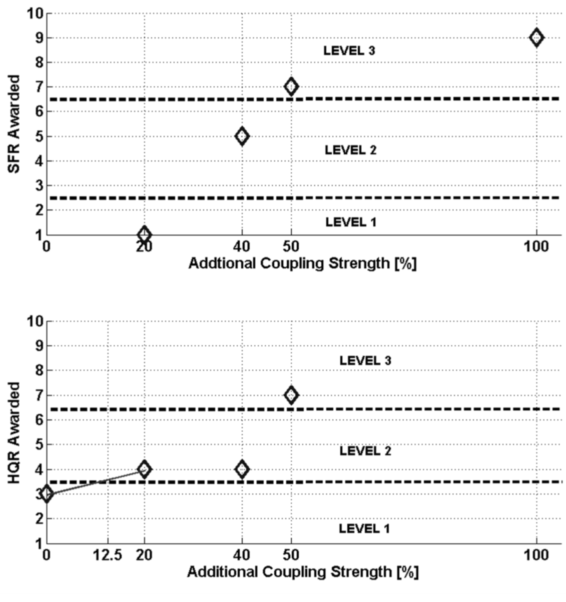

Figure 19 - Pilot B HQRs and SFRs for increasing cross coupling

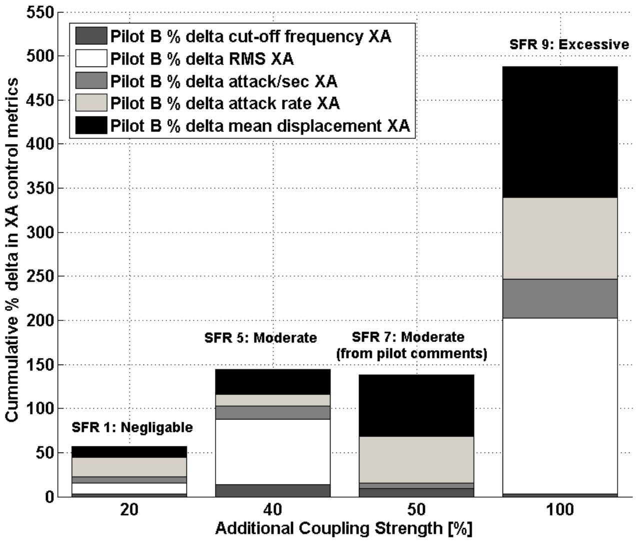

Figure 20 - Breakdown of combined control activity metrics against awarded level of adaptation - Pilot B

This study of two manoeuvres (the AD MTE and the PH MTE) has shown that pilot sensitivity to flight model variations is task dependent. The AD and PH MTE manoeuvres are well defined, handling qualities assessment manoeuvres. These manoeuvres were chosen for this initial study due to their simplicity and well defined structures. However, for ‘real world’ application, the methodology will be utilised to assess more complex training tasks. New manoeuvres allowing the assessment of synthetic training device fidelity are currently being developed at UoL. The specification is that they should be realistic training tasks, be quantifiable in terms of performance requirements, and be repeatable to allow for gathering of statistically significant results. Two manoeuvres to be assessed will be the Turn-around-Tail and a level turn with quick-stop into wind. The long term aim of this research is that the SFR scale can be utilised for all 70 training tasks outlined in the ICAO 9625 volume II ed. 3 [6] to provide a means of assessing simulator fidelity on a task by task basis.

The methodology has received a degree of validation in flight tests at the NRC where similar trends and sensitivities to fidelity have been found. The scale is considered sufficiently sensitive to capture pilot adaption to changes in flight model. The effects of cueing environment and control strategy on perceptual fidelity are equally as important as effects of the flight model. Continuing research will explore these issues further.

CONCLUDING REMARKS

This paper has reported results from research into the development of a quantification and assessment methodology for simulation fidelity. In particular, a new Simulator Fidelity Rating (SFR) scale has been used to assess the level of fitness-for-purpose of a synthetic training device. It is proposed that such a method could be used to complement the current functional fidelity assessment of simulators as well as provide evidence and justification for quantitative fidelity tolerances.

The paper has detailed simulator trials where one simulation (the trainer) is compared to another (the real aircraft) using the HELIFLIGHT-R facility at UoL, to isolate effects of the mathematical flight model on transfer of training. Initial results from piloted trials have been presented to highlight some influencing factors, emerging trends and the relationship between quantitative metrics and subjective ratings given by evaluation pilots.

Pilot variability, both from one to another and from run to run, is an effect that must be considered during the piloted trials. A calibration test is to be explored to determine piloting strategy that might allow for weighting or normalisation of SFRs.

The way in which pilots perceive their adaptation has a significant effect on SFRs, of course, and can be the difference between a rating that deems the simulator fit for purpose or not. The current study has highlighted the need of clear guidance and training as well as briefing and debriefing protocols, alongside quantitative data gathered in real time to support the effective use of the scale.

The results show that subjective fidelity is task dependant; in this study the pilots were more sensitive to pitch/roll cross coupling errors in the more aggressive manoeuvre, the Accel-Decel. While the moderate amplitude coupling HQ metric may be suitable for fidelity assessment, further investigation is required to define where the boundaries should lie as the awarded SFRs suggest a higher tolerance to cross coupling than the HQ metric.

The current methodology looks at forward transfer of training; assessing how flight agrees with initial exposure in the simulator. It is also planned to use the scale in reverse, for backward transfer of training. As mentioned in the Introduction, pilots that have experienced in emergency situations in the aircraft are able to comment on the fidelity of the simulation in that flight regime; using the SFR scale would formalise this process. However, the absence of real world validation data for emergency situations will always limit the scope of training.

ACKNOWLEDGEMENTS

The research reported in this paper is funded by the UK EPSRC through grant EP/G002932/1 and the US Army International Technology Center (UK) (reference W911NF-11-1-0002). The authors would like to thank all of the test pilots involved in the work so far, without whom this work would not be possible, especially Andrew Berryman for his continuing support in this work. The contributions of test pilots Lt Cdr Lee Evans (RN), Lt Cdr John Holder (RN), Lt Christopher Knowles (RN) and Flt Lt Russ Cripps (RAF) of the UK Rotary Wing Test and Evaluation Squadron are also gratefully acknowledged The use of the Bell 412 ASRA facility is gratefully acknowledged, along with the support of FRL at the NRC, Ottawa, particularly Stephan Carignan. The authors would also like to thank all who attended the 'Simulation Fidelity Workshop' (http:/www.flightlab.liv.ac.uk /fidelity) at the 67th AHS forum in May 2011; comments and suggestions made at this event have contributed greatly to the research effort. The partnership with Advanced Rotorcraft Technology Inc. in their continued support in the development of the simulation facility. The acquisition of the HELIFLIGHT-R was made possible through support from the University of Liverpool.

REFERENCES

[1] Ebbutt, G., “Jane’s Simulation and Training Systems”, IHS Jane’s, ISBN 9780710629388, 2010.

[2] Anon., “The Impact of Flight Simulation on Aerospace”, Royal Aeronautical Society Flight Simulation Group, http:/www.raes-fsg.org.uk/30/Flight_Simulation_Group_Specialist_Paper, 2008.

[3] Anon., “JAR-FSTD H, Helicopter Flight Simulation Training Devices”, Joint Aviation Authority, 2008.

[4] Anon., “AC 120-63, Helicopter Simulation Qualification”, Federal Aviation Authority, 1994.

[5] Padfield, G.D., Casolaro, D., Hamers, M., Pavel, M., Roth, G. and Taghizad, A., “Validation Criteria for Helicopter Real-Time Simulation Models: Sketches from the Work of GARTEUR HC-AG12”, European Rotorcraft Forum, 2004.

[6] Anon. “Manual of Criteria for the Qualification of Flight Simulation Devices, Volume II – Helicopters”, International Civil Aviation Organization , ICAO 9625 Third Edition - Draft. 2011.

[7] Weaver, Sallie J. et al., “Simulation-based team training at the sharp end: A qualitative study of simulation-based team training design, implementation, and evaluation in healthcare”, Journal of Emergencies, Trauma and Shock, pp. 369-377, 2010.

[8] Holman, G., L., “Training Effectiveness of the CH-47 Flight Simulator”, Research Report 1209, U. S. Army, 1979.

[9] Mitchell, D. G., Roger H. Hoh, He, C., and Strope, K., "Determination of Maximum Unnoticeable Added Dynamics," Atmospheric Flight mechanics Conference, AIAA 2006-6492, Keystone, Colorado, USA, 21st - 23rd August 2006.

[10] Burki-Cohen, J. S., "The Effect of Simulator Motion Cues on Initial Training of Airline Pilots", American Institute of Aeronautics Modelling and Simulation Technologies Conference, San Francisco, CA, USA, 15-18th August, 2005.

[11] Sparko, A. L., Burki-Cohen, J. S., "Transfer of Training From a Full Flight Simulator vs. A High Level Flight Training Device with a Dynamic Seat", American Institute of Aeronautics Modelling and Simulation Technologies (MST) and Global Navigation and Control (GNC) Conferences, Toronto, Canada, 2nd - 5th August, 2010.

[12] White, M. D., Perfect, P., Padfield, G.D., Gubbels, A. W. and Berryman, A. C., "Acceptance testing of a rotorcraft flight simulator for research and teaching: the importance of unified metrics." Proceedings of the 35th European Rotorcraft Forum, Hamburg, Germany, September 2009.

[13] Gubbels, A.W., Ellis, D.K., "NRC Bell 412 ASRA FBW Systems Description in ATA100 Format", Institute for Aerospace Research, National Research Council Canada, Report LTR-FR-163, 2000.

[14] Anon., "Organisational and Financial Requirements for Finalising, Testing and Validating an Objective Motion Cueing Fidelity Criterion with Project Milestones" RAeS FSG, 2010.

[15] Anon., "JAR-FSTD-A - Aeroplane Flight Simulation Training Devices", Joint Aviation Authority, 2008.

[16] White, M. D., Perfect, P., Padfield, G., "Progress in the Development of Unified Fidelity metrics for Rotorcraft Flight Simulators", 66th American Helicopter Society Forum, Phoenix, Arizona, US, 11-13 May 2010.

[17] Perfect, P., Timson, E., White, M. D., Padfield. G., "A Rating Scale for Subjective Assessment of Simulator Fidelity", 37th European Rotocraft Forum, Milan Italy, 13 -15th May, 2011.

[18] Anon. "ADS-33E-PRF, Handling Qualities Requirements for Military Rotorcraft", U. S. Army, 2000.

[19] Perfect, P., Gubbels, A. W., Berryman, A., "Integrating Predicted and Perceived Fidelity for Flight Simulators", 36th European Rotorcraft Forum, Paris, France, 7-9th September, 2010.

[20] Cooper, G. E. and Harper, Jr. R. P., "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities", NASA Technical Note, TN-D-5153, April 1969.

[21] DuVal, R.W., "A Real-Time Multi-Body Dynamics Architecture for Rotorcraft Simulation", The Challenge of Realistic Rotorcraft Simulation, RAeS Conference, London, November 2001.

[22] Perfect, P. "ADS-33 Predictive Handling Qualities – FLIGHTLAB Model with NRC Control Augmentation", FS&T Report FST-LS-REP007, University of Liverpool, 2011