Methods to Assess the Handling Qualities Requirements for Personal Aerial Vehicles

Philip Perfect[1], Michael Jump[2] and Mark D. White[3]
The University of Liverpool, Liverpool, United Kingdom

This paper describes the development of a methodology to assess the handling qualities requirements for Personal Aerial Vehicles (PAVs). It is anticipated that a PAV would be flown by a 'flight-naïve' pilot who has received less training than is typically received by today's general aviation private pilots. The methodology used to determine handling requirements for a PAV cannot therefore be based entirely on existing best practice – the use of highly experienced test pilots in a conventional handling assessment limits the degree to which results apply to the flight-naïve pilot. This paper describes an alternative set of methods based on both the subjective and objective analysis of performance and workload of flight-naïve pilots in typical PAV tasks. A highly reconfigurable generic flight dynamics simulation model that has been used to validate the methodology is also described. Results that highlight the efficacy of the various methods are presented and their suitability for use with flight naïve pilots demonstrated.

Nomenclature

δ_{lat} = Lateral stick input [-1:1]

δ_{lon} = Longitudinal stick input [-1:1]

$\delta_{lon,gust}$ = Longitudinal stick input component from gust [-1:1]

$\delta_{lon,pilot}$ = Longitudinal stick input component from pilot [-1:1]

$\delta_{lon,total}$ = Longitudinal stick input from gust and pilot combined [-1:1]

ϕ_{cmd} = Commanded Bank Angle [rad]

ω_{lat} = Roll Response Natural Frequency [rad/s]

[1] Research Associate, Centre for Engineering Dynamics, School of Engineering
[2] Lecturer, Centre for Engineering Dynamics, School of Engineering, Member AIAA
[3] Lecturer, Centre for Engineering Dynamics, School of Engineering

ζ_{lat} = Roll Response Damping Ratio [-]

A = Aptitude Test Score [-]

A_{lon} = Turbulence Amplitude [-]

K_{lat} = Lateral Control to Roll Attitude Gearing [rad]

L_w = Turbulence Scale Length Parameter [m]

P = Precision Metric [%]

R^2 = Coefficient of Determination [-]

U_0 = Mean Wind Speed [m/s]

W = Workload metric [1/sec]

W_{min} = Theoretical minimum workload for an MTE [1/sec]

W_{noise} = White Noise Input [-1:1]

I. Introduction

A. Background

The development of aviation transport technology in the last half century has largely followed an evolutionary, rather than a revolutionary path. There is good reason for this; revolutionary developments carry much more risk, and typically incur higher costs. So, whilst the evolutionary approach has led to significant gains in performance, efficiency and safety, the airframes of today, particularly in the civil transport arena, would be recognizable to the air traveler of 50 years ago. As a response to this perceived lack of revolutionary innovation in the air transport industry, the European Commission (EC) funded the 'Out of the Box' study (Ref. [1]) to identify new and potentially disruptive concepts for air transport for use in the second half of the 21st century.

One of the six final concepts short-listed by the 'Out of the Box' study to be given further consideration was that of the Personal Aerial Transportation System (PATS). The rationale behind the need for a PATS is the continued increase in the volume of road traffic in and around the world's cities (Refs. [2, 3]), and the congestion that results during peak times. In major European cities such as London, Cologne or Amsterdam, a road-bound commuter might expect to spend over 50 hours per year in traffic jams. Across Europe, delays due to road congestion have been estimated to cost approximately €100bn per year (Ref. [4]). A radical solution to these problems, which will only become worse if road traffic continues to grow as predicted, would be to move commuting traffic from the ground into the air using a PATS.

For a PATS to be successful, it would be necessary to combine the benefits of conventional road transportation (door-to-door, available to all) and air transportation (high speed, comparatively free of congestion), whilst simultaneously avoiding the need for costly infrastructure such as airports, roads etc. The PATS would have to be capable of supporting air traffic flow volumes that are much higher than the present day (which, of course, may conflict with the need to avoid new infrastructure) whilst mitigating any environmental impact (both in terms of fuel efficiency and noise footprint). The PATS would also have to ensure appropriate safety levels through the application of pilot-vehicle interaction and collision avoidance technologies, to name but a few of the additional challenges that would need to be solved for such a vision to be realized. At the same time, any PATS would have to be designed giving due consideration for the general population's needs and wants, including cost effectiveness and affordability.

The results from the 'Out of the Box' study were used to inform the direction of some of the recent EC 7th Framework Programme (FP7) funding calls. One of the subsequent projects funded by FP7 was *myCopter – Enabling Technologies for Personal Aerial Transportation Systems* (Ref. [5]). The aim of the four year *myCopter* project, launched in 2011, was to develop some of the enabling technologies that would ultimately enable a PATS to be realized. This paper provides an introduction to some of the methods developed during the project, using a handling qualities approach, to assess the candidate vehicle response types and the corresponding occupant training requirements for phases of flight that might require some form of human closed-loop control input.

B. The Personal Aerial Vehicle

A key part of any PATS would be the Personal Aerial Vehicles (PAVs) that the traveling public would occupy. Vehicles that might be considered to be PAVs (as distinct from General Aviation (GA) aircraft) have been in existence for over half a century. Since the 1950s, a number of vehicle designs claiming to combine the benefits of the car and the aircraft have been produced. These have included 'roadable aircraft', vehicles which can be driven on the road and are also capable of conventional fixed-wing flight, such as the Taylor 'Aerocar' (Ref. [6]), the Carplane (Ref. [7]) and the Terrafugia Transition (Ref. [8]). Similarly, a number of rotary-wing designs, such as the PAL-V (Ref. [9]), Carter PAV (Ref. [10]), Moller Skycar (Ref. [11]) and Urban Aeronautics X-Hawk (Ref. [12]) have been proposed or have reached the prototype stage for personal transportation purposes.

Each of these designs may be considered as meeting some or all of the criteria for a PAV. However, while some are in the process of being developed for the market, to date, none has achieved mass production; there are many

possible reasons for this. First, the start point for each of these projects has been the vehicle design. The method by which it would be operated and how it would integrate with existing road and air transportation has not always been given due consideration. The roadable aircraft, in particular, still require long runways to operate to and from, reducing their utility for relatively short commuter journeys. Finally, all of the vehicles would still require the user to obtain a pilot's license, imposing further financial and skills testing barriers to mass-adoption.

C. The *myCopter* Project

The *myCopter* project has approached the challenges of personal aviation from a different direction to the more traditional vehicle design projects. The idea is to first identify how such a system would work and how PAVs would operate within a PATS. The actual design of the PAV, which is not part of the myCopter project, could then follow using the outputs of the *myCopter* project as a basis.

The *myCopter* consortium consists of six partner institutions in Germany, Switzerland and the UK, and the project's research activities cover three main themes:

1. Human-Machine Interaction (HMI), including cockpit technologies for inceptors and displays, and vehicle handling characteristics;
2. Autonomous flight capabilities, including vision-based localization and landing point detection, swarming and collision detection and avoidance;
3. Socio-economic aspects of a PATS – the requirements for such a system to become accepted and widely adopted by the general public.

In order to inform the direction of the research, a broad specification for a potential PAV configuration was drawn up in the early stages of the *myCopter* project (Ref. [13]). It is envisaged that the PAV will take the form of a small (1-2 seat) Vertical Take-Off and Landing (VTOL) vehicle capable of cruising at 80-120kts over a range of 50-60 miles. The *myCopter* PAV would not have road-going capabilities.

To meet the requirement for general access to the PATS for all, it is further envisaged that it will be necessary to make significant reductions in the costs associated with traditional GA – including training, operating and maintenance costs. It is the first of these that this paper will focus on. In order to try to reduce the cost of training, two approaches have been considered in *myCopter*. The first of these is to implement autonomous capabilities on the PAV so that the occupant is not required to fly manually. This, however, implies software and hardware designed to very high safety assurance levels. In turn, this implies expensive testing and certification processes

which will add to the development costs of the vehicles. A second, alternative option (perhaps for early adopters of such a technology as the traffic densities would almost certainly have to be lower) is therefore to ascertain the required PAV Handling Qualities (HQs) such that the degree of 'skill' associated with piloting the vehicle is significantly reduced compared to that required for a traditional GA rotorcraft, for example. While very high levels of safety assurance will also be required for the flight control systems that will be associated with these excellent HQs, the process to achieve this should be more straightforward for manual flight than for fully autonomous flight. The University of Liverpool (UoL) is working within the first of the themes described above to develop these HQ requirements for PAVs operating within the PATS. The objective of this research is to identify required response types and boundaries for predictive metrics for the PAV in much the same way as ADS-33E-PRF (Ref. [14]), the US Army HQ performance standard, does for military rotorcraft.

D. Paper and problem overview

While methods for HQ assessment of conventional rotorcraft have become widely accepted using standards such as ADS-33E-PRF, these have been developed for use by professional test pilots who are assessing the vehicle's capabilities as a proxy for a well-trained and well-motivated professional pilot. The much broader spectrum of potential PAV occupants (as with car driving, from what might be termed 'naïve' i.e. lacking experience, through to highly skilled) means that it was considered unlikely to be the case that these methods on their own would be directly applicable or even useable for PAV HQ assessments.

It was therefore considered necessary, using the existing body of knowledge as a basis, to develop additional methods to support vehicle HQ analysis when the intended pilot is non-professional. This paper documents those tools and techniques used by the myCopter project team for this purpose. A companion paper will report upon the test results and conclusions obtained using these methods. Section II describes the experimental test environment, including the vehicle flight dynamics modeling employed, the Mission Task Elements (MTEs) used, the flight simulation facility itself and the environmental modeling methods required. Section III reports upon the configuration development and subsequent analysis of the different vehicle models using conventional HQ analysis techniques. Section IV then describes the methods used to describe and assess the PAV pilot subjects and illustrates the data used to validate them. Finally, Section V brings the paper to a close with conclusions regarding the work presented.

II. Testing Environment

The main facility employed for the PAV HQ evaluations has been the HELIFLIGHT-R flight simulator at UoL (Ref. [15]), using a MATLAB/Simulink model of a generic VTOL aircraft – the Generic PAV Dynamics Model (GPDM, Ref. [16]). This section describes the creation of the test environment that is used throughout the PAV HQ assessment process; the key features of the HELIFLIGHT-R simulator; the GPDM itself; the Mission Task Elements (MTEs) flown during the evaluations and finally the simulation of the harsh environment.

A. HELIFLIGHT-R simulator

The HELIFLIGHT-R simulator (Fig. 1) has been the main research tool for HQ criteria development in the *myCopter* project. HELIFLIGHT-R features a two-seat crew station inside a 12ft diameter dome and a simulation engineer's station at the rear. The crew station and projection dome are mounted on top of a hexapod platform offering six degrees-of-freedom motion cueing. The outside world scene is rendered using the Vega Prime image generator (Ref. [17]), and projected onto the dome by three HD projectors. The output from each Vega Prime display channel is warped and blended to create a seamless image on the surface of the dome covering a field of view of approximately 210° by 70°. This is extended in the region ahead of and below the pilot by a pair of 'chin' windows.

Fig. 1 The HELIFLIGHT-R Simulator at the University of Liverpool

Four axis (lateral and longitudinal cyclic, collective and pedals) dynamic control-loading permits the inceptor force-feel characteristics to be tuned to represent specific configurations or adjusted to investigate the impact on

vehicle handling. While these standard rotorcraft inceptors have been used in the HQ evaluations to date, one of the areas of study as the *myCopter* project progresses will be to identify whether these are the most suitable inceptors for a PAV, or whether an alternative configuration, such as a yoke, or even a steering wheel, as would be found in a car, would be more suitable. This is, however, beyond the scope of the current paper.

The crew station's instrument panel features a pair of reconfigurable 'glass cockpit' style Primary Flight Displays (PFDs, Fig. 2) together with two smaller displays on a center console. For the *myCopter* HQ evaluations, the instrument panel displays were configured to show a slightly modified (through the addition of a radar altimeter and Height and Direction Hold engaged indicators (HH and DH respectively)) representation of the Garmin G1000 GA glass cockpit PFD (Ref. [18]). The G1000 glass cockpit was selected as it is one of a small number of such products that are currently available in the GA market. It was anticipated that the clear presentation of the PFD symbology would permit the flight-naïve pilots taking part in the *myCopter* trials to learn to interpret the display quickly.

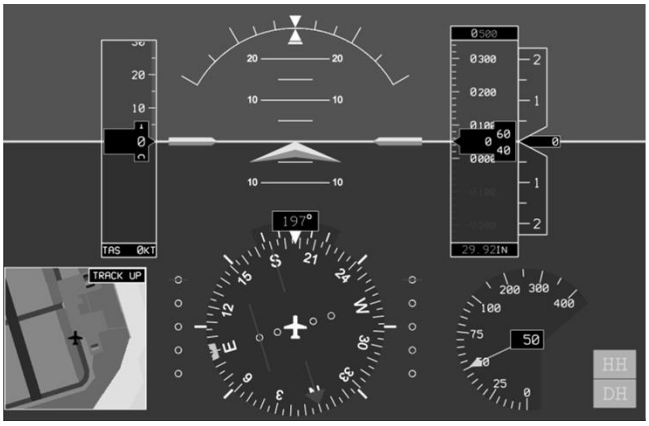

Fig. 2 Garmin G1000 PFD as Implemented at the University of Liverpool

In addition to the head-down display symbology offered by the G1000 panel, a set of basic Head-Up Display (HUD) symbology has been developed, which is overlaid onto the outside world scene (Fig. 3). The HUD symbology includes: a Malcolm horizon line (Ref. [19]) spanning the full field of view of the simulator (rendered in orange-brown); a flight path vector indictor showing the current direction of flight (white); an attitude indicator showing attitude relative to the horizon (green); numerical readouts of current airspeed (and, for the longitudinal acceleration-rate-based response type (described in the following Section), commanded airspeed), heading and height above the terrain (green apart from the commanded airspeed, which is red); and during decelerating flight, a

conformal display element that indicates the point on the ground above which the vehicle will come to a stop assuming that the deceleration rate remains constant (cyan).

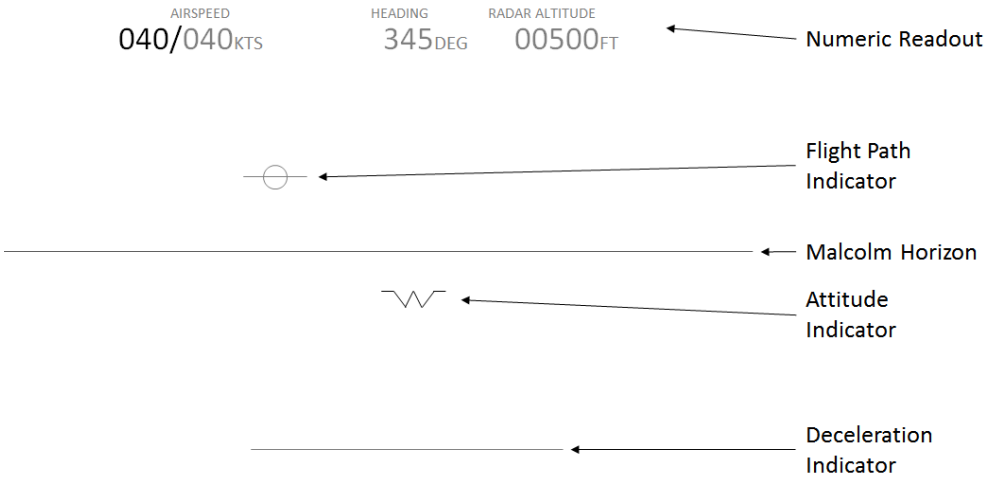

Fig. 3 Schematic of Head-Up Display Symbology

The HUD symbology has been kept deliberately simple and sparse so as to facilitate assimilation and interpretation by flight-naïve pilots. The exact requirements for such a display are one of the areas of study for the *myCopter* project (Ref. [20]) and it is not claimed here that this is the final or optimum set.

B. PAV Model Configuration

The underlying response types offered by the GPDM are either Rate Command (RC) or Attitude Command, Attitude Hold (ACAH) in the pitch and roll axes, and RC in the yaw and heave axes. These responses are created through 1st order (RC) or 2nd order (ACAH) transfer functions (Ref. [21]). Equation 1 illustrates this method; it shows how the roll axis dynamics are specified for the ACAH response type.

$$\frac{\phi_{cmd}}{\delta_{lat}} = \frac{K_{lat}}{\frac{1}{\omega_{lat}^2}s^2 + \frac{2\zeta_{lat}}{\omega_{lat}}s + 1} \tag{1}$$

The use of transfer functions for the rotational motion permits the vehicle dynamics to be tuned rapidly and to obtain precise HQs (e.g. a specified bandwidth value can be set directly), facilitating evaluation of multiple configurations. Of course, this approach does not lend itself to a direct physical interpretation of the individual elements of the vehicle, but this was considered to be acceptable for this project, given that an actual myCopter PAV

design does not yet exist. In the heave axis, the collective lever controls a vertical 'lift' force, which, when tilted via pitch and roll control, creates longitudinal and lateral accelerations through standard rigid body dynamics.

The basic responses can be augmented through 'outer loop' feedback to create, for example, a Translational Rate Command (TRC) response type for pitch and roll, or a flight path angle response type (γC) in the heave axis.

Three baseline vehicle configurations have been developed to facilitate the *myCopter* HQs research. Each configuration can be considered as being representative of a different level of 'augmentation' of a vertical flight vehicle's responses. The three configurations are as follows:

1. *Configuration 1: 'Rate Command'*

RC responses in pitch and roll are combined with RC in heave. In yaw, the response type in the hover is RC, but as the speed increases, directional stability is introduced through sideslip angle feedback, providing a sideslip angle command (βC) response type at forward flight speeds greater than 25kts. Additionally, in forward flight, turn coordination inputs are applied to the roll, pitch and yaw controls to ensure that the vehicle performs smooth turns without additional pilot activity. Apart from these coordination inputs, inter-axis coupling is not present in the model. This is on the basis that such a coupling would be highly undesirable in a PAV, would complicate the analysis of individual response types and would likely compromise the ability of flight-naïve pilots to complete the specified tasks. It is not envisaged that a PAV with strong cross-coupling would be a viable solution. The dynamics of this configuration have been tuned to offer predicted Level 1 HQs for the 'All Other MTEs' category of tasks according to the US Army rotorcraft handling qualities specification, ADS-33E-PRF (Ref. [14]). The rate-based response types of this configuration may be considered as being approximately representative of a current light GA helicopter, albeit one with excellent HQs.

2. *Configuration 2: 'Attitude Command, Attitude Hold'*

The second configuration may be considered as being approximately similar to a modern, augmented helicopter. It is configured as described above for the RC response type. The difference is the primary response in the pitch and roll axes, where an ACAH response type is used rather than the RC response type of Configuration 1. Again, the dynamics of this configuration have been tuned to offer predicted Level 1 HQs according to ADS-33E-PRF.

3. Configuration 3: 'Hybrid'

The 'Hybrid' configuration has been designed so that the response type offered to the pilot in each axis changes with flight condition, allowing the dynamics to be more closely matched to the demands of a particular task than is the case with configurations 1 and 2. For the tests described in this paper, there are sets of response types for hover and low speed maneuvering (at speeds up to 15kts), and for forward flight maneuvering (at speeds above 25kts). Smooth blending occurs between the hover response types and the forward flight response types as the speed increases from 15kts to 25kts and vice versa.

In the hover and low speed segment of the flight envelope, the response type for the pitch and roll axes is TRC. Yaw and heave are RC as with configurations 1 and 2. In forward flight, yaw behaves in the same way as with configurations 1 and 2, but in the heave axis the response type changes to γC. In roll, the response type changes to ACAH. In pitch, the response type changes to Acceleration Command, Speed Hold (ACSH). The ACSH response type generates, for a fixed displacement of the longitudinal controller, a constant rate of change of airspeed; releasing the controller to the zero force position results in the currently commanded airspeed being held.

The transition between TRC and ACSH modes during deceleration does not follow the general pattern of blending between 15kts and 25kts. Instead, the ACSH mode is maintained throughout the deceleration until the vehicle comes to a stop. At this point, the response type is switched back to TRC ready for the next pilot input. In addition, the Hybrid configuration is equipped with pilot selectable Height Hold (HH) and Direction Hold (DH) functions.

The response characteristics of the Hybrid configuration in each airspeed region are summarized in Table 1.

Table 1 Summary of Hybrid Configuration Response Types

Speed Range	Pitch	Roll	Yaw	Heave
<15kts	TRC	TRC	RC	VRC
blend	Instantaneous at 15kts (accel) and 0kts (decel); internal logic to eliminate transients	Smoothed transition between 15-25kts	Smoothed transition between 15-25kts	Smoothed transition between 15-25kts
>25kts	ACSH	ACAH	βC + TC	γC

The philosophy behind the selection of response types for the Hybrid configuration has been, where possible, to minimize the number of control inputs required to perform a maneuver.

As with the other configurations, the dynamics of the Hybrid configuration have, where appropriate criteria exist, been tuned to offer predicted Level 1 HQs for 'All Other MTEs' according to ADS-33E-PRF. The HQs of the

inner loop ACAH responses are identical to those of the ACAH configuration itself. For the response types not covered by ADS-33E-PRF (such as the ACSH mode), subjective tuning has been performed to create a satisfactory response. This tuning was conducted on the basis of feedback from the first group of Test Pilots (TPs) and flight-naïve Test Subjects (TSs) to fly the GPDM in the Hybrid configuration.

4. Comparison of responses of PAV configurations

It is instructive to consider the way in which the various configurations behave when a pulse control input is applied to a specific axis of motion. Inputs such as these form the basis for many of the ADS-33E-PRF predictive analyses, and so examination of the time histories can help to understand the HQ results. Figure 4 shows the various responses in the pitch axis, Fig. 5 the responses in the roll axis and Fig. 6 the responses in the yaw axis. In the cases of Figs. 4 and 5, the magnitude and duration of the pulses for each response type have been selected to generate rate and attitude responses of comparable magnitude. The progressively increasing stability of the responses as the response type progresses from RCAH through ACAH to TRC is evident in the velocity traces for each axis.

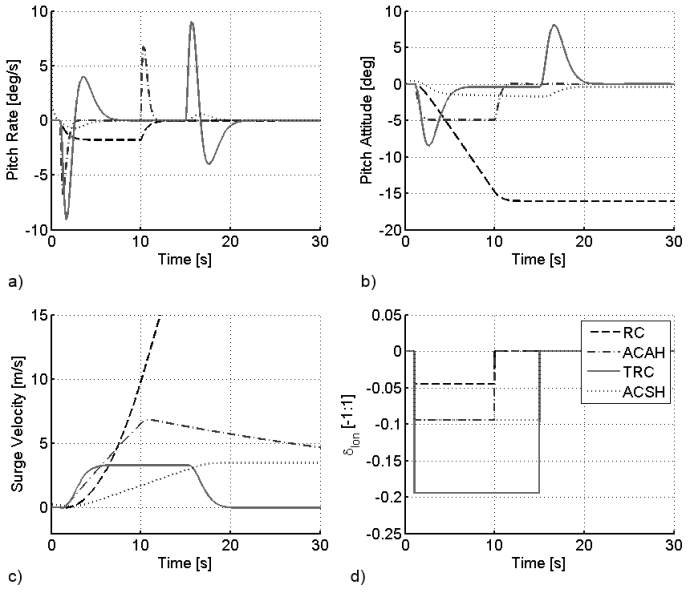

Fig. 4 myCopter PAV Model Pitch Axis Responses to Pulse Inputs

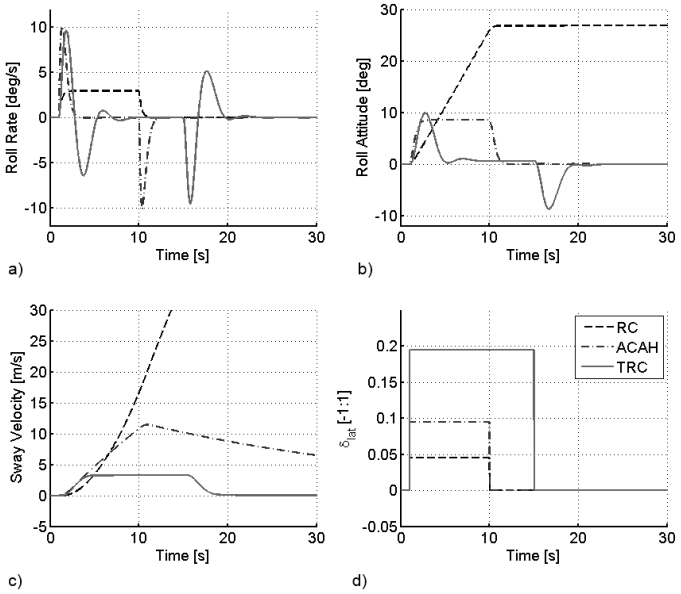

Fig. 5 myCopter PAV Model Roll Axis Responses to Pulse Inputs

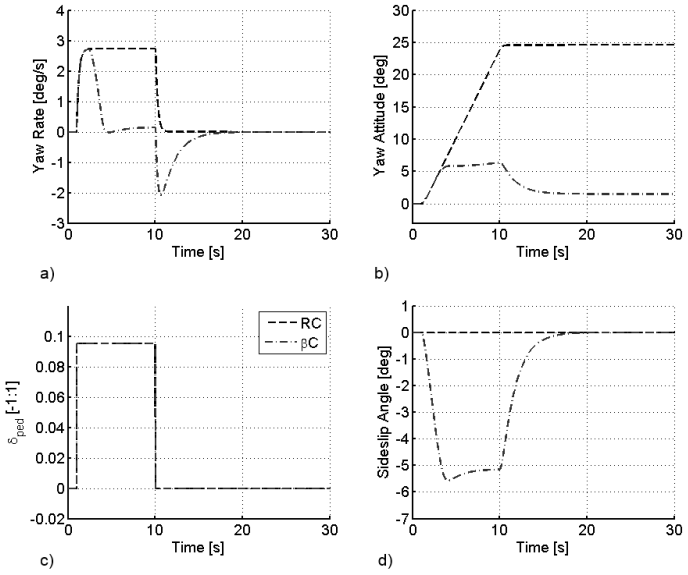

Fig. 6 myCopter PAV Model Yaw Axis Responses to Pulse Inputs

C. Mission Task Elements

As part of the wider *myCopter* project, a commuting scenario was developed, whereby the PAV flight would begin with a vertical take-off from a rural or suburban region (Ref. [13]). The PAV would accelerate and climb into a cruise towards its final destination, typically the central business district of a major city. Upon arrival at the destination, the PAV would descend and decelerate to hover at a designated PAV landing point, before executing a vertical landing. From this general commuting scenario, a series of Mission Task Elements (MTEs) appropriate to

the PAV role have been identified and a subset of 5 hover and low speed MTEs were selected for use in the investigations reported in this paper. The 5 MTEs used were the Hover, Vertical Reposition, Landing, Decelerating Descent and Aborted Departure. Where possible, the outline of the task has been drawn from ADS-33E-PRF; some of the task performance requirements have, however, been modified (generally relaxed) to reflect the nature of the PAV role.

1. Hover MTE

The hover maneuver is initiated from a stabilized hover at a height above ground level of 20ft and the aircraft is accelerated towards the target hover position. The target hover point is oriented at 45° relative to the heading of the aircraft. The ground track is such that the aircraft will arrive over the hover point, and the aircraft should translate at a ground speed between 6 and 10kts. Upon arrival at the hover point, a stable hover should be captured and held for 30 seconds. The transition to hover should be accomplished in one smooth movement; it is not acceptable to accomplish most of the deceleration well before the hover point and then to creep up to the final position. The performance requirements for this task are shown in Table 2, and the test course used in the piloted simulations is shown in Fig. 7 – the board and pole together provide the pilot with cueing for desired and adequate vertical and lateral positioning, whilst the cones on the ground around the target hover point indicate the desired and adequate longitudinal position tolerances.

Table 2 Hover MTE Performance Requirements

Parameter	Desired	Adequate
Maintain longitudinal position within ±X ft of the target hover point	3	6
Maintain lateral position within ±X ft of the target hover point	3	6
Maintain heading within ±X°	5	10
Maintain height within ±X ft	2	4

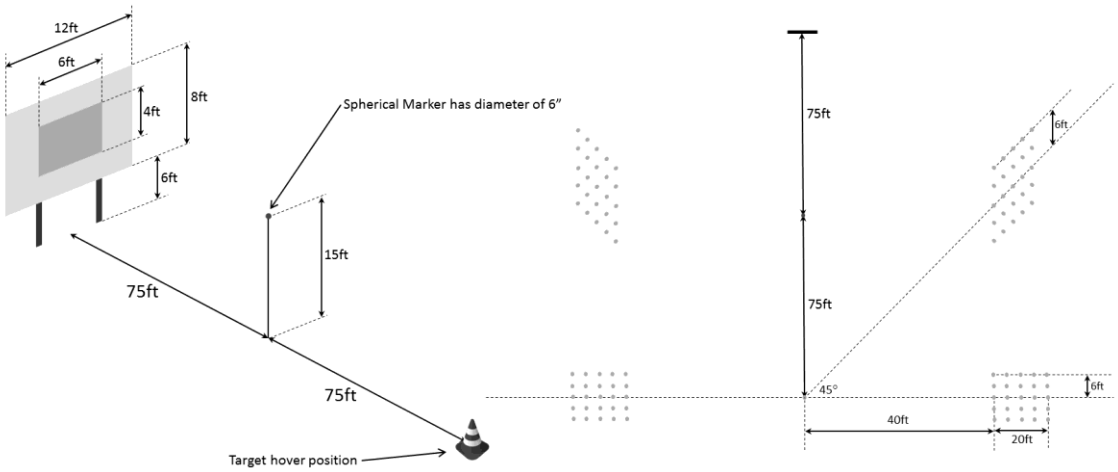

Fig. 7 Hover MTE Test Course

2. Vertical Reposition MTE

The vertical reposition maneuver starts in a stabilized hover at an altitude of 20ft with the aircraft positioned over a ground-based reference point. A vertical climb is initiated to reposition the aircraft to a hover at a new altitude of 50ft within a specified time. Overshooting the end point is not permitted. The maneuver is complete when a stabilized hover is achieved. The performance requirements for the vertical reposition maneuver are shown in Table 3, and the test course used in the piloted simulations is shown in Fig. 8.

Table 3 Vertical Reposition MTE Performance Requirements

Parameter	Desired	Adequate
Maintain longitudinal position within ±X ft of the target hover point	5	10
Maintain lateral position within ±X ft of the target hover point	5	10
Maintain heading within ±X°	5	10
Capture new height within ±X ft	2	4
Complete the manoeuvre within X seconds	10	15

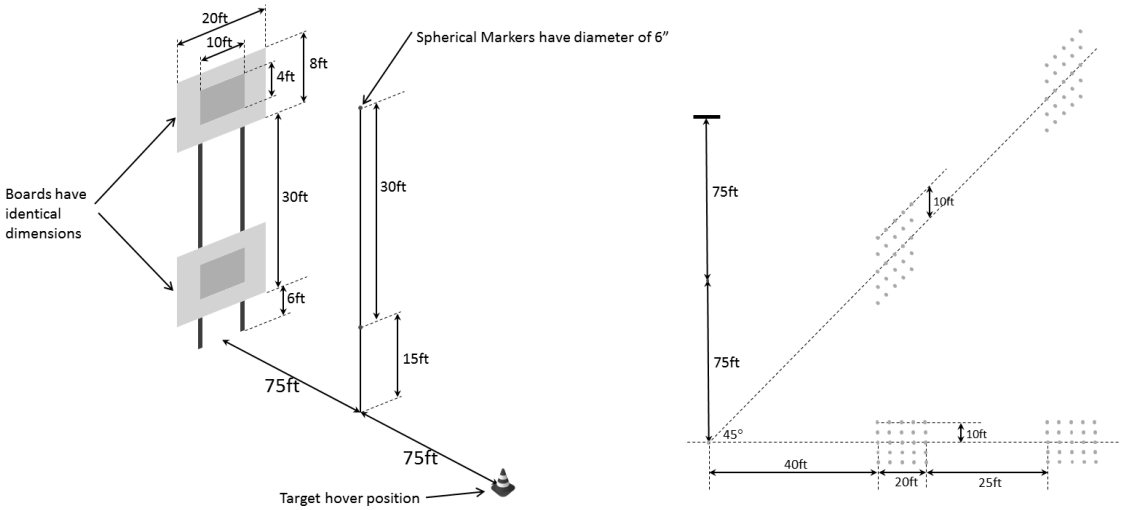

Fig. 8 Vertical Reposition MTE Test Course

(note: ground markings are repeated to the left of the hover point)

3. Landing MTE

The landing maneuver starts with the vehicle in a stable hover at a height of 20ft, offset laterally and longitudinally from the prescribed landing point. Following a repositioning phase to place the vehicle in a hover directly above the landing point, an essentially steady descent to the landing point is conducted. It is acceptable to arrest sink rate momentarily to make last-minute corrections prior to touchdown. The performance requirements for the landing maneuver are shown in Table 4, and the test course used in the piloted simulations is shown in Fig. 9.

Table 4 Landing MTE Performance Requirements

Parameter	Desired	Adequate
Accomplish a gentle landing with a smooth continuous descent, with no objectionable oscillations	✓	N/A
Once height is below 10ft, complete the landing within X seconds	10	N/A
Touch down within ±X ft longitudinally of the reference point	1	3
Touch down within ±X ft laterally of the reference point	0.5	3
Attain rotorcraft heading at touchdown that is within ±X° of the reference heading	5	10
Final position shall be the position that existed at touchdown	✓	N/A

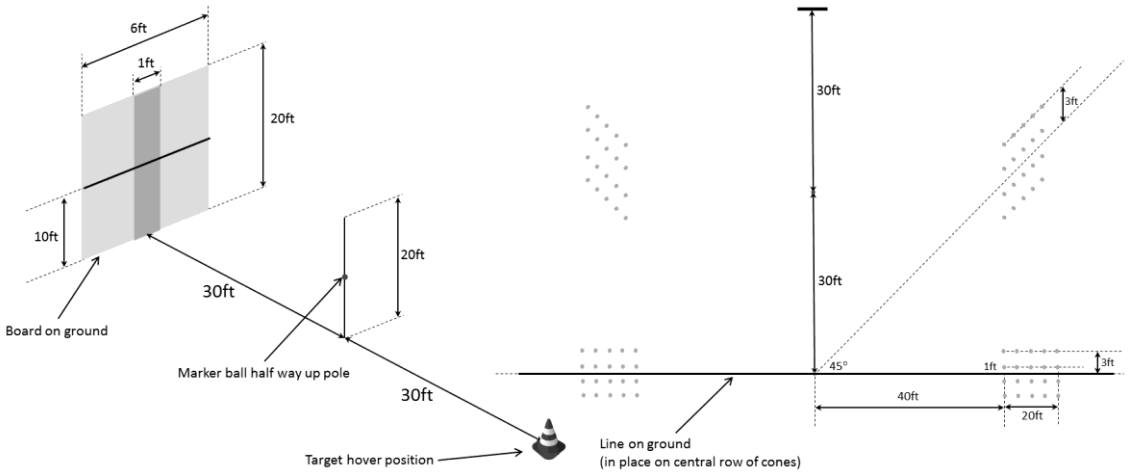

Fig. 9 Landing MTE Test Course

4. Decelerating Descent MTE

The decelerating descent maneuver begins with the aircraft in a stable cruise at 60kts at a height of 500ft above the ground. Once a specified ground marking has been reached, the pilot initiates a descent and decelerates towards a target hover point and a height of 20ft. The approach is configured to give a mean glideslope angle of 6 degrees. The maneuver is complete when the aircraft has been stabilized over the marked maneuver end point. Overshooting the approach beyond the front longitudinal adequate tolerance, or the lower vertical adequate tolerance is not permitted. The performance requirements for the decelerating descent maneuver are shown in Table 5, and the test course used in the piloted simulations is shown in Fig. 10.

Table 5 Decelerating Descent MTE Performance Requirements

Parameter	Desired	Adequate
Maintain the lateral position within ±X ft	20	50
Maintain heading within ±X°	10	15
Stabilise target height within ±X ft	5	10
Stabilise hover point within ±X ft longitudinally of marked position	10	20

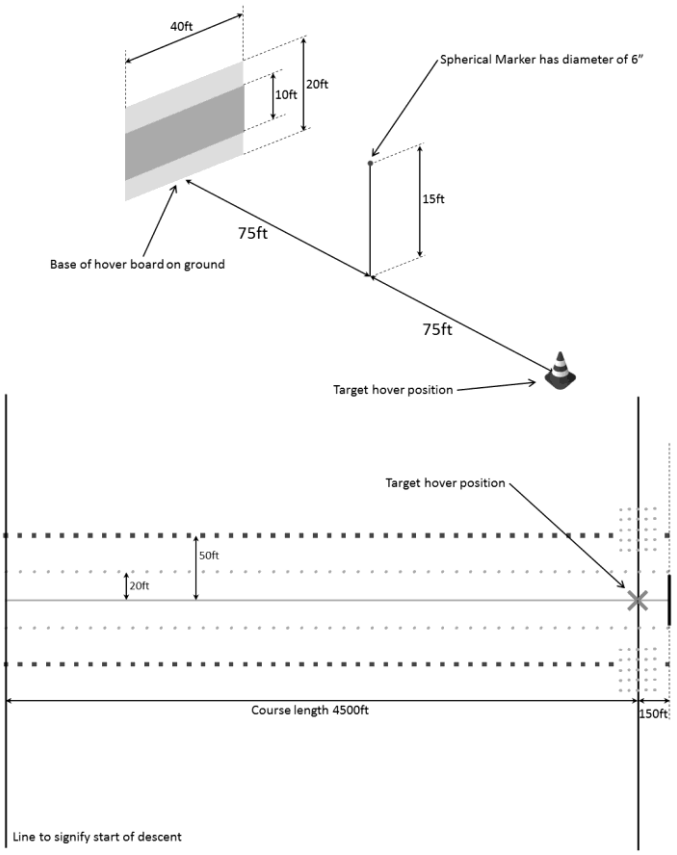

Fig. 10 Decelerating Descent MTE Test Course
(note: upper figure shows final 150ft of test course only for clarity)

5. Aborted Departure MTE

The aborted departure begins in a stabilized hover at an altitude of 50ft. A normal departure is initiated by accelerating the aircraft longitudinally along a target trajectory (using a nose down pitch attitude of approximately 15°). When the groundspeed has increased to 40kts, the departure is aborted and the vehicle is decelerated to a hover as rapidly and as practicably as possible. The acceleration and deceleration phases should each be accomplished in single, smooth maneuvers. The maneuver is complete when control motions have subsided to those necessary to maintain a stable hover. The performance requirements for the aborted departure maneuver are shown in Table 6, and the test course used in the piloted simulations is shown in Fig. 11.

Table 6 Aborted Departure MTE Performance Requirements

Parameter	Desired	Adequate
Maintain the lateral position within ±X ft	10	20
Maintain heading within ±X°	10	15
Maintain height within ±X ft	10	20
Complete the manoeuvre within X seconds	25	30

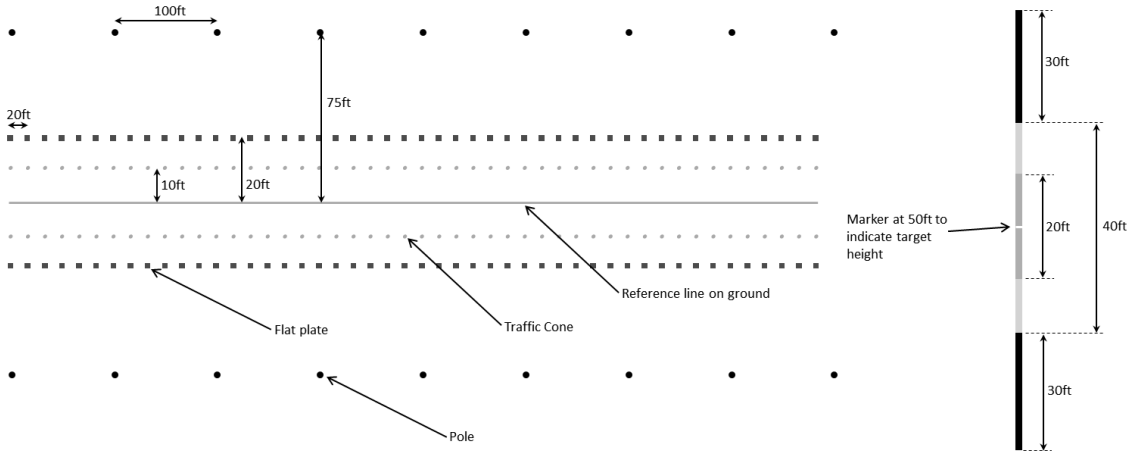

Fig. 11 Aborted Departure MTE Test Course

D. Simulating the Harsh Environment

1. Introduction to Harsh Environments

A newly-qualified GA pilot is typically restricted to operating in daylight hours when the visibility is good (greater than 4500ft [22]). However, these conditions are not always available. For example, the three main airports in the North West of England, at Liverpool, Manchester and Blackpool, have recorded an average of 21.7 days with fog (visibility less than 3000ft) per year over the period 1993-2013 (Ref. [23]). A comparable figure for three airports in the North Eastern United States and Eastern Canada (New York John F. Kennedy; Chicago O'Hare; and Montréal Dorval) is 33.0 days with fog per year (Ref. [23]). If a PAV was unable to operate in these conditions (or in other degraded visibility conditions such as rain, sleet or snow), it would be a significant impediment to the utility of the vehicle for one of its anticipated primary roles: commuting to and from a place of work. It is therefore important that the impact on PAV HQ requirements of degrading the visual cues is considered.

The US military performance specification for the Handling Qualities (HQs) of helicopters, ADS-33E-PRF (Ref. [14]) defines minimum acceptable response types for various stages of degradation in the Useable Cue Environment (UCE). For UCE=1 (excellent task cueing available), a basic rate response type is all that is required for Level 1 handling. In UCE=2 conditions (some degradation in either attitude and/or translational rate cueing), a more strongly stabilized ACAH response type is required to maintain Level 1 handling, while in UCE=3 conditions (severe degradation in attitude and/or translational rate cueing), a TRC response type is required. These requirements were developed on the basis that the pilot flying the helicopter would be a well-motivated professional pilot, with extensive training in the skills required to control the helicopter. For the potentially flight-naïve PAV pilot, the relationship between UCE and required response type needs to be established.

Atmospheric disturbances can occur in either free stream conditions or as a result of the air passing around obstructions such as trees or buildings (Ref. [24]). It is expected that PAVs will operate into and out of the central business districts of major cities in their commuting role. Thus, it is important that the PAV pilot is able to perform precision take off, landing and low speed maneuvering tasks in the presence of disturbances that may typically be found in these locations.

2. Degraded Visual Conditions

Degradation in the UCE was achieved by restricting the visibility range of the UoL HELIFLIGHT-R simulator's Vega Prime image generation system. Vega Prime employs an atmospheric illumination model to automatically darken the scene and provide a 'natural' fog onset in response to limiting the visibility. In this context, visibility is measured as the maximum range from the observer at which no further world features can be seen.

In the simulated 'benign environment', the visibility range was set to a sufficiently large value that there were no apparent reductions in scene contrast and no obscured features within the regions scanned by the pilot (Fig. 12).

Fig. 12 Hover Test Course in Good Visual Conditions

For the investigation into the effect of degraded UCE, the visibility range was reduced to 800ft (Fig. 13). This range just provided the pilot with visibility of the task cues from the starting point for each of the hover and low speed MTEs. All natural horizon references were obscured together with many of the vertical features of the terrain database. Additionally, the micro-contrast of the textures used on the ground was significantly reduced.

Fig. 13 Hover Test Course in Poor Visual Conditions

The effect of the introduction of the fog model was assessed by a test pilot who awarded Visual Cue Ratings (VCRs, Ref. [14]) to determine the UCE level of the environment. With reduced visibility and micro-contrast, the VCRs were increased (Fig. 14, where the MTE identifications are as follows: MTE 1 = Hover; MTE 2 = Vertical Reposition; MTE 3 = Landing and MTE 4 = Aborted Departure), creating a Degraded Visual Environment (DVE). Taking the hover test course as an example, for attitude control, the VCR awarded by the assessing test pilot increased from 1.5 to 4.5, whilst the translational rate VCR increased from 3 to 5. Together, the effect of these changes is to degrade the UCE from UCE=1 to UCE=3.

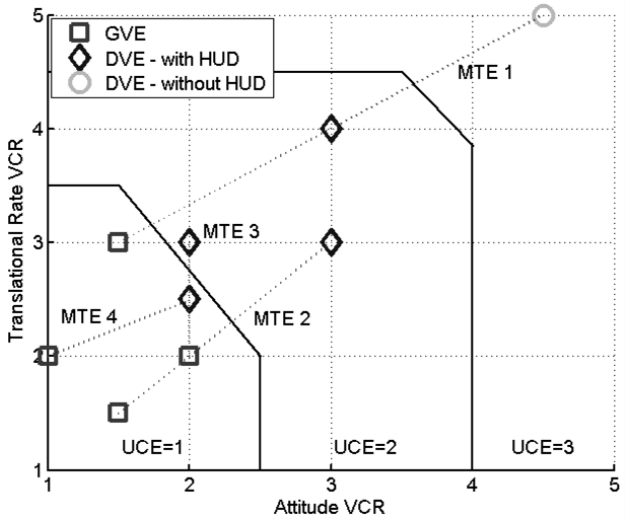

Fig. 14 Determination of UCE for Hover and Low Speed MTEs

The flight-naïve TSs were not, however, asked to fly the MTEs in UCE=3 conditions. As described above, the PAV simulation is equipped with a HUD (Fig. 3). Two features of the HUD in particular provide enhancements to the UCE. The first is the wide field of coverage Malcolm horizon. This, together with the attitude indicator, provides the pilot with a strong reference for vehicle attitude and attitude rate in any external visual conditions. The second key HUD feature is the flight path indicator. This marker shows the pilot the current direction of travel (in three dimensions) of the PAV, and hence can be used to augment cueing of the ratio of longitudinal to lateral translational movement.

In general therefore, the harsh environment evaluations were conducted in UCE=2 conditions. The exception was the Aborted Departure MTE, where the very strong task cueing provision meant that even in the presence of the simulated fog, sufficiently strong cues were provided to still achieve UCE=1 (Fig. 15).

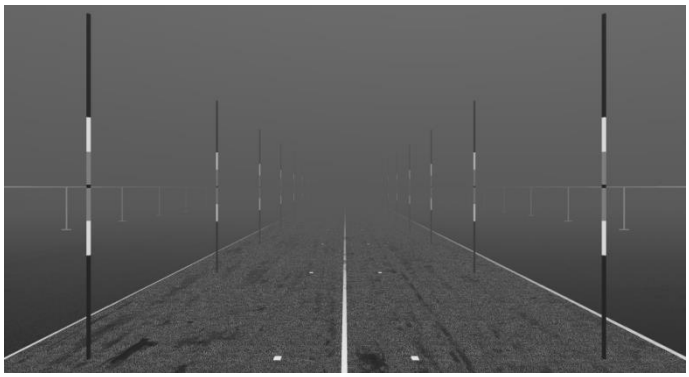

Fig. 15 Aborted Departure MTE Test Course in DVE

The Decelerating Descent MTE was not used in the harsh environment testing due to the strong reliance placed on distant visual cues in this task. In the GVE, UCE=1 conditions prevailed for this task.

3. Atmospheric Disturbances

Many different models of atmospheric disturbances have been created and applied to the simulation of helicopters (Ref. [25]). Among the most popular of these is the von Karman method (Ref. [26]). This method models continuous gusts with specified Power Spectral Density (PSD) characteristics for gust magnitudes in surge, sway and heave; angular gust components are also modeled. Wind speed variations are generated for insertion into the simulation by passing white noise signals through filters that are designed to approximate the von Karman PSD characteristics.

Due to the non-physical nature of the PAV simulation model, however, it was not feasible to generate disturbances using the von Karman method. Instead, the Control Equivalent Turbulence Input (CETI, Ref. [27]) method was adopted. The CETI approach uses a similar technique to the von Karman method, in that it passes white noise through appropriately designed filters to generate disturbance signals. However, rather than applying the output of the filters as changes to the atmospheric model around the aircraft, the outputs are applied as control inputs (lateral and longitudinal cyclic, collective and pedals) that generate 'equivalent' aircraft responses to those that would have been experienced by the vehicle when exposed to the originally-modeled gusts. The CETI technique is constrained relative to the von Karman method in that it can only create disturbances that are achievable using the controls of the aircraft, but it has the advantage that it can be realized in any simulation without the need to control the local atmospheric properties.

The structure of the CETI models used in this study was adopted from Ref. [27]. For example, the structure of the CETI model for the longitudinal cyclic is shown in Equation 2 below:

$$\frac{\delta_{lon,gust}}{w_{noise}} = A_{lon}\frac{1}{\left(s+\frac{U_0}{L_w}\right)} \tag{2}$$

Initial parameter settings for the turbulence filters were also drawn from Ref. [27]. However, as the PAV is considered to be a small air vehicle (mass ≈ 500kg), and the aircraft used to generate the parameters was considerably larger (an EC-135 helicopter, mass ≈ 2800kg), the parameter settings were subjectively tuned using the opinions of a current light helicopter pilot to create a turbulence response that was considered appropriate for a vehicle of the size of the PAV. Frequency spectra for the transfer functions used in the harsh environment study are shown in Fig. 16 below (the settings can be found in the Appendix).

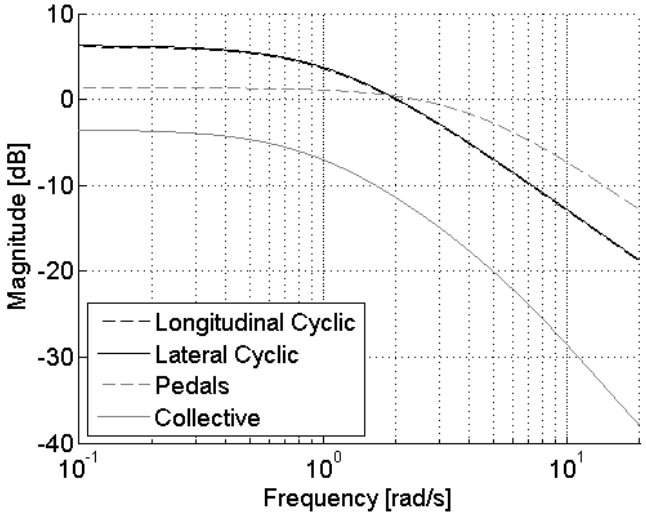

Fig. 16 Comparison of CETI Filters for PAV Simulation

When driven by white noise generators, these transfer functions produce control input signals (such as those shown in Fig. 17) which command angular rate (or vertical rate in the case of the heave axis) perturbations.

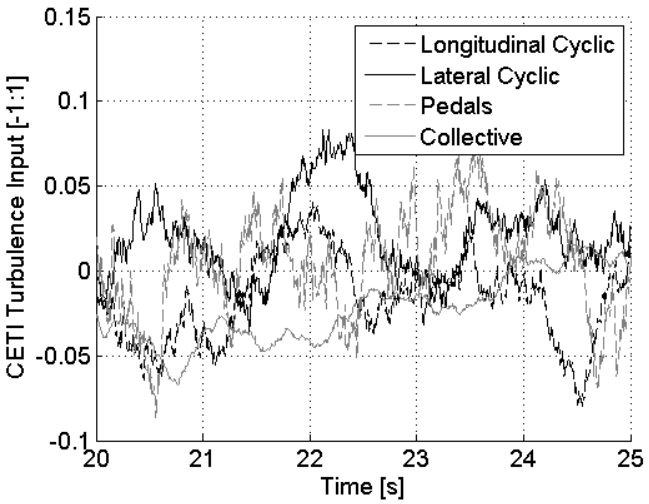

Fig. 17 Samples of Typical Turbulence Inputs to PAV Simulation

For the RC configuration, the CETI signal is fed directly into each channel of the model, as shown in Equation 3 for the pitch axis:

$$\delta_{lon,total} = \delta_{lon,pilot} + \delta_{lon,gust} \tag{3}$$

This approach models a system where there is no closed-loop feedback of vehicle response into the flight control system – in other words, an unstabilized response.

For the ACAH and Hybrid configurations, the CETI signal in pitch and roll is firstly integrated to create a commanded attitude disturbance, and the integrated signal is fed into the appropriate control channel. It is assumed with these configurations that closed-loop feedback of vehicle attitude is present within the 'virtual' flight control system. Therefore, delayed feedback of the commanded attitude disturbance is also applied to each control channel. The resultant implementation is shown in Equation 4:

$$\delta_{lon,total} = \quad \delta_{lon,pilot} + \left(\int \delta_{lon,gust}.\,\mathrm{dt}\right) - \left(e^{-\tau s} \int \delta_{lon,gust}.\,\mathrm{dt}\right) \tag{4}$$

The time delay of the 'sensor' (τ in Eq. 4) was modeled as 100ms, this figure being representative of the measurement and associated digital data processing time that would be required in a real sensor feedback system. The turbulence implementation for the yaw and heave axes were also subject to closed-loop feedback in the ACAH and Hybrid configurations, meaning that, although the underlying vehicle response to a control input is identical to the RC configuration in these axes, the turbulence response is different.

The effect of the turbulence on the RC and ACAH configurations is shown in Fig. 18. The attitude response of the Hybrid configuration is similar to that of the ACAH configuration seen in Fig. 18, but the translational response is more stable due to the additional closed-loop feedback of velocity present in the TRC control loop.

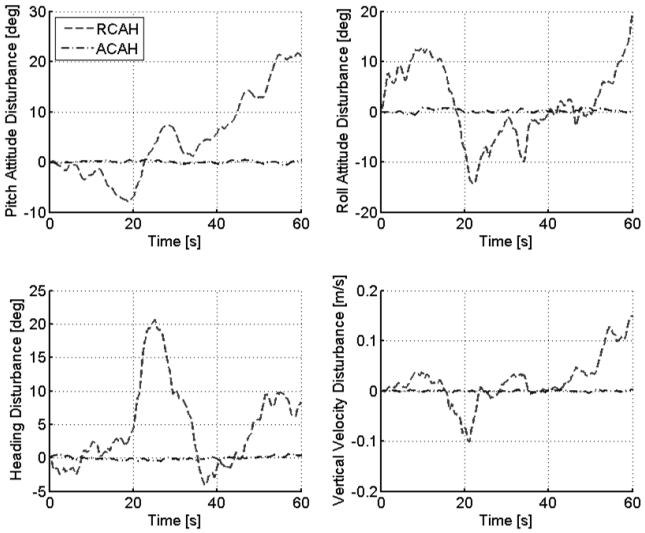

Fig. 18 Turbulence response of RCAH and ACAH configurations

The strength of the turbulence was assessed subjectively by a test pilot using the Turbulent Air Scale (Ref. [28], repeated in Table 7). For the 'unaugmented' RC configuration, a rating of 5, occasionally increasing to 6, was awarded, indicating moderate turbulence intensity.

Table 7. Turbulent Air Scale

Rating	Definition	Air Conditions
1	-	Flat calm
2	Light	Fairly smooth, occasional gentle motion
3		Small movements requiring correction
4	Moderate	Continuous small bumps
5		Continuous medium bumps
6		Medium bumps with occasional heavy one
7	Severe	Continuous heavy bumps
8		Occasional negative 'g'
9	Extreme	Rotorcraft difficult to control
10		Rotorcraft lifting several hundred feet

III. Conventional Handling Qualities Evaluation

The three PAV configurations have been assessed against the ADS-33E-PRF hover and low speed criteria for 'All Other MTEs'. The results in this Section focus on vehicle responses in the hover, as this is the condition in which the majority of the piloted simulation tests have been performed. However, as the rotational dynamics of the GPDM are created through transfer function models, these predicted HQ values will remain constant across the flight envelope.

In the pitch axis, the bandwidth of the RC and ACAH configurations is as shown in Fig. 19, while the attitude quickness is as shown in Fig. 20. The bandwidths of the two configurations are quite different; this is a result of the configurations being tuned to exhibit similar attitude quickness properties. The different structures used to implement the RC and ACAH response types in the GPDM prevent an exact match, and additionally lead to different bandwidth results. For each criterion, the handling qualities are predicted to lie within the Level 1 region.

It should be noted that, for all of the bandwidth analyses shown below, the results incorporate a time delay of 80ms. This is representative of the inherent stick-to-visuals transport delay of the HELIFLIGHT-R simulator – with the exception of the 'sensor' delays described above, the GPDM itself does not include any additional delay elements.

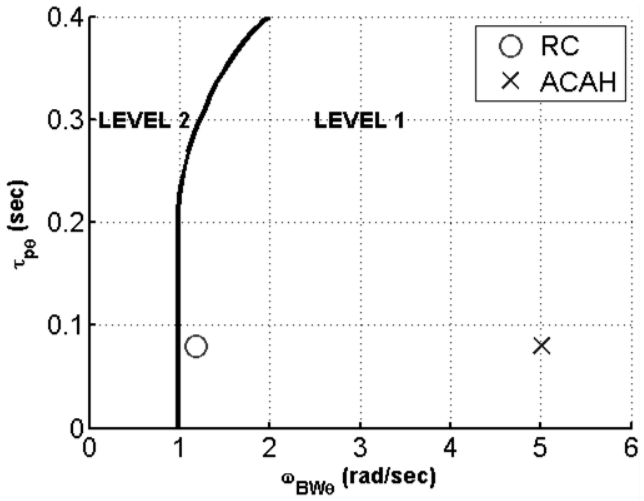

Fig. 19 Pitch Axis Bandwidth

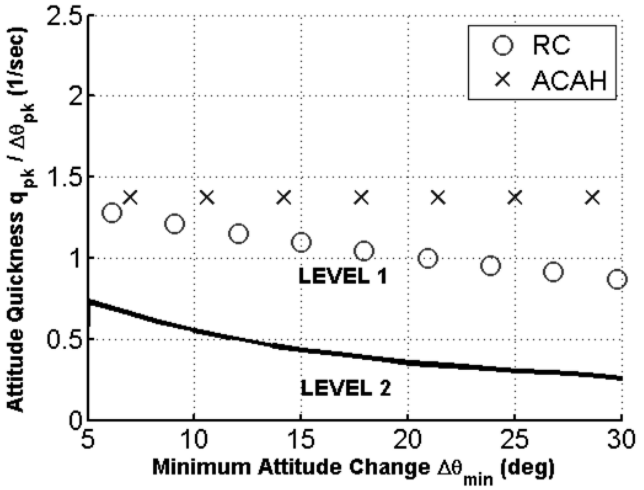

Fig. 20 Pitch Axis Attitude Quickness

In the roll axis, the bandwidth is shown in Fig. 21, and the attitude quickness is shown in Fig. 22. The attitude quickness results could not be matched so closely in roll as they were able to be matched in pitch – increasing quickness at smaller attitude changes (the left-hand half of Fig. 22) for the ACAH configuration would have resulted in the bandwidth increasing to very high values. This would have resulted in an aircraft that was extremely sensitive to small control inputs, which was considered to be undesirable for flight-naïve pilots. Conversely, reducing the quickness for smaller attitudes with the RC configuration would have resulted in the bandwidth becoming unacceptably close to the Level 1/Level 2 boundary.

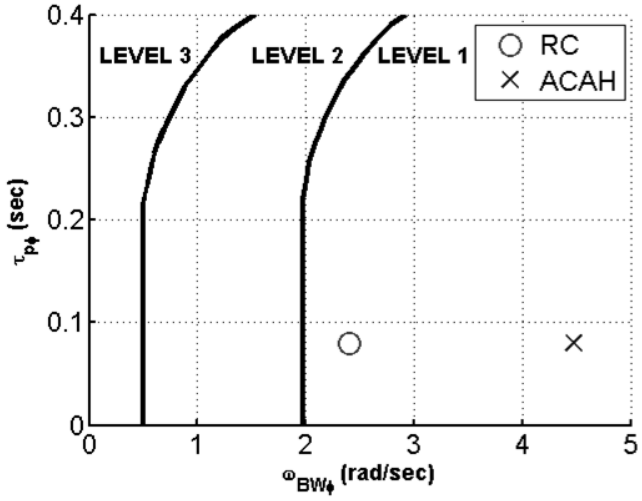

Fig. 21 Roll Axis Bandwidth

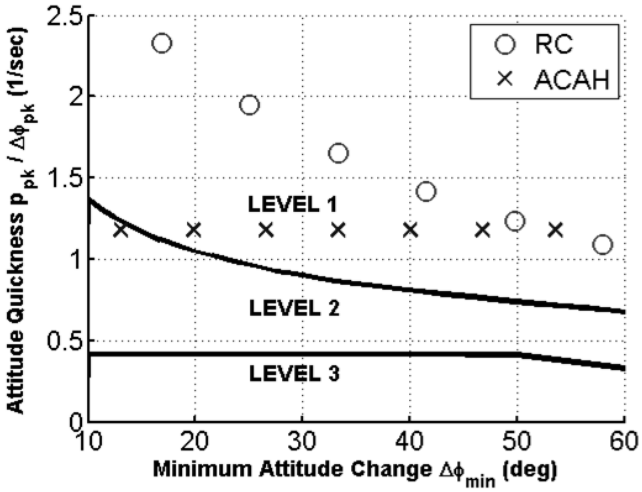

Fig. 22 Roll Axis Attitude Quickness

In yaw, all configurations employ the same RC response type in the hover. The bandwidth for this response is shown in Fig. 23, and the attitude quickness is shown in Fig. 24.

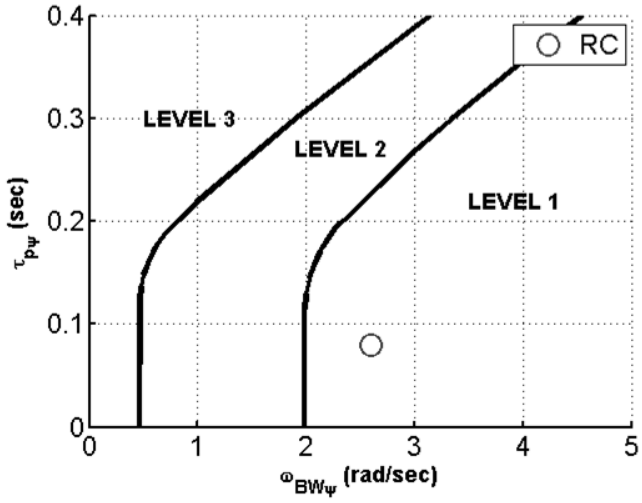

Fig. 23 Yaw Axis Bandwidth

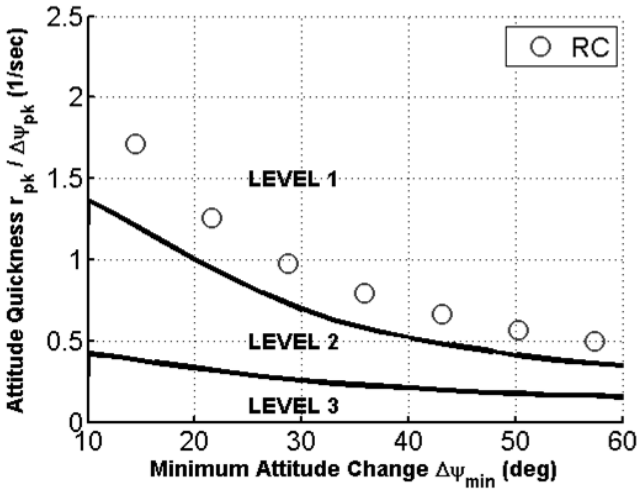

Fig. 24 Yaw Axis Attitude Quickness

The TRC response type of the Hybrid configuration is created through a velocity feedback loop around the ACAH dynamics described above. Therefore, the initial attitude response of the Hybrid configuration will be the same as that of the ACAH configuration. The velocity feedback loop has been configured to offer a rise time of 2.5 seconds in both the pitch and roll axes. The magnitude of the surge and sway velocity response for a given controller deflection is set as a constant 11ft/s/in for any deflection size. The rise times meet the ADS-33E-PRF Level 1 requirement for a TRC response type. The velocity gradient is somewhat higher than that required for Level 1 handling for low velocities, but is acceptable for higher velocities (ADS-33E-PRF recommends a non-uniform velocity gradient to improve sensitivity around hover). The constant velocity gradient has been adopted for this study to increase the predictability of the vehicle response to a change in control position for flight-naïve pilots.

Turning to piloted assessment of the three PAV configurations, Fig. 25 shows the HQRs (Ref. [29]) awarded in each of the five MTEs. The ratings were awarded by a single TP over the course of a one day simulation trial.

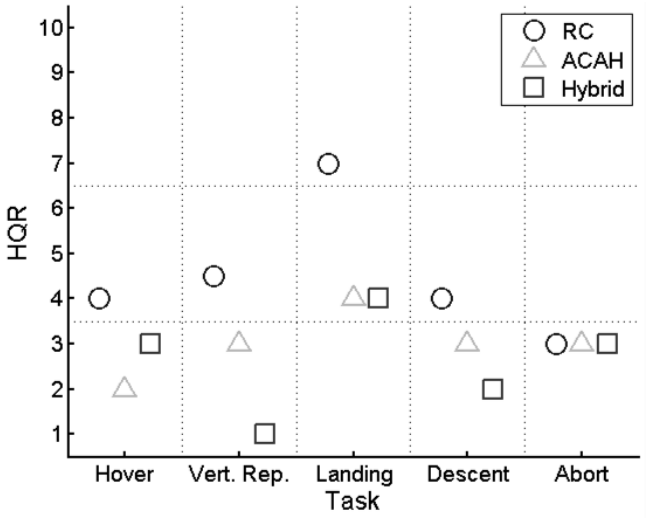

Fig. 25 PAV Handling Qualities Ratings

It can be seen in Fig. 25 that, despite all three configurations offering predicted Level 1 handling, a number of Level 2 HQRs and a single Level 3 HQR were awarded to the RC configuration. The TP found that, although desired performance could generally be achieved in all tasks apart from the Landing MTE, the workload associated with achieving this was higher than he considered desirable. The pilot commented on a number of occasions about a PIO susceptibility with the RC configuration – note the marginally Level 1 attitude bandwidth shown in Figs. 19 and 21 above, and hence a requirement to modify his control strategy to avoid exciting oscillations.

In terms of HQRs, the Hybrid configuration was shown to be as good as, or better than, the RC and ACAH configurations in all MTEs. The only exception to this was in the Hover MTE, but it should be noted that the HQR=3 awarded to the Hybrid configuration was during the TP's first exposure to this configuration, and it is possible that a lack of familiarity with its responses may have played a part in this rating.

The Landing MTE generally resulted in poorer HQRs than would be desired. The TP found the requirement to control the PAV's position within ±1ft longitudinally and ±0.5ft laterally to be demanding, even in the ACAH and Hybrid configurations. In the RC configuration, this level of accuracy could not be attained, resulting in the touchdown being made outside the position limits of the task. Note in Fig. 14 that the VCRs place the UCE close to the UCE=1/UCE=2 boundary for the Landing MTE. As the RC configuration will be more susceptible to UCE degradation than the other configurations (Ref. [14]), this may explain the larger difference in HQRs observed here compared to the other tasks.

The Level 1 HQRs suggest that the Hybrid configuration would be highly suited for use by a typical helicopter pilot of today. Further, with the exception of the RC configuration, the results also show that there is generally a good agreement between the predicted HQs according to ADS-33E-PRF and the assigned HQRs, serving to validate the GPDM and the wider simulation. In the case of the RC configuration, the TP was generally able to meet each task's desired performance standards, indicating that high precision was attainable, albeit at the expense of higher than desired workload. The improvement in HQRs as the response type is changed from RC through ACAH to TRC is as expected given the stability improvements accorded by the changes from rate to attitude, and from attitude to translational rate response types.

Although the results presented in Fig. 25 are from a single TP, a total of five other TPs have also taken part in HQ assessments during various stages of the development of the PAV simulation (Refs. [16,30]). While these TPs were not flying the final versions of each configuration as described in this paper, the results from these assessments show good correlation with the results for the final configurations presented in Fig. 25.

While the results in Fig. 25 indicate that the Hybrid configuration is highly suitable for current helicopter pilots, there is insufficient evidence upon which to draw conclusions regarding its suitability for the less experienced, flight naïve pilots who might be expected to be flying PAVs.

IV. PAV Handling Qualities Assessment Procedure

In order to determine HQ requirements for potential PAV 'flight-naïve' pilots (i.e. pilots with a broad spectrum of previous experience – ranging from JAR PPL(H) or UK PPL(A) holders and those learning to fly, to those with no previous flight experience), it is necessary to look beyond the traditional TP evaluation. It is proposed that, by assessing the performance of the flight-naïve pilots themselves in representative tasks with each candidate vehicle configuration, it is possible to make judgments regarding the suitability of those configurations for use in a PAV.

As the 'pilots' who took part in the assessments did not possess training in HQ evaluations, alternative approaches to those described above for the assessment of conventional rotorcraft with TPs had to be employed. Workload in each task was assessed subjectively through the NASA Task Load Index (TLX) rating (Ref. [31]). Task performance was then evaluated through a quantitative analysis of the precision with which the task was completed and the amount of control activity required to perform the task.

With this data in place, the suitability of the different candidate vehicle configurations could be assessed in terms of the skill required to meet various levels of performance, and hence also infer the amount of training required to be capable of operating the PAV safely and precisely.

While it is possible to broadly categorize these 'pilots' via their level of prior experience, it is to be expected that considerable variations in skill level would be evident within an experience tier. Therefore, each participant in the evaluations undertook a series of psychometric tests to determine their underlying aptitude towards flying before attempting the PAV tasks.

This Section of the paper describes the development of the various methods that have utilized in the evaluation of PAV handling requirements.

A. Aptitude

The suite of psychometric tests used to determine a subject's aptitude to pilot a PAV consisted of 9 separate computer-based tests examining different aspects of the piloting task. The tests were created at UoL from elements of the US Air Force Basic Attributes Test (Ref. [32]) and a kit of standard psychometric tests (Ref. [33]) to produce a broad assessment of the skills required to fly the PAV.

The 9 psychometric tests can be grouped into a number of categories relevant to the piloted simulation tests to be performed as follows: hand-eye coordination i.e. the ability to apply appropriate control inputs relative to visual stimuli (e.g. positional errors); visual (including pattern recognition, and spatial reasoning) i.e. the ability to develop spatial awareness; decisiveness i.e. the ability to make rapid decisions regarding the correct course of action; memory i.e. the ability to remember task instructions; and problem solving i.e. the ability to work out the correct control inputs for a given response type.

The 9 components of the *myCopter* aptitude test were as follows:

1. Two Handed Coordination – the test subject (TS) is required to track a circling target using separate controllers for horizontal and vertical position. This is a test of hand-eye coordination.
2. Complex Coordination – the TS is required to align a crosshair (vertical and horizontal motion) and a 'rudder bar' (horizontal motion only) in the face of continuous disturbances. One hand controls the crosshair, the other controls the rudder bar. As with the two handed coordination task, this is a test of hand-eye coordination.

3. Card Rotations – the TS is presented with a series of reference images together with derivations of that reference image. The subject must identify which of the derivations have just been rotated relative to the reference image, and which have been mirrored in addition to being rotated. This is a test of visual pattern recognition.
4. Dot Estimation – the TS is shown pairs of windows containing randomly dispersed dots. The subject must determine as rapidly as possible which of the pair of windows contains the greater number of dots. The dot estimation task is a test of a participant's decisiveness.
5. Identical Pictures – the TS is shown a series of reference images together with a group of candidate images. The subject must identify which one of the candidate images is identical to the reference image. This test examines a participant's visual pattern recognition and speed of mental processing capabilities – there are ninety six questions to be answered in three minutes.
6. Line Orientation – the TS is shown pairs of lines radiating from a central point. Using a reference array of lines, the subject must identify which of the reference lines correspond to the pair of lines. The line orientation task again examines pattern recognition abilities.
7. Locations – the TS is shown four lines each with a pattern of dashes and spaces. There is a single cross on each line. The subject must identify the pattern connecting the location of the cross on each of the lines, and apply that pattern to a fifth line to determine the location in which the cross would be found. This task examines a participant's problem solving ability.
8. Picture-Number Test – the TS is shown a set of pictures, and must memorize the numbers associated with each picture. The positions of the pictures on the screen are then shuffled, and the subject must recall the numbers that correspond to each picture. The picture-number test is a measure of a participant's memory capacity.
9. Shortest Roads – the TS is shown a series of images of three routes connecting two points on the screen. For each image, the subject must identify which of the three routes represents the shortest distance between the two points. The shortest roads test is a measure of a participant's spatial reasoning capabilities.

The methods used to compute the scores for each individual test, and to combine the results to produce a TS's overall aptitude score are described in the Appendix. From the nine tests, the theoretical maximum achievable score

is fifteen; scores closer to the maximum indicate a greater aptitude for the skills needed to successfully complete the PAV flight tasks.

Figure 26 shows the test scores achieved by 22 subjects (19 male and 3 female with an age range of 19-43, the mean of which is 25) who have taken the aptitude test. The TSs have also been broadly categorized by their prior flight experience:

- No Experience – these TSs have no prior experience of flight, either real or simulated;
- Simulator Experience – these TSs have experienced flight simulation, either on the desktop level through PC-based flight simulation games, or in a full flight simulator such as HELIFLIGHT-R.
- Flight Experience – these TSs have undergone some elementary flying training, and have generally achieved solo flight;
- Flight Qualified – these TSs have completed some form of elementary flying training (either military or civilian), and are qualified pilots. The most experienced pilot in this group has just over 200 hours of flight time.

The Figure shows that each category of pilot contains a reasonably broad range of aptitudes but that, as might be expected, the trend is for a higher aptitude amongst those with increasing flight experience. Also as might be expected, there is an overlap between the aptitudes demonstrated between the groups. Aptitude here is meant as the innate ability of a subject to perform a task. It is entirely possible that someone with a good aptitude for a subject has little experience in it (the former does not imply the latter), they simply have not yet been able to obtain any experience in it. For those with flight experience, the categorizations above make no distinction between those with fixed-wing experience and those with rotary-wing experience. The vast majority of the TSs came from a fixed-wing background. It is, however, interesting to note that the two pilots with a rotary-wing background achieved the two highest aptitude scores amongst those TSs in the 'Flight Qualified' category. Given that these two subjects have reached this level of qualification, it may be expected that they possess natural aptitude towards the operation of vertical flight vehicles. Their high scores in the aptitude test described in this paper and the trends previously described therefore provide confidence that the aptitude testing methodology is returning an appropriate and useful grading of the test subjects.

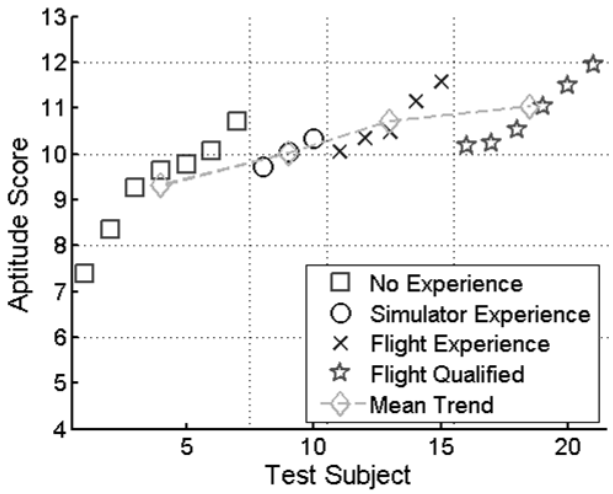

Fig. 26 Aptitude Test Scores

The impact of TS aptitude on flying ability can be seen in Fig. 27. This Figure shows the average time spent within the desired performance boundaries of the five MTEs defined above, for a number of TSs of varying aptitude (each point represents an individual TS). The TSs were flying the RC configuration. The result for one of the TSs (A=10.3) falls well above the trend line for the other TSs. This TS had considerable previous simulator experience with vehicles which exhibit responses similar to those of the RC configuration, and it is possible that this aided them with understanding the demands of this particular exercise. Apart from this TS, it is clear that increasing aptitude closely correlates with increased ability to complete the PAV MTEs more precisely. This provides evidence that the aptitude assessment process described here is an effective discriminator to identify the flying ability of flight-naïve TSs.

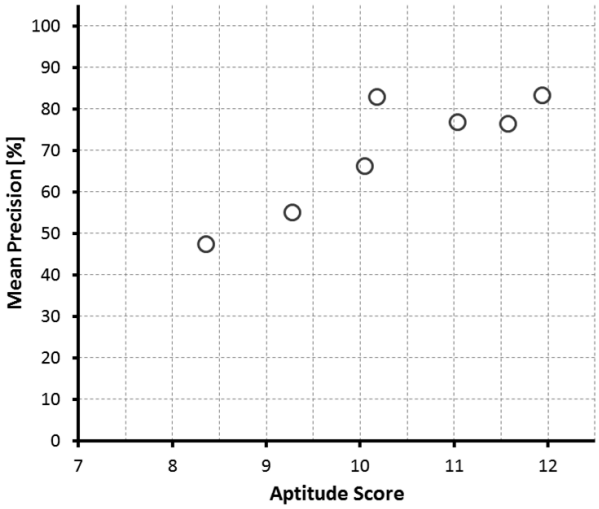

Figure 27. Improvement in Attained Precision with Increased Aptitude for RC Configuration

B. Task Load Index

The Task Load Index (TLX, Ref. [31]) is a workload rating system developed by NASA. It was designed to be applicable to the assessment of the workload involved in any task, and for it to be straightforward for new users to understand the concepts and processes involved in its use.

The TLX rating involves the assessment of six aspects of workload – mental demand; physical demand; temporal demand; performance; effort and frustration. The ratings for each of these aspects are then combined together using a weighting system, in which the TS compares each of the workload elements to the other elements and decides in each case which represented the greater contribution to the overall workload of the task. This process allows a single workload score for each task to be produced. The final TLX rating lies in the range $0 < TLX \leq 100$, where lower numbers indicate a reduced workload when compared to higher TLX scores.

The TLX rating, as a measure of the workload associated with a task, provides an indication of the extent to which a TS is comfortable with the demands of a given configuration. This is illustrated in Fig. 28, which shows the mean TLX rating awarded by a group of TSs to the ACAH configuration in the five PAV MTEs. It can be seen that there is a close relationship between the aptitude of a given TS and the workload that they associate with the tasks – the higher the aptitude, the more straightforward the TS is likely to find this configuration. As with Fig. 27, outlying data is present in Fig. 28 in the A=10.2 region. Again, it is believed that this is a result of previous exposure of these TSs to response characteristics similar to those of the ACAH PAV configuration.

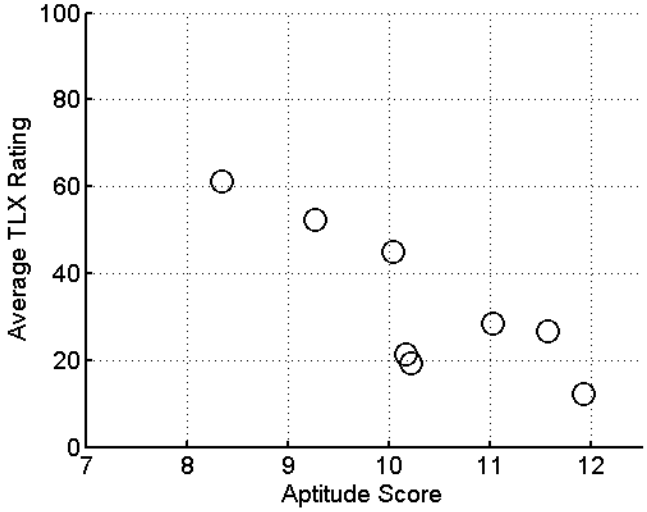

Fig. 28 Effect of Aptitude on Awarded TLX Ratings with ACAH Configuration

C. Task Performance Assessment

For the quantitative assessment of task performance, two key parameters have been identified. The first of these is the accuracy with which a given MTE could be performed. This has been measured as the percentage of time spent within each of the MTE's desired performance boundaries. The results for each performance requirement are averaged to produce an overall precision rating (P) for an MTE. Higher P values correspond to more accurate performance in the task. Values of P range from 100% if the vehicle remains within all desired performance limits for the duration of the task, to 0% if any of the desired performance limits are never entered at all during the task.

The second parameter is a quantitative measurement of the task workload (W), captured in terms of the amount of control activity required to complete an MTE. While this can be measured in many ways (for example cut-off frequency analysis (Ref. [34]), attack analysis (Ref. [35]) etc.), the objective technique used to assess workload for the trials reported in this paper was to count the number of discrete movements of the controls (above a threshold of 0.5% of full stick deflection – this is implemented to prevent measurement noise from affecting the analysis), and averaged against the time required to complete the task – giving a number of control inputs made per second in each axis. This metric has been found to be sensitive to pilot control strategy and reflective of a pilot's subjective opinion of the physical workload associated with a task (Ref. [35]). The control input rate is averaged across the four control axes to produce a single value for each MTE. With the control activity evaluation, the ability to perform a task with fewer control inputs is preferable, as this equates to a lower amount of pilot effort. For flight-naïve pilots, minimal control effort is expected to be key to ensuring safe, reliable operation of the PAV. Typical workload measurements fall in the range $0.1/\text{sec} < W < 1/\text{sec}$ depending on the nature of the task, the configuration being flown and the pilot's strategy.

It is acknowledged that a metric such as this does not capture all aspects of a pilot's workload, and that there can in fact be cases where a low amount of control activity correlates with a high workload (a good example of this would be a situation where a large time delay is present in a system – the pilot then has to apply considerable mental effort to reduce their control activity in order to prevent the excitation of pilot induced oscillations). However, it is considered that the benefits of having a single metric to capture a basic representation of the workload outweigh these disadvantages, provided that the subjective workload assessments are also considered to ensure that the correlation between low workload and low control activity holds.

Taking the quantitative analysis a stage further, it is possible to combine the metrics used to assess precision and workload into a single metric to represent the overall performance achieved in a given MTE. At a basic level, the precision and control activity metrics can be combined directly:

$$\text{performance} = \frac{P}{W} \tag{5}$$

However, if it is considered that ability to achieve an MTE's desired performance requirements is of greater importance than achieving a minimal workload for a given task, the relative weighting of precision and workload metrics can be adjusted so that the overall metric becomes:

$$\text{performance} = \frac{P^2}{\sqrt{W}} \tag{6}$$

Finally, it is possible to define a theoretical maximum value for each of the precision and workload metrics for each MTE, and hence a maximum value for the overall performance metric. Maximum precision should be 100% time spent within the desired performance requirement in every case. Theoretical minimum workload can be computed by determining the fewest control inputs required to complete a given MTE. To give an example for the hover MTE and Hybrid configuration, assuming that the HH and DH functions are engaged, one movement of the cyclic at 45° is required to initiate the translation to the hover point, and a second movement is required to decelerate to the hover. As each of these inputs occurs in both the longitudinal and lateral cyclic axes, there are a minimum of four discrete inputs required to complete the hover task (assuming that no control activity is required during the stable hover phase). By estimating a duration of 40 seconds (10 seconds of translation followed by 30 seconds of stable hovering), the average number of control inputs per second can be computed as 0.025/sec in each axis.

These theoretical maximum performance values for each MTE are used to normalize the values of achieved performance. This has been called the Task Performance Index (TPX) for this paper:

$$TPX = \frac{P^2\sqrt{W_{min}}}{100^2\sqrt{W}} \tag{7}$$

With the TPX, a rating of 1.0 means that the pilot was able to achieve maximum precision in the task through the use of the minimum possible control effort. TPX ratings of less than 1.0 indicate that either the control effort was higher, or the precision lower, than would be ideal.

To improve the statistical reliability of the TPX scores, average values of *P* and *W* are calculated using the final three attempts at a given task by each TS. The TPX is therefore representative of the overall 'experience' of flying a given task, rather than a snapshot of the events of a single run.

Figure 29 shows a comparison of the measured TPX scores and awarded TLX ratings taken during the HQ assessment process in the *myCopter* project. A total of 209 individual test points are represented, flown by 13 flight-naïve TSs. Notwithstanding the comments above regarding the *W* component of the TPX score only considering physical, rather than mental workload, and the likelihood of variation in the TS-to-TS interpretation of the TLX rating scale, a relationship can be seen between the TPX scores and the TLX ratings. Plotting the computed TPX scores on a logarithmic scale yields a reasonably linear relationship with the subjectively awarded TLX ratings. An exponentially-shaped best fit to the data (shown as the dashed line on Fig. 29) has been found as:

$$TPX = 0.54e^{(-0.033 \times TLX)} \tag{8}$$

where the coefficient of determination (Ref. [36]), $R^2 = 0.988$. A greater scatter in the points is visible in Fig. 29 at higher TLX ratings – this is partly due to the logarithmic presentation of the TPX scores (resulting in the significance of a small change in TPX at already small TPX scores being increased), and partly due to the higher TLX scores being associated with test subjects struggling to achieve the specified tasks – and hence potentially returning ratings that do not truly reflect performance in the task.

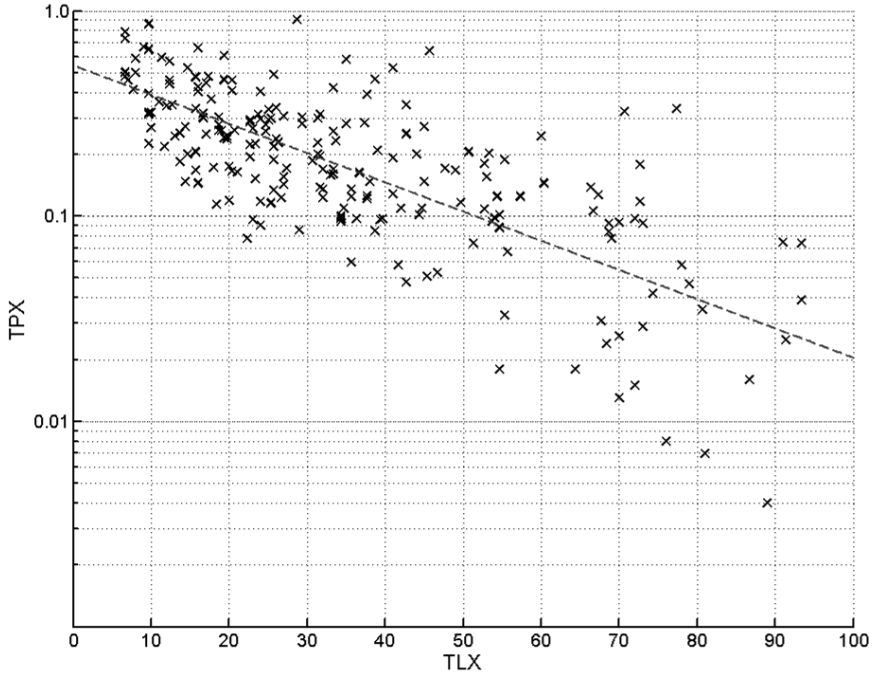

Fig. 29 Comparison of TPX Scores with TLX Ratings

The TLX rating is predominantly intended to function as a measure of the workload experienced when completing a specified task. However, the presence of 'performance' as one of the six contributory factors to the overall workload means that the evaluator is considering the level of success attained in a task as part of the process

of awarding a TLX rating. The incorporation of both *P* and *W* into the TPX means that the metric is evaluating similar factors to the TLX rating (albeit with an increased emphasis placed on performance relative to workload). Hence, the existence of a coherent relationship between the TLX rating and the TPX metric demonstrates that TPX is an effective measure of the performance of a TS in a given flight task.

The use of the TPX metric is illustrated in Fig. 30, which shows scores achieved by a group of flight-naïve pilots flying the Hover MTE using the ACAH configuration in DVE conditions with turbulence. It can be seen that there was a progressive increase in achieved performance as the aptitude of the TS increased. However, there are also two notable deviations from this trend, with two TSs in the A=10.3 region achieving a higher level of performance than any of the other TSs. Both of these TSs had considerable previous simulator experience with vehicles similar to the ACAH configuration, and it is possible that this aided them with understanding the demands of the particular exercise.

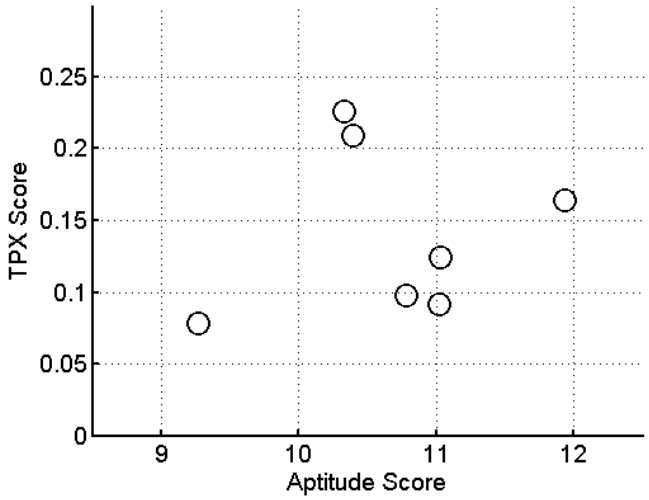

Fig. 30 Sample TPX Scores for Hover MTE with ACAH Configuration in Harsh Environment

D. Determination of the Suitability of Candidate PAV Configurations

The results presented above indicate that, despite the generally good HQRs awarded by the TP to the PAV configurations, all of the flight-naïve TSs were not able to complete the MTEs to an acceptable standard. The low level of achieved precision and the high reported workload for some TSs do not concur with the HQRs awarded by the TP.

The methods described in the preceding Sections can be used to compare the different configurations. By considering the relationship between aptitude and the various metrics (TLX, *P*, TPX), it is possible to determine the

range of *A* across which a given configuration can be flown 'successfully' – i.e. safely and repeatably perform the PAV MTEs to the required level of accuracy. As the suitability of a given configuration for use in a PAV increases – in other words, the more straightforward and intuitive the response characteristics are for flight-naïve pilots to learn and master – TSs with progressively lower *A* will be able to operate the PAV with similar levels of performance as the TSs with high *A*. The results from such assessments of the candidate PAV configurations will be reported in the earlier-mentioned future companion publication.

V. Conclusions

This paper has briefly described the activities of the *myCopter* research project and the development of a simulation environment to allow the assessment of PAV handling qualities requirements and training needs. The paper has also described the development of an aptitude assessment process designed to determine the flying abilities of 'flight-naïve' test subjects, and quantitative metrics for the assessment of performance and workload in Mission Task Elements that these subjects have been asked to fly in simulated flight vehicles with response types configured to behave as though they are equipped with different levels of augmentation.

The main conclusions that can be drawn from the paper are:

- The conventional handling qualities assessment process is insufficient for the analysis of PAV handling requirements. Despite good HQRs being awarded by a test pilot for all configurations, in many cases flight-naïve test subjects were unable to perform to a comparable level of precision, and their corresponding level of workload was higher than desirable.
- A computer based aptitude test battery has been created, and has been shown to be a good predictor of the ability of flight-naïve test subjects to achieve precise vehicle control in flight tasks.
- As aptitude score increases, there is an expected strong correlation with perceived reduction in workload, showing that the Task Load Index (TLX) rating can be effectively utilized by flight-naive test subjects.
- A Task Performance Index (TPX) metric has been created. A coherent relationship has been shown to exist between the recorded TPX and subjective TLX values, showing that TPX provides a valid method for the objective assessment of workload and task performance of flight-naïve test subjects.

VI. Appendix

A. CETI Turbulence Model Parameters

This Appendix details the settings used for the PAV CETI turbulence model. For longitudinal, lateral and pedal inputs, the structure of the turbulence filter is:

$$\frac{\delta_{gust}}{W_{noise}} = A\frac{1}{\left(s+\frac{U_0}{L}\right)} \tag{A1}$$

For the longitudinal filter, the settings were:

A = 2.29

U_0/L = 1.13

For the lateral filter, the settings were:

A = 2.33

U_0/L = 1.13

For the pedal filter, the settings were:

A = 4.68

U_0/L = 4.00

For collective inputs, the structure of the turbulence filter is:

$$\frac{\delta_{gust}}{W_{noise}} = A\frac{\left(s+20\frac{U_0}{L}\right)}{\left(s+0.63\frac{U_0}{L}\right)\left(s+5\frac{U_0}{L}\right)} \tag{A2}$$

The settings used were:

A = 0.153

U_0/L = 3.85

B. Aptitude Test Details

This Appendix provides additional details regarding the nine individual components of the aptitude test battery.

1. Two-Handed Coordination

The two-handed coordination test requires the TS to align the crosshair with the target (the aircraft symbol), using a pair of controllers – one determines the vertical position of the crosshair, the other the lateral position. The test subject is given an unscored practice period of 3 minutes, following which the assessment takes place over a period of 5 minutes.

Scoring for the two-handed coordination test is based on the proximity of the crosshair to the target. A proximity score at each time step is calculated as follows:

1) The distance (measured in number of pixels on the screen) between the crosshair and the target is computed

2) This distance is normalized using the vertical screen resolution (in pixels)

3) The score is computed as *proximity_score = 1 – normalized_distance*

The overall score for the task is computed as the numerical mean of the individual scores at each time step.

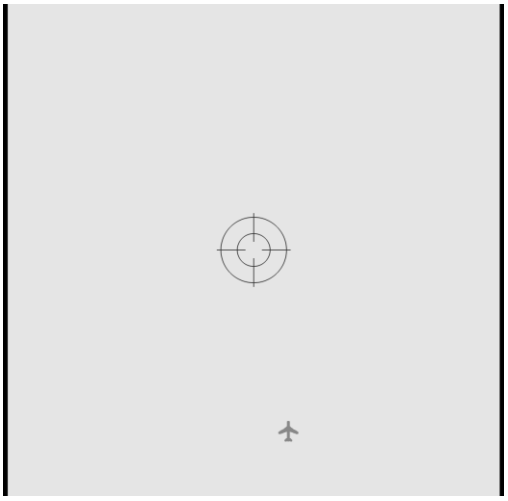

Fig. B.1 Two-Handed Coordination Test

2. Complex Coordination

The complex coordination test requires the TS to align a crosshair and a rudder bar with a pair of reference markers. A pair of controllers is used – one determines the position of the crosshair (both vertical and lateral), the other the lateral position of the rudder bar. The crosshair and rudder bar are disturbed from their reference positions by pseudo-random signals created as a sums of sine waves. The test subject is given an unscored practice period of 3 minutes, following which the assessment takes place over a period of 5 minutes.

Scoring for the complex coordination test is based on the proximity of the crosshair and rudder bar to their targets. A proximity score at each time step is calculated as follows:

1) The separation (in pixels) between each controlled parameter (i.e. crosshair vertical; crosshair horizontal and rudder bar horizontal) and the reference position is calculated

2) The root-mean-square of the three separations is calculated, giving an ‘average’ position error

3) This error is normalized using the vertical screen resolution (in pixels)

4) The score is computed as *proximity_score = 1 – normalized_error*

The overall score for the task is computed as the numerical mean of the individual scores at each time step.

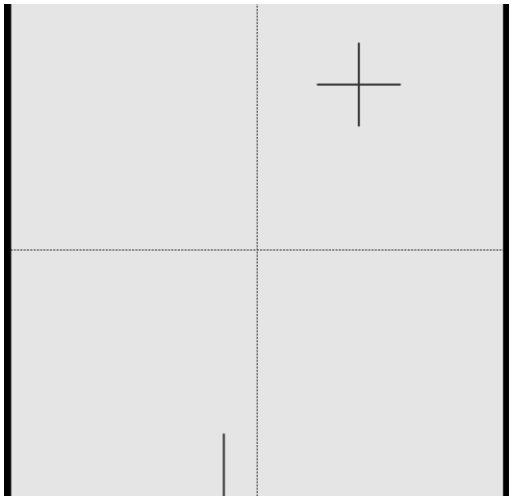

Fig. B.2 Complex Coordination Test

3. Card Rotations

The card rotations test presents the test subject with one reference shape and eight candidate shapes. The candidate shapes are rotated and/or mirrored relative to the reference shape. The test subject is required to identify all of the candidate shapes that are only rotated (i.e. are not mirrored). A total of 28 questions are attempted, and the test subject has 8 minutes to complete the test (a one minute break is provided after 14 questions). One point is gained for finding all of the correct answers to a question, while one point is lost for failing to find all of the correct answers.

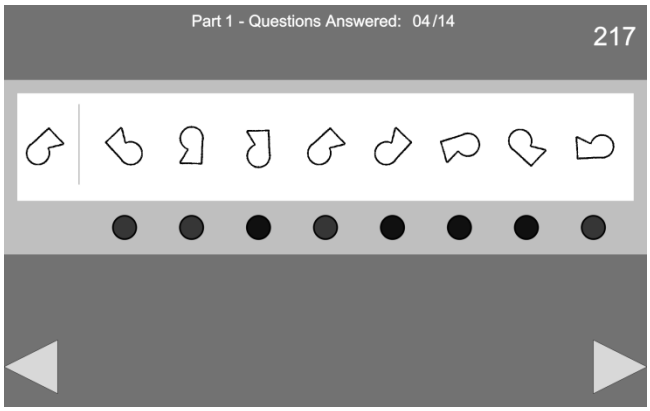

Fig. B.3 Card Rotations Test

4. Dot Estimation

The dot estimation task requires the test subject to make a judgement regarding which of a pair of boxes contains a greater number of 'dots'. As the test progresses (with 55 pairs of boxes to assess in total), greater numbers of dots

populate each of the boxes. For each pair, one of the boxes contains one additional dot compared to the other box. At the start, one box contains 10 dots and the other 11. By the end of the test, each box will contain either 50 or 51 dots.

The score for the dot estimation task considers both the accuracy of the judgements made and also the time taken to make the judgements. The total number of correct answers is summed, and this is divided by the average time required by the test subject to decide on their answer for each pair of boxes. Hence, slow decision-making (such as counting all fifty dots in each box at the end of the test) is penalized.

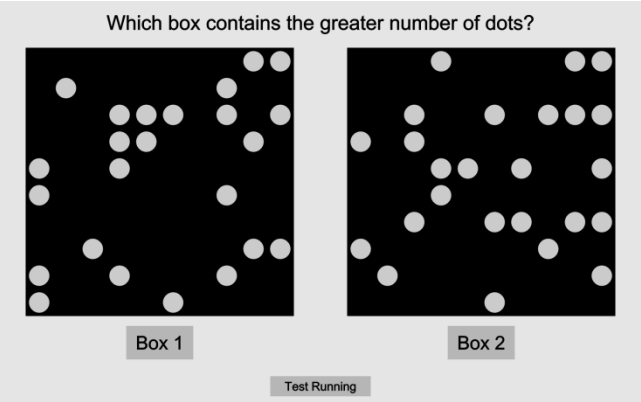

Fig. B.4 Dot Estimation Test

5. Identical Pictures

The identical pictures test requires the test subject to determine which of five candidate shapes is identical to a reference shape. The correct shape is not rotated or mirrored. The test consists of 96 questions, and the test subject has three minutes to answer them all. A thirty second break is provided after the first 48 questions.

For each correct answer, a test subject's score is increased by one point. An incorrect answer results in a penalty of 0.5 points.

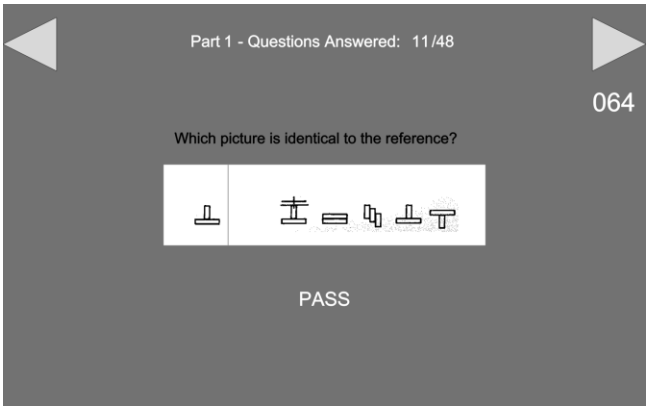

Fig. B.5 Identical Pictures Test

6. Line Orientation

The line orientation test requires a test subject to match the orientations of two candidate lines with a grid of reference lines. The test consists of thirty questions, and the test subject has two minutes to complete the test. For each correct answer (i.e. both candidate lines correctly matched), one point is scored. An incorrect answer results in a penalty of 0.5 points.

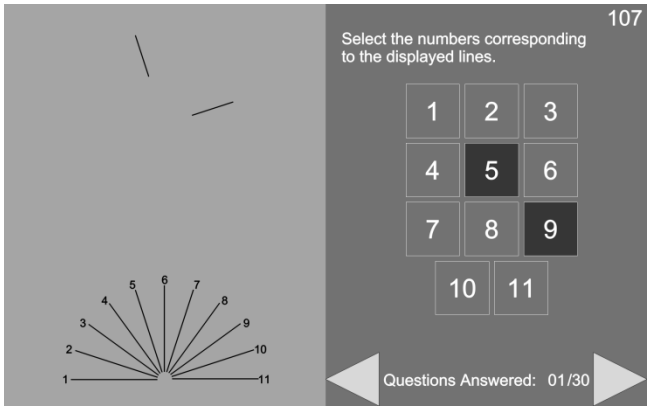

Fig. B.6 Line Orientation Test

7. Locations

The locations test presents the test subject with four patterns, consisting of dashes and spaces. In each pattern, one X is placed instead of a dash according to a certain rule. The test subject must identify the rule and then apply it to a fifth pattern in order to determine which of five potential positions is correct for the X. The test consists of 28 questions, and the test subject has 12 minutes to complete the test (with a one minute break after the first 14 questions). One point is scored for each correct answer in this test, with a 0.5 point penalty applied for an incorrect guess.

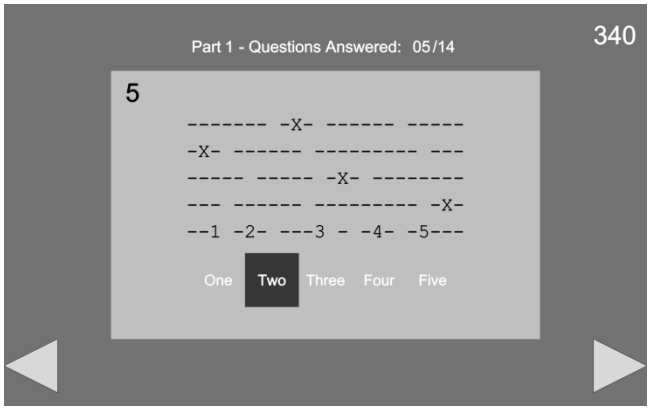

Fig. B.7 Locations Test

8. Picture-Number Test

The picture-number test assesses a test subject's memory. A screen of drawings of everyday objects is presented to the test subject; each drawing has a number associated with it. This screen is displayed for four minutes, during which time the test subject attempts to memorize as many of the combinations of object and number as possible. At this point, the numbers are removed and the order of the pictures rearranged. The test subject then has three minutes to enter the numbers that correspond to each object. For each correctly memorized and recalled combination of object and number, one point is scored. This test does not penalize incorrect guesses.

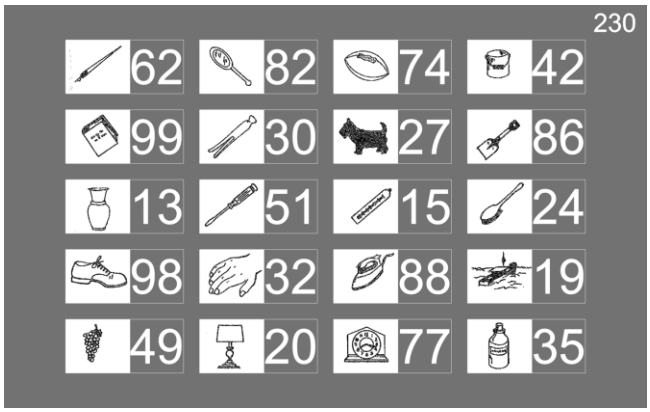

Fig. B.8 Picture-Number Test

9. Shortest Roads

In the shortest roads test, three routes between a 'start' point and an 'end' point are shown. The test subject is required to identify which of the three routes represents the shortest distance between the two points. This test consists of 56 questions, and the test subject has four minutes to answer all of the questions. A one minute break is provided after the first 28 questions. One point is scored for each correct answer in this test. An incorrect answer is penalized by a 0.5 point deduction.

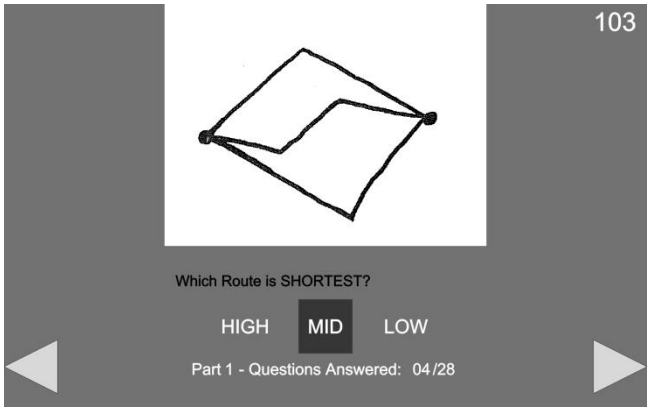

Fig. B.9 Shortest Roads Test

10. Determination of Overall Aptitude Score

The first stage in the determination of the overall aptitude score is normalization of the scores for each of the individual tests. This is performed by dividing the test score by the theoretical best score for each test. Following this step, all nine test scores are measured as a fraction, with 1 being the best score. The scores for the two-handed coordination and complex coordination tests are weighted by multiplying by 4. The final aptitude score is then found as a sum of the weighted, normalized scores for each individual test. The maximum achievable score in the overall test is 15 points.

VII. Acknowledgements

The work reported in this paper is funded by the EC FP7 research funding mechanism under grant agreement no. 266470. The authors would like to thank all those who have participated in the simulation trials reported in this paper for their contribution to the research.

VIII. References

1. *Out of the Box: Ideas about the Future of Air Transport*. Part 2, European Commission, Brussels, November 2007
2. *Transport Trends: 2009 Edition*, Department for Transport, London, 2009
3. *Energy, Transport and Environment Indicators, Eurostat Pocketbook*, EUROSTAT, Publications Office of the European Union, Luxembourg, 2009
4. *Green Paper – Towards a new culture of urban mobility, COM(2007) 551 final*, Commission of the European Countries, Brussels, September 2007
5. Jump, M. *et al*, *myCopter: Enabling Technologies for Personal Air Transport Systems*, Royal Aeronautical Society conference "The Future Rotorcraft – Enabling Capability through the application of technology", London, UK, June 2011
6. *Aerocar 2000*, 2002 [cited 2013 10th April]; Available from: http://www.aerocar.com/.
7. *Carplane Road/Air Vehicle*, 2013 [cited 2013 10th April]; Available from: http://www.carplane.com/.
8. *Terrafugia*, 2012 [cited 2013 10th April]; Available from: http://www.terrafugia.com/.

9. *PAL-V. Personal Air and Land Vehicle*. 2013 [cited 2013 10th April]; Available from: http://www.pal-v.com/.
10. Carter, J., *CarterCopter -- A High Technology Gyroplane* in *American Helicopter Society Vertical Lift Aircraft Design Conference*. 2000, AHS: San Francisco, CA.
11. *Moller International*, 2013 [cited 2013 10th April]; Available from: http://www.moller.com/.
12. *Urban Aeronautics X-Hawk Aerial Vehicle*, 2013 [cited 2013 10th April]; Available from: http://www.urbanaero.com/category/x-hawk.
13. Jump, M. *et al*, *myCopter: Enabling Technologies for Personal Air Transport Systems – An Early Progress Report*, Proceedings of the 37th European Rotorcraft Forum, Gallarate, Italy, September 2011
14. *Handling Qualities Requirements for Military Rotorcraft*, ADS-33E-PRF, USAAMC, March 2000
15. White, M.D., Perfect, P., Padfield, G.D., Gubbels, A.W. and Berryman, A.C., "Acceptance testing and commissioning of a flight simulator for rotorcraft simulation fidelity research", *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, Volume 227 Issue 4 April 2013 pp. 655 – 678, 2013
16. Perfect, P., Jump, M. and White, M.D., *Development of Handling Qualities Requirements for a Personal Aerial Vehicle*, Proceedings of the 38th European Rotorcraft Forum, Amsterdam, Netherlands, September 2012
17. *Presagis Vega Prime*, 2014 [cited 2014 3rd March]; Available from: http://www.presagis.com/products_services/products/modeling-simulation/visualization/vega_prime/
18. *Garmin G1000*, 2014 [cited 2014 4th June]; Available from: https://buy.garmin.com/en-GB/GB/aviation/flight-decks/g1000-/prod6420.html
19. *Peripheral Vision Horizon Display*, NASA Conference Publication 2306, November 1984
20. Chuang L.L., Nieuwenhuizen F.M. and Bülthoff H.H., *Investigating Gaze Behavior of Novice Pilots during Basic Flight Maneuvers*, International Conference on Human-Computer Interaction in Aerospace, Brussels, Belgium, 2012
21. Schönenberg, T., *Design of a Conceptual Rotorcraft Model Preparing Investigations of Sidestick Handling Qualities*, Proceedings of the 67th Annual Forum of the American Helicopter Society, Virginia Beach, VA, USA, May 2011

22. *Guide to Visual Flight Rules (VFR) in the UK*, UK Civil Aviation Authority, London, 2011
23. *TuTiempo.net*, 2014 [cited 2014 27th February]; Available from http://www.tutiempo.net/en/Climate/
24. Roth, M., "Review of Atmospheric Turbulence over Cities", *Q.J.R. Meteorol. Soc.* Vol. 126, Issue 564, pp. 941-990, 2000
25. Padfield, G.D., *Helicopter Flight Dynamics*, Second Edition, pp. 346-348, Blackwell Science, 2007
26. von Karman, T., *Progress in the Statistical Theory of Turbulence*, Turbulence: Classical Papers in Statistical Theory, Interscience Publications, New York, 1961
27. Seher-Weiss, S. and von Gruenhagen, W., "Development of EC 135 Turbulence Models via System Identification", *Aerospace Science and Technology*, Vol. 23, pp. 43-52, 2012
28. *General Handling Flight Test Requirements – Subjective Definition of Turbulent Air Standards*, DEF STAN 00-970 Vol. 2 Leaflet 900/4, October 1986
29. Cooper, G.E. and Harper, Jr. R.P., *The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities*, NASA TN-D-5153, April 1969
30. Perfect, P., White, M.D. and Jump, M., *Towards Handling Qualities Requirements for Future Personal Aerial Vehicles*, Proceedings of the 69th Annual Forum of the American Helicopter Society, Phoenix, AZ, May 2013
31. Hart, S.G. and Staveland, L.E., *Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research*, In P.A. Hancock and N. Meshkati (Eds.) Human Mental Workload. Amsterdam: North Holland Press, 1988
32. Carretta, T.R., *Basic Attributes Tests System: Development of an Automated Test Battery for Pilot Selection*, AFHRL-TR-87-9, US Air Force Human Resources Laboratory, September 1987
33. French, J.W., Ekstrom, R.B. and Price, L.A., *Manual and Kit of Reference Tests for Cognitive Factors*, Educational Testing Service, 1963
34. Blanken, C.L. and Pausder H.-J. "Investigation of the effects of bandwidth and time delay on helicopter roll-axis handling qualities", *Journal of the American Helicopter Society*, July 1994, 39, (3), pp. 24-33
35. Perfect, P., White, M.D., Padfield, G.D. and Gubbels, A.W., "Rotorcraft Simulator Fidelity: New Methods for Quantification and Assessment", *The Aeronautical Journal*, Vol. 117, Number 1189, pp. 235-282, March 2013

36. Rao, C.R., *Linear Statistical Inference and its Applications*, 2nd Edition, Wiley, New York, 1973