Increasing Workload on Simulated Remotely Piloted System Interaction and Task Completion – Gamers versus Non-Gamers

by

George A. H. Bedford

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of

Doctor of Philosophy of Loughborough University

(2015)

Abstract

With the current high rate of development and deployment of Remotely Piloted Aerial Systems (RPAS) for both commercial and military sectors globally, it is key to understand the implications this technology has on current and future RPAS operators and the consequential effect on licensing, training and performance measurement.

This thesis investigates aspects of training and potential objective performance measurement of RPAS operators, this is carried out by reviewing current literature relating to RPAS and associated human factors thus a gap analysis was undertaken and a set of experiments/evaluations were devised to provide important new insights. Attention is drawn to the type of skill set required for future RPAS operations. A factor has been to understand whether a regular computer games player displays differing simulator interaction, in this case information gathering and analysis patterns, to that of someone with limited to no computer games experience.

To achieve the aims of the research experimentation had to be carried which required the development of an appropriate simulator followed by the inclusion of a case study and the creation of bespoke performance data analysis software, SimPACT.

Although performance differentials have been observed through action it was hoped to be able to identify performance differential characteristics through the means of evaluating the use of disparate physical data sets; the research, in fact, identified no significant difference between data set use and it must be concluded that any pre-action performance differential cannot be measured, at least not with the equipment available. However computer gamers, rather than having differing information acquisition strategies, have differing and more effective information retention and processing pathways likely to have been developed through continuous gaming which can be applied to any game-type environment and, potentially, any type of interactive task. These results have been proven to be statistically viable and observable.

This research has contributed to the understanding of human performance measurement within the RPAS sector, including the addition of new data processing software, as well as provide new evidence relating to difference within human data gathering and processing between groups of differing experiences.

Acknowledgements

I must take the opportunity to thank all the people who have aided this research with both support and knowledge, without these people this research may never have seen the light of day.

Prof. Roy Kalawsky

For the advice, support and direction as well as the stories and recovery strategies. Also not forgetting the occasional kick up the behind to keep me moving forward.

Mr. Luan Le-Ngoc, Mr. Craig Wright

For being fantastic sounding boards as well as being a wealth of interdisciplinary knowledge.

Dr. John O'Brien, Dr. Petri Vitiello

For passing on their experience and valuable thesis knowledge.

Mr. Mike Southworth

For the invaluable industry knowledge and contacts and always being available for a chat.

The Royal Air Force (RAF)

for their knowledge and advice during the course of the research

My Parents, Mr. Tony Bedford & Mrs. Jane Bedford

For giving me the support, drive and opportunity to attempt this study

and finally to Miss Emma Bush

for all the love and support that she has given me during the entirety of this research. Without her understanding and patience this acknowledgements page, as well as the thesis, may never have been written.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AADA - Acquire, Analyse, Decide, Act

ACL - Autonomous Control Level

AFRL - Air Force Research laboratory

AVRRC - Advanced Virtual Reality Research Centre (Loughborough University)

BTT - Basic Training Target

CAA - Civil Aviation Authority

CGCS - Common Ground Control Station

COTS - Commercial-Off-The-Shelf

CPU - Central Processing Unit

CSV - Comma Separated Variable text file

DLM - Delimited text file

DUO - Designated UAV Operator

DVLA - Driving and Vehicle Licensing Agency

EBT - Evidence based training

FWV - Fixed-Wing Vehicle

GCS - Ground Control Station

GUI - Graphical User Interface

HALO - High Altitude Low Observability

HMI - Human Machine Interface

HUD – Heads Up Display

IAI - Israel Aircraft Industries

IAT - Integrated Aircrew Training

KSE - Knowledge, Skill and Experience

LoA - Level of Autonomy

MAP - Mobility, Acquisition and Protection

MEC - Mission Essential Competency

MoD -Ministry of Defence (UK)

NASA - National Aeronautics and Space Administration

NATO - North Atlantic Treaty Organization

OODA - Observe, Orient, Decide Act

PETS - Performance Effectiveness/Evaluation Tracking System

R/C - Remote Control

RAF - Royal Air Force

RMSE - Root Mean Squared Error

RPAS - Remotely Piloted Aircraft System

SAM - Surface to Air Missile

SC - Supporting Competency

SME - Subject Matter Expert

SSD - Solid State Drive

STANAG - Standard Agreement (NATO)

SWAT - Subjective Workload Assessment Technique

UAV - Unmanned Aerial Vehicle

UCAV - Unmanned Combat Aerial Vehicle

UDP - User Datagram Protocol

USAF - United States Air Force

Chapter 1 Introduction

This chapter introduces the purpose of this thesis and goes on to outline the research aims and objectives and the relevant research questions to be answered whilst also presenting potential hypothesis associated with the research questions.

1.1 Overview

The Remotely Piloted Aircraft System (RPAS), also known as the Unmanned Aerial Vehicle (UAV), are rapidly becoming widely used platforms in both the commercial and military sectors [1]. With an increase of use of these systems there has also been an increase in the demand for, and subsequent training, of operators for these system*s* [2]. Until recently only experienced military pilots were being considered for RPAS operator roles, this pool of available military personnel of this type is often small and this creates a recruitment problem [3].

With the need for the use of drones within the military sector increasing, it has been identified that the potential pool of viable military candidates was not sufficient to meet the demands of this increasing sector [2], this culminated in the USAF introducing a new training pipeline for RPAS operators which no longer requires seasoned pilot experience as a prerequisite. Pilots are now being trained specifically to fly drones.

This research initially investigates a hypothetical training structure based around the Mission Essential Competency (MEC).

This structure is utilised by applying to a performance measurement system; this system relies upon the observation of the way in which participants use different types of information to carry out a specific task. To understand what different types of information are observed, head tracking is used to understand the orientation of the participants gaze in relation to the position of a specific information set.

Although the objective head-tracking information may indicate that a participant is viewing a certain information set, this does not mean that the participant is actually using that information set as a current source of information. To understand and compensate for this, stepped increases in workload were used to force the participant to, eventually, only focus upon task and time relevant information sets. In this case workload is represented by the amount of control input required of the participant to be able to perform the task; this not only increases the amount of physical input required but also an increase in different types of information required to perform a correct input and/or maintain high performance by monitoring of information.

To test whether this hypothesised measurement system could be effective participants were required. The most commonly available sets, which would likely display different information gathering and analysis strategies, were those who play computer games and those who do not (or at least to a very limited extent). The potential for 'Gamers' to be utilised as future RPAS operators has also not gone unnoticed; the potential for a 'Gamer', to be a much more effective initial RPAS trainee (and thereby reducing training times and costs).

To analyse this information the author created a data capture program, which was used during experimentation to capture and sort participant data, and a data analysis program which performed multiple logical mathematical functions upon each users data to allow the author to perform statistical analysis much more easily. Both of the programs contained original programming in terms of formula, function and graphical user interface (GUI).

1.2 Aims and Objectives

A need has been identified within the scientific community for further research investigating optimal operator selection, more specifically targeted at the potential of using experienced video game enthusiasts as potentials for RPAS or as a remotely operated systems operator training [4]. With a statistically viable group of subject matter expert RPAS operators being unavailable the research focus upon the differences between experienced video gamers, a person who plays a large amount of computer games on a regular basis, versus non-gamers, those who don't play computer games on a regular basis, and aims to identify these differences as potential cues for RPAS operator training.

Aim1: To understand task based analysis of performance and apply it to semi-automated performance measurement systems

Aim 2: To demonstrate that decision-making ability and data processing capability is enhanced by experience of interactions with computer game based environments

Objective 1: To identify the literature available in the RPAS and task analysis domain applied to platform classification and the training of operators (A1)

Objective2: Create a potential task analysis system based around RPAS operation and based upon current Air Force doctrine based upon the MEC system (A1)

Objective 3: To define the components of a semi-automated performance measurement system in order to create a demonstrator/simulator (A1)

Objective 4: Creation of a simulator, with bespoke data acquisition software, based upon a semi-automated performance measurement system (A1)

Objective 5: To create and test software to acquire and analyse data received from the simulation (A1)

Objective 6: To identify whether information set usage can be considered a viable indicator of performance when applied to a task based system (A2)

Objective 7: To identify key factors in participant selection, experimental suitability (pre, current and post) and empirical data collection via creation and application of questionnaires

Objective 8: To identify if a more experienced gamer manage stepped increases in workload more effectively and accurately than a non-gamer with respect to flight stability and object spotting and identification (A2)

Objective 9: To identify if an experienced gamer's information set usage and data acquisition strategy differ from that of a non-gamer (A2)

Objective 10: To identify inter-group similarities with information set usage as well as similarities with degradation of performance with increasing workload (A2)

Objective 11: To identify if increased workload affect information set usage in both groups (A2)

1.3 Thesis Outline

This thesis consists of 6 chapters. Outlined below is the content of each of these chapters.

Chapter Two: Literature Review

This chapter provides an investigation into the current state of Unmanned Aerial Vehicles (UAVs), also known as Remotely Pilot Aerial Systems (RPAS) and their operators. It will also try to understand and adapt the Mission Essential Competency training ethos. Human factors associated with task based measurement and analysis will also be investigated as they pertain to later methods and conclusions

It will also outline the relevant research gaps which this research hopes to fill.

Chapter Three: Methodology

This chapter initially contains details of the initial simulator creation and the case study carried out to initially investigate the problems pertaining to objective pilot performance measurement. It will also provide a brief investigation and analysis of the case study empirical results.

This is followed by the main experiment, adapted and influenced by the case study, and will detail the experimental process as well as solutions created to negate the problems identified in within the previous chapter.

Chapter Four: Findings and Analysis

This chapter contains three distinct sections: Case Study findings, SimPACT software creation and main experimental analysis. The Case Study section will investigate, although not analyse the findings of the case study and how the findings impact on future experimentation and analysis.

The SimPACT software section will detail the creation of software designed by the author to mitigate some of the issues highlighted by the Case Study.

The main analysis (sections 4.3 & 4.4) investigate, using advanced statistical analysis techniques, links and differences between the two participating groups and identify why these exist.

Chapter Five: Discussion

This chapter examines the relevant findings derived in the previous chapter and will apply these findings in the context of the research questions

Chapter Six: Conclusion

The final chapter will highlight the main contributions of this research as well as go on to explore potential further areas of research and interesting topics

Chapter 2 Literature Review

This chapter reviews current literature regarding ‘autonomous’ systems and derives conceptual insights into the use and workings of these systems. This chapter then proceeds to review literature associated with task related training and the association with the Mission Essential Competency (MEC) and provide potential conceptual and theoretical development of these areas pertaining to their use with future experimentation. It also investigates literature and conceptual development of performance measurement associated with later experimentation.

2.1 Autonomous Systems

Autonomy (Ancient Greek: αὐτονομία autonomia from αὐτόνομοςautonomos from αὐτο- auto- "self" + νόμος nomos, "law", hence when combined understood to mean "one who gives oneself their own law"

-Oxford Dictionary

2.1.1 Remotely Piloted Aerial Systems

UAV - Unmanned Aerial Vehicle -

"A powered, aerial vehicle that does not carry a human operator, uses aerodynamic forces to provide vehicle lift, can fly autonomously or be piloted remotely, can be expendable or recoverable, and can carry a lethal or non-lethal payload. Ballistic or semi ballistic vehicles, cruise missiles, and artillery projectiles are not considered unmanned aerial vehicles" - Department of Defence Dictionary of Military and Associated Terms, JP1-02, April 2001

A Brief History

Unmanned Aerial Vehicles (UAV) or Remotely Piloted Aerial Systems (RPAS) have been in development for nearly a century; initial development of remotely piloted systems can be seen during 1914 to 1918 with the Hewitt-Sperry Automatic Airplane [5]; this aircraft worked using a rudimentary autopilot that would form the basic structure for most future autopilot systems. The purpose was military in nature and the intended outcome was to create an aerial unit that would proceed a certain distance along a set trajectory and then deliver a payload to a target area.

Of course, with the rudimentary design being inaccurate, it was never deployed as an operational weapon; in fact the First World War had nearly finished by the time the N-9 variant made a successful, unmanned launch by which point they were no longer required and the project was shelved. It would be over a decade before the United States military began to undertake serious RPAS research again. Although unsuccessful, Hewitt and Sperry's innovation paved the way for future RPAS development.

Some of the best examples of early RPAS systems are those developed between 1930 and 1950.

The RAE Larynx [6] and The German V1 rocket can be seen as the evolutionary descendant of the Hewitt- Sperry flying bomb, it used similar autopilot technology for control along a specific trajectory as well as height. The V1 and the Larynx were classed as a guided missile rather than a true RPAS system but the concept of an unsupervised autopilot system performing a task could still be thought of as being similar to certain aspects of modern remotely piloted systems in which the pilot sets a goal and then only monitors the craft rather than operating. [7]

The De-Havilland Queen Bee (circa 1935) [8] could be interpreted as the first true type of RPAS or UAV system developed; it was the first returnable and reusable unmanned aerial system and may have been the origin for the term ‘drone’ due to the system being called the "Queen Bee" . Radio controlled and utilising pneumatic servos for flight control, it could operate at over 16,000ft. Over 400 were built during its 12 year service period.

Target radio controlled drones became popular during the second world war with not only the Queen Bee produced in large numbers but also smaller Radio plane OQ-2 drones being used by the USAF for gunnery practice. These drones, the brainchild of Reginald Denny, were based around a 6hp twin piston engine inside a 12ft by 8ft monoplane design and were intended for both military and hobbyist; eventually bought by the military, thousands were produced and used.

Post war development of unmanned drones was continued by Radioplane (later to be incorporated into Northrop) in the form of a family of BTT (Basic Training Target) drones; this line of drones was still in production into the 1980's. Based around the original OQ-2 concept of a single-engined remote controlled vehicle, they incorporated autopilot systems with the remote control feature allowing for both manual and automated flight. Size, power and endurance were also increased as well as adding radar enhancement devices to provide the aircraft with a definable radar signature. The MQM- 57 Falconer, part of the Radioplane family, is an example of an early reconnaissance UAV; unlike other OQ-2 variants and descendants it was fitted with cameras and illumination flares [9].

McDonnell Aircraft, later to merge with Boeing, was the first company to build an operational decoy drone; this drone was designated the ADM-20 Quail [10] and, unlike the Radioplane BTT series, was jet powered instead of propeller allowing it to mimic speeds of the current, subsonic, jet engined aircraft of the period. The autopilot system allowed for the drone to make multiple, pre- programmed course and speed corrections while in operation. [11]

As fighter speeds increased so did the need for a target drone to be able to match these speeds. This led to the development of the Northrop AQM-35, [12] with early versions being able to reach speeds of Mach 1.55 and later versions reaching speed over Mach 2. Although the AQM- 35 reached its operational criteria it was never considered successful with only a small number of 25 units produced.

Rather than building bespoke aircraft some current manned platforms were also retrofitted for unmanned, radio controlled use during the 1950’s and 60s, the most notable were the B-17 Flying Fortresses which were remote controlled by two separate teams; one team based in a jeep at the airfield for take-off and landing and a second team for main flight based in another, nearby, B-17. This mimics many modern day RPAS operation standards where two different teams are employed for launch and recovery and main flight. The retrofitted B-17’s were deployed to closely monitor the Bikini Atoll nuclear test and collect samples from inside the radioactive cloud. B-17 drones were also used, in the same role, in further nuclear experiments along with modified Lockheed P-80 drone fighters.

With the target and decoy drones proving themselves to be effective and with the innovation of the MQM-57 Falconer reconnaissance aircraft, more development of reconnaissance drones began in earnest during the early 1960's. The Ryan Aeronautical Company Firebee, which had been in operation since the early 1950's in one form or another, was adapted to reconnaissance in two new forms, the Model 147A Fire Fly and the Model 147B Lightning Bug both of which were fitted with static and motion cameras. The reconnaissance Fire Fly and Lightning bug didn't see operational use until 1964 but they and variants of these platforms were still in active use until the 1990's and were heavily used during the Vietnam War.

During the 1970's individual companies in Israel began to develop RPAS technology in response to the Yom Kippur war (1973). It took the Israeli military until the late 1970's to show positive interest in developing and procuring RPAS’s.[13] By 1982 the Taridan Mastiff and the IAI Scout were both in operation use and were deployed very successfully as both decoys and reconnaissance against Syrian SAM sites during the 1982 Lebanon crisis. The Israeli RPAS program came to the attention of the US military during this conflict and subsequently the US military ordered both Israeli Mastiff and Scout RPASs; in producing these RPASs Taridan and IAI partnered with AAI and created the IAI Malat division. The Mastiff and Scout, later to be replaced by the Malat Searcher, saw active service until the early 1990's.

The world’s first UCAV (Unmanned Combat Aerial Vehicle) also appeared around this time; carrying 6 RPG's the Iranian RPAS was deployed against Iraqi targets; although few details are available it was likely highly inaccurate but it has paved the way for future unmanned combat vehicles. [14]

Current Systems, Licensing and Regulations

Current military autonomous systems operate in all environments (air, sea, land, space) with varying degrees of development in each medium. As this research focuses purely on training for autonomous aerial vehicles other potential autonomous systems will not be discussed. However, it would be naive to ignore the potential of this research to be used on non- aerial vehicles operating in different mediums, as the theoretical training structure will be highly adaptable to other applications.

The use of RPAS systems has not been lost on non-military industries with the RPAS commercial market expanding rapidly. Public sector services such as the fire brigade and the police force are already using micro RPAS's for fire/hot spot detection and crowd monitoring [15] and surveillance. However, there are current limitations to civil use of RPASs due to current airspace restrictions. The Civil Aviation Authority (CAA) currently has this general guidance for registering a RPAS in the UK:

"Unmanned aircraft with an operating mass in excess of 20 kg are required to be registered unless they are flying under an exemption or under the provisions of a 'B Conditions' approval issued to an organisation under BCAR A8-9. Unmanned aircraft with an operating mass of more than 150 kg must be registered with the CAA. Once the CAA has processed the application, the aircraft will be issued with a registration ID consisting of five characters starting 'G-' (e.g. G- ABCD) and the details will be entered into the aircraft register. The registration must be displayed permanently on the aircraft in accordance with Part 3 of Schedule 3 to the ANO 2009." [16]

"Unmanned aircraft with an operating mass of more than 150 kg are subject to European Regulation (EC) No. 216/2008. Accordingly, the design and manufacture of the aircraft must be in accordance with the relevant Certification Specifications similar to manned aircraft and they must be issued with a Certificate of Airworthiness or Permit to Fly. More information is available on the EASA website.

An unmanned aircraft with an operating mass of between 20 kg and 150 kg is required to qualify for a Certificate of Airworthiness under UK regulations. However, if the aircraft is to be flown within a 500 m radius and below 400 ft, or within segregated airspace, the CAA may be prepared to exempt from the requirement for a Certificate of Airworthiness if there is a level of airworthiness assurance appropriate to the UAS and the intended flights. The CAA may issue an exemption on the basis of its own investigations or by recommendation from an organisation approved under BCAR A8- 22." [17]

The CAA CAP 722 document [18] covers the issue of licensing both operator and aircraft more comprehensively. According to the CAA website [19] there is no current legal license for RPAS operators (as of 2014):

"At the present time there are no RPA pilot licenses recognised in aviation law. However, it is essential that pilots of any aircraft have at least a basic understanding of the applicable regulations, in particular the Rules of the Air Regulations. Therefore, the CAA will require a potential RPA operator to demonstrate that the pilot is appropriately qualified before any operating permission is issued.

The Basic National UAS Certificate for Small Unmanned Aircraft (BNUC-STM) is available from EuroUSC and is a UAV-specific qualification. Pilots must demonstrate the necessary skills and knowledge to pass the ground exam and flight test. The CAA recognises the BNUC-STM as evidence of pilot competence." - [20]

The regulation of both operators and systems is a growing concern globally with the need for current and future systems to be interoperable in terms of hardware, software and operator training as well as capable of adapting to existing manned aircraft and air traffic control protocols with an interaction mimicking that of a manned aircraft. One of the milestones for this is the development of standardisation of the Human Machine Interface (HMI); until a standardised human interface is created there is no possibility of standardised RPAS operator training.

The principle behind this can be seen with a comparison to the DVLA (Driving and Vehicle Licensing Agency). A car or light vehicle license (Category B) permits the holder to drive any car or light vehicle set within a standardised model; however, if the holder has only passed on a vehicle with automatic transmission this does not make them eligible to drive a vehicle with manual transmission as they have not demonstrated the necessary skill set. A manual transmission license holder will be allowed to drive a light vehicle with automatic transmission.

Applying this principle to a non-standardised RPAS platform licensing environment would mean that each operator would only be licensed to fly the individual platform with which competence was demonstrated and would have to re-qualify for other platforms with different interfaces. Obviously this would create a licensing system that would be ineffective as well as overly complicated. With a standardised interface, then, a DVLA style of licensing system could be potentially applied with categorised interfaces for platforms with a standardised operation. Due to the ethical need for a human redundant element within any current RPAS a single basic license would need to be issued for fundamental remote control of the system, this would be required no matter how complex or automated the system.

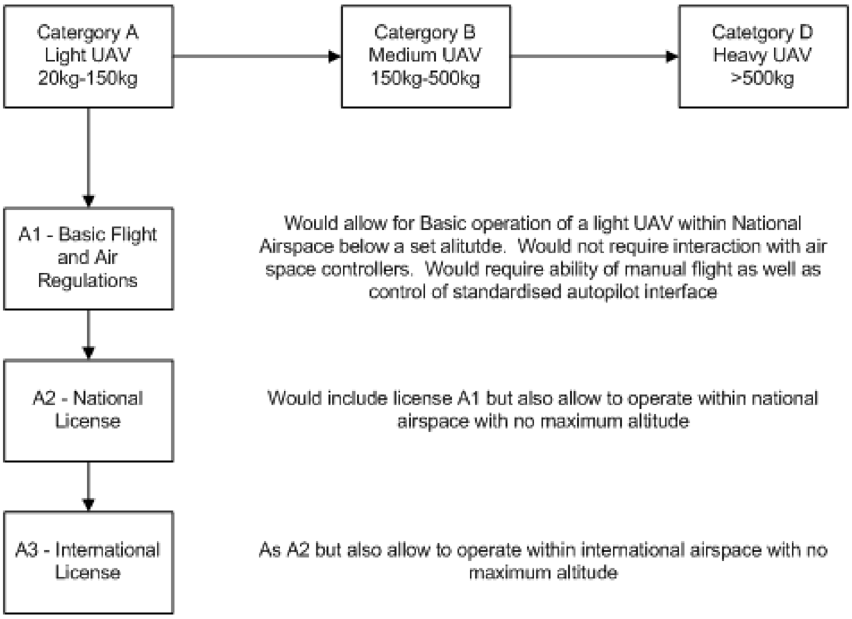


Figure 2.1- Concept of standardised training system

From that point further bolt-on types of operation could be incorporated and trained independently of the main licensing body, as long as the licensing body considers to bespoke training to be appropriate and not impact on basic RPAS operation. This would allow the systems producing company to independently train potential operators on bespoke elements of their systems design while still retaining the basic operator flight qualification.

The conceptual training system, pictured above (Figure 2.1), only shows the concept of a three- tier light RPAS license; with larger RPAS the first license would almost certainly have to include access to National Airspace with either an improved maximum altitude or no set limit for maximum altitude. It is likely that a category A (light) and category B (medium) RPAS could have differing HMI (Human Machine Interface) requirements, due to change of role between light and medium systems. This means that licensing for each category is completely independent and will likely be based on type of standardised HMI.

The need for interoperability and operator training is not just a civil issue but is also reflected in the changing military needs; NATO, in its STANAG (Standardization Agreement) documents [21] has stated the need for interoperability between RPAS platforms and communications as well as relevant operator training guidelines.

"The aim of this agreement is to promote interoperability of present and future UAV systems in a NATO Combined/Joint Service Environment. Interoperability is required because it will significantly enhance the war fighting capability of the forces. Interoperability will increase flexibility and efficiency to meet mission objectives through sharing of assets and common utilization of information generated from UAV systems. “[21: p. vi]

"The aim of this agreement is to establish a broad set of training guidelines and the skills required of a Designated UAV Operator (DUO) to operate a UAV in all classes of airspace." [22: p.vi]

Hopefully these documents will help to form a basis for both military and civil RPAS interoperability and, in so doing, create the right environment for a standardised HMI and licensing system to be created. Without a standardised licensing system the commercial aspect of the RPAS remains underdeveloped.

Current RPAS in the UK, vary widely in terms of use, size, crew and level of autonomy (which is discussed later). At this moment in time RPAS systems within the military still have limited roles, these roles being reconnaissance, surveillance and air- to-ground type of operations; there is scope, however, for RPAS to fill other roles such as supply and defence.

The main role of RPAS systems, at this time, appears to be for reconnaissance and surveillance with vehicles such as Globalhawk being the latest in the evolutionary line. The scale of this type of operation is limited to the design of the system with some being single operator, human launch vehicles, such as Raven, which up the scale to Predator and Globalhawk levels. The smaller systems are designed to be used as support for individual land units and to be deployed whenever that unit sees fit; this makes the smaller systems local-area specific as they are only short range.

At the other end of the scale Globalhawk is designed as a long endurance, large area reconnaissance or HALO (High Altitude Low Observability) vehicle; this type of vehicle can cover a very large area over an extended period of time and, with a team of analysts, can produce detailed surveillance/reconnaissance information using its battery of sensors and cameras. This makes Globalhawk much better at providing an overview of the 'big picture' but less useful when specific local area information is required quickly.

RPAS systems have also recently been used in a weaponised air-to-ground capacity (UCAV) in the form of MQ-9 Reaper [23]; this vehicle is loosely based on Predator but can carry air-to-ground weapons which are deployable with the operator’s discretion. This capability allows for offensive use within a hostile environment that may be considered too dangerous for the deployment of strike aircraft; it also allows for troop support as the Reaper can loiter in an area until required.

The differences between a manned and unmanned system are completely dependent on the level of autonomy afforded to the unmanned system. As an example the Predator is piloted using stick and thrust controls, similar to an aircraft and can be seen to be, in terms of its operation, a simulation of a manned aircraft. At the other end of the scale, a system like Globalhawk does not require manual operator input in the same way as a manned aircraft; the control is input through keyboard and mouse.

It can be seen that a list of differences can only be compiled when comparing a specific system or level of autonomy to that of the manned system and, with the current climate of secrecy surrounding RPAS systems, it is very challenging to fully compare a specific RPAS system to manned flight.

Fortunately, the comparison is not limited to a specific RPAS system but can be based on a generalised version of the BAE System’s PACT levels (refer to Table 2.3); though not as complete or accurate as a specific comparison it none the less gives a feel for the increasing differences with increasing levels of autonomy.

Using the aforementioned PACT levels and some generalised information regarding RPAS systems the first observation of fundamental differences can be made between the most basic level of autonomy, level 0 - Remote operation.

Looking at the first level (level 0):

At this level there are few differences between manned and unmanned flight, but the differences that occur are fundamentally important; the first being the remote operation itself.

The pilot is no longer located in the system but is now based at a fixed location from which the RPAS system is operated; this leads to loss of physically intuitive data that may be gathered from the system (human interpretation of G-Forces, visual stimuli, auditory stimuli, vibration etc.). This loss of intuitive information is difficult to replace and to replicate at the ground station.

The second key difference is the way in which the pilot must control the system as well as interpret information received from the system. The pilot's view has now been limited from a full cockpit view to that of the cameras and sensors available on the system; this information is then displayed on multiple monitors which is of limited size. This limits the amount of information that a pilot can realistically observe.

An RPAS system has a much reduced FoV (Field of View) than a counterpart manned system; this means that not only is there less information available but also that the information that is available will be displayed in a much more concise, and possibly confusing, way. This could lead to a higher level of incorrect interpretation and inaccuracy in completing the task.

The control interface also differs; the system takes over several of the pilot-monitored/controlled functions in the cockpit, freeing the pilot to concentrate on other areas.

This can be seen as transference of elements of manned piloting being transferred, and built into the RPAS systems architecture; as the autonomy increases many more of these elements will be incorporated into the RPAS system architecture thereby freeing the operator, and his data analysis, to concentrate on higher level operation and interpretation. As the levels increase after this, control systems are the first things likely to change, with a movement towards more autonomy and away from conventional stick and throttle control systems. It is also possible that there will be more of a loss of first hand situational awareness (which could be replaced with system awareness); it will also be more likely that, due to the lowered involvement of the operator, that long term concentration levels could possibly drop.

Taking into account the new thinking regarding dynamic autonomy levels it is likely there will be combinations of conventional and autonomous control as well as differing levels of situational awareness and concentration required at different points during the mission; however, in general, as the systems become more intelligent there will be a move away from conventional controls and a lowering of SA which will lead to a drop in concentration.

2.1.2 Levels of Autonomy

Level of Autonomy (LoA) is a complex and still a relatively uniformly undefined field but it will have a huge bearing on the future use and development of RPAS. Relating to LoA to section 2.1.1 it can be seen that the principle of semi-autonomy can be applied but only to a limited extent as most platform systems are still merely automated rather than semi-autonomous.

The concept behind level of autonomy is that an RPAS has a specific level of decision making power as well as operator control, this level could vary for that given system dependent on several factors:

- Operator competence
- Role type
- Role importance
- Mission type
- Mission Complexity
- Mission importance
- Rules of Engagement (RoA)
- Ethical impact

Sheridan, one of the founders of LoA thinking first writes that automation should be decided by considering human tasks process: Acquire, Analyse, Decide, Act [24]. This is subsequent to decades of investigation and research with one of the earliest LoA documents being that of Sheridan and Verplank's 1978 Human and Computer Control of Teleoperators [25]. Although this is not specifically LoA directed it is a comprehensive study of undersea UXV human/machine operation and includes task analyses of the human/machine interaction. The Acquire, Analyse, Decide, Act (AADA) process is supported by USAF Colonel John Boyd’s OODA loop and complements this thesis authors own early and independent analysis of human tasks processes which exactly matched that of Sheridan (see figure 2.2). [26]

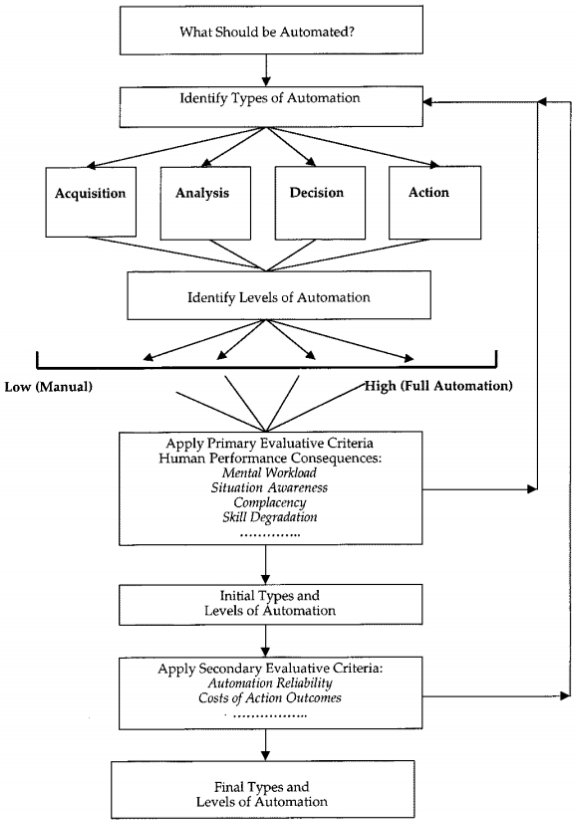


Figure 2.2 - Sheriden *Acquire, Analyse, Decide, Act (AADA)* [24: p.290]

The use of the four step AADA approach leads to individually partitioning the capabilities of an automated system; where one partition may have a high level of automation another may be low. An example of this could be a system that has a high ability to acquire information but would rely heavily on a human aspect to interpret the information as part of the analysis partition; this can be equated to a low level RPAS (such as BAE Systems HERTI) in which the operator must interpret all of the information feeds and then base their decision on those feeds. There are however many more aspects to the LoA of an unmanned vehicle than just these simplified partitions, a system may indeed have highly complex analytical or decision making abilities but these may only be present for a limited amount of situations and, with this being the case, Sheridan et al [24] conclude that the above framework for identification of LoA is therefore purely a simplified concept that may aid in future development of more complex and complete LoA identification tool sets.

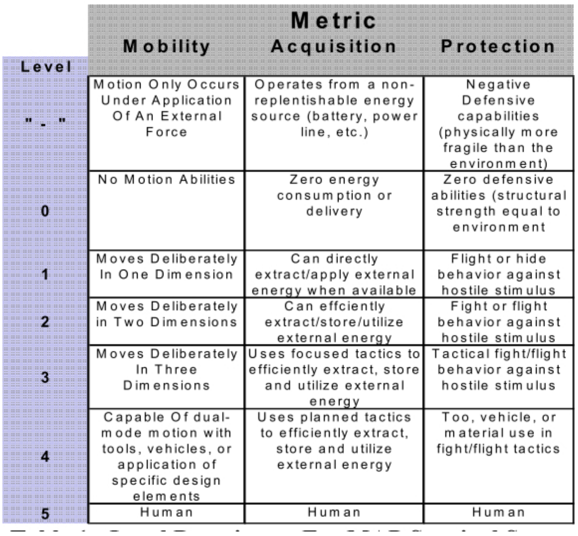
Since the publication of this work there have been many studies geared towards obtaining a generic standard for classification of RPAS LoA, many of which cite Sheridan’s work as a basis. Although this work is only complementary to many of the other projects undertaken it seems that it has formed a crucial direction of thinking within the community.

Currently there are multiple taxonomies of RPAS's available, some of which are built on older taxonomies. Clough et al Metrics Schemetrics [27] offers a reasonably concise tool set for identifying LoA known as Autonomous Control Level (ACL), the quotation below shows that, up until comparatively recently, there has been no commonly agreed upon or effective LoA identification tool:

"The Fixed-Wing Vehicle Initiative (FWV) has broad goals across numerous vehicle technologies. One of those areas is mission management of RPASs. Our broad goal is to develop the technology allowing RPASs to replace human piloted aircraft for any conceivable mission. This implies that we have to give RPASs some level of autonomy to accomplish the missions. One of the cornerstones of the FWV process is the establishment of metrics so one know that a goal is reached, but what metrics were available for measuring UAV autonomy? Our research, in conjunction with industry, determined that there was not any sort of metric as we desired. Thus we set out to define our own [Note 1]." [27: p.1]

The AFRL themselves used principles developed by Tilden & Hasslacher [28] and the Charles Stark Draper Laboratory [29]. In the course of his work simplistic robots Tilden required a method by which to measure their level of autonomy, he enlisted the aid of Physicist Hasslacher and they produced the Mobility, Acquisition and Protection (MAP) metric which graded each of the previous aspects of the robots function and incorporated them into a definable chart, Table 2.1.

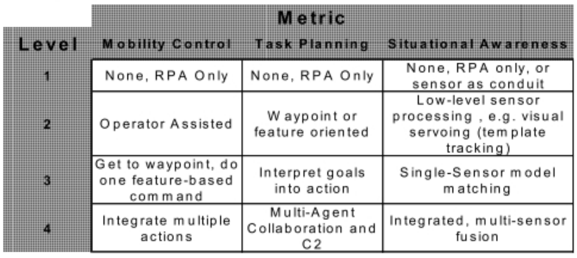
Table 2.1 - Tildn/Hasslacher Mobility, Acquisition and Protection, from Metrics Schemetrics! - Clough [27: p.2]



Level	Metric Mobility	 Acquisition	 Protection
" - "	Motion Only Occurs Under Application Of An External Force	Operates from a non-replentishable energy source (battery, power line, etc.)	Negative Defensive capabilities (physically more fragile than the environment)
0	No Motion Abilities	Zero energy consumption or delivery	Zero defensive abilities (structural strength equal to environment
1	Moves Deliberately In One Dimension	Can directly extract/apply external energy when available	Flight or hide behavior against hostile stimulus
2	Moves Deliberately in Two Dimensions	Can effciently extract/store/utilize external energy	Fight or flight behavior against hostile stimulus
3	Moves Deliberately In Three Dimensions	Uses focused tactics to efficiently extract, store and utilize external energy	Tactical fight/flight behavior against hostile stimulus
4	Capable Of dual-mode motion with tools, vehicles, or application of specific design elements	Uses planned tactics to efficiently extract, store and utilize external energy	Too, vehicle, or material use in fight/flight tactics
5	Human	Human	Human

The Draper Laboratory found several different ways in which to measure a systems autonomy but Clough focused on the 3D Intelligence Space (defined in table 2.2).

Table 2.2 - Draper Laboratory 3D Intelligent Space [27: p.3]

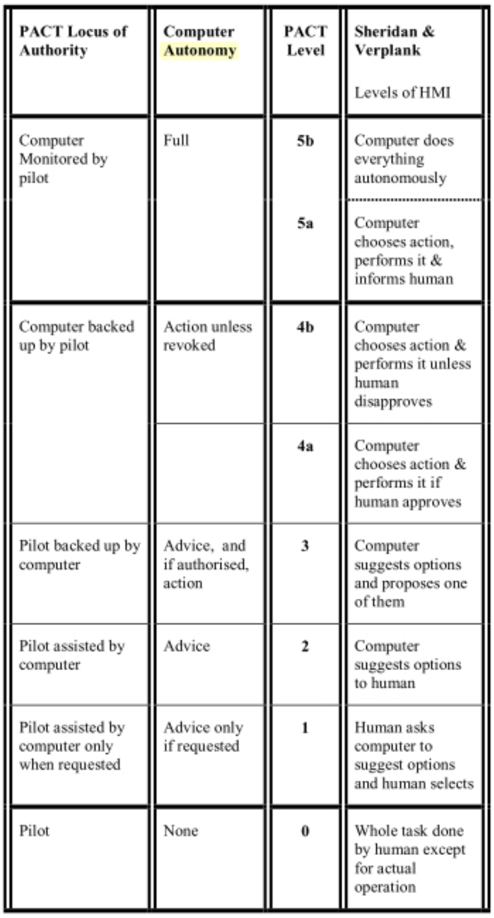


Level	Metric Mobility Control	 Task Planning	 Situational Awareness
1	None, RPA Only	None, RPA Only	None, RPA only, or sensor as conduit
2	Operator Assisted	Waypoint or feature oriented	Low-level sensor processing , e.g. visual servoing (template tracking)
3	Get to waypoint, do one feature-based command	Interpret goals into action	Single-Sensor model matching
4	Integrate multiple actions	Multi-Agent Collaboration and C2	Integrated, multi-sensor fusion

By integrating the above two approaches and revising and iterating the AFRL developed its ACL metrics. Since its inception this system has proven successful for the AFRL in planning and operation of unmanned vehicles but it is considered by some to be limited in its function due to its lack of consideration of unmanned vehicles at the functional level [30].

BAE Systems own approach to the problem was to use the UK Ministry of Defence (MoD) COGPIT programme [31], documentation regarding COGPIT is unfortunately unavailable and cannot currently be commented on. The produced LoA metric were based on the MoD developed Pilot Authority and Control of Tasks or PACT and comprises a more simplified view of LoA but also allows for a varying degree of autonomy within the system (table 2.3); Sheridan's early work has been cited as an aid to this project. So far this seems to be one of the more comprehensive solutions to the defining of LoA.

Table 2.3 - BAE Systems modified PACT levels for autonomous systems [30]



PACT Locus of Authority	Computer Autonomy	PACT Level	Sheridan & Verplank Levels of HMI
Computer Monitored by pilot	Full	5b	Computer does everything autonomously
		5a	Computer chooses action, performs it & informs human
Computer backed up by pilot	Action unless revoked	4b	Computer chooses action & performs it unless human disapproves
		4a	Computer chooses action & performs it if human approves
Pilot backed up by computer	Advice, and if authorised, action	3	Computer suggests options and proposes one of them
Pilot assisted by computer	Advice	2	Computer suggests options to human
Pilot assisted by computer only when requested	Advice only if requested	1	Human asks computer to suggest options and human selects
Pilot	None	0	Whole task done by human except for actual operation

Generically across the multiple types of LoA metric it can be seen that a decreasing amount of operator input (in terms of the AADA approach) would lead to an approximately inversely proportional amount of system autonomy; this, of course has major implications for operator training within the RPAS environment. With classic aeronautical training there is always the same base need for manual control as well as operation of basic automated systems but within the varying LoA world of the RPAS training an operator at one LoA may preclude him from operating a system at a different LoA

This creates possibilities for varying ways in which RPASs are controlled, either through their own systems or by an operator; this mainly is due to the wide variety of companies and systems, both hardware and software available at this present time. An automated function on one platform may not be present within the framework of a differing companies system. In a way this returns again to Sheridan's work by having a differing level autonomy within the AADA as the system may require more or less interaction at every partition. Until standardisation is reached it will be difficult to create any form of licensing system or interoperability.

2.1.3 Pilots and the Ground Control Station

“There is an art . . . to flying. The knack lies in learning how to throw yourself at the ground and miss.”

— Douglas Adams, The Hitchhikers Guide to the Galaxy

The Ground Control Station

The Ground Control Station, or GCS, can be defined as the main Human Machine Interface of any part of an entire RPAS platform; the GCS is the point at which any data received or relayed to the RPAS is processed and managed. This includes visual data, control data, situational awareness data as well as a host of bespoke (dependent on producing company) features which will be tailored specifically to that vehicle.

With the advent of remotely operated vehicles there has always had to have been an operator to provide control for the system. There is still a blurred line for the distinction between a Radio Controlled plane pilot and an RPAS pilot, with the systems PACT level of 0, as theoretically the only difference is the location of the operator. The Radio Controlled plane heritage can still be seen in several military and commercial RPAS with the main control unit being that of a modified R/C controller interface or a controller reasonably similar to a games console controller (see fig 2.6).

HMI's for micro RPAS can also be purely based on automated take-off and landing and, therefore, go without a manual element leaving any higher level inputs to a standard mouse and keyboard type interface. This same theme progresses to the larger RPAS where, instead of a ‘pad’ for manual control a more appropriate aeronautical themed stick and throttle are often used, larger RPAS are, however, more likely to have higher levels of automation and, therefore, the manual aspect could potentially less important [32]

Figure 2.3 - R/C RPAS interfaces. Top: IStart UAV, Bottom: Kutta Technologies handheld UAV Control device

GCS design and usage varies wildly from company to company and platform to platform (figure 2.3); this is not just seen within the hardware architecture but also within the software as many, if not all, GCS control software being designed by the producing company. The disparity between systems can be seen with a basic online search.

Although this diversity helps to promote innovation it also constrains interoperability between companies, both production and consumer, as well as nations; this also leads to the problem of requiring specific training for each and every type of RPAS system currently in use within an organisation which leads to excess costs in training and from an enlarged personnel base.

For the time being there is no optimum GCS design or data protocol due to the very individual requirements of each RPAS platform. This, however, is undesirable due to inadequate interoperability and has been recognised by NATO as an issue. NATO has issued several documents within the STANAG protocols [21 & 22] detailing changes to RPAS systems to enable ease of data exchange as well as control; it is difficult to judge whether these documents have had much effect as much of the further progress will be classified under company or national laws.

Pilots

As RPAS were originally developed and are predominantly maintained by military institutions it is not surprising that the predominant number of larger RPAS operators are of military pilot origin with prerequisite experience on manned platforms. The military RPAS operators were, up until recently, drawn only from this experienced manned pilot pool; this is mainly due to the military not having a specific training pipeline from training initiation to final course completion.

Military RPAS operators are required to have a fundamental knowledge of basic aircraft flight as well as communications standards, this can apply to any military pilot and it has been observed [30] that military RPAS operators have varying flight backgrounds, from fast jet flight to transport duties. This indicates that only a base knowledge of flight is required to begin RPAS training and no specialist skills associated with a particular manned platform are required, although there may be internal preferences.

A case study of an available RPAS operator (in this case Reaper) was carried out by the researcher (see appendix D) supports the fact that military transport pilots are being used as RPAS pilots; the subject showed over 1000 hours having been flown on a C-130 Hercules which accounts for over 30% of his manned military flight time. Since transferring to an RPAS he has flown 1250 hours, which is a greater number of hours than that of his transport flight time; with an age of 34 it appears that he was transferred nearly half way through his military career.

As operators seem to require only basic flight training and standard communications knowledge it is entirely feasible that a recruit could begin training with other manned platform recruits and then specialise to an RPAS platform; this would follow many military training doctrines in which most recruits start with generic training and then specialise later on during the training timeline.

However, the RPAS physical requirements differ from that of normal pilot selection as the potential RPAS operator will not be subjected to the same physical and environmental experiences of a manned aircraft; this could open the door to recruits who would not otherwise be selected for flight training (due to disabilities, physical characteristics etc.). The inclusion of these potentials would further increase the candidate pool for RPAS training.

As already mentioned RPAS operators are still predominantly experienced manned aircraft operators that have already seen some active service in one form or another; there may be a myriad of reasons why manned aircraft operators transfer to an unmanned system, from social, physical to financial. There still seems to be an overriding ego associated with many pilots who are seen in the public view that flying an unmanned system piloted remotely [33] as not being a true pilot, Lee states this as such in his summary.

> "It is difficult to see how representations of RPAS operations and crews in the media will shift from the negative connotations now commonly portrayed to something more positive. The contrast with the long established and deeply embedded public perception of fighter pilot and fast-jet operations in particular provides TV and print media journalists with easy and convenient labels on which to hang their stories. Consequently, those who opt to serve as remote aircrew will have to accept that they will never be viewed in the romantic or daring light of aircrew elsewhere" [33: p.16]

This may inform that the RPAS training route may be more difficult to recruit for from a pool of candidates whose goals are centred on the desire to fly manned aircraft and be part of the ‘knights of the sky ethos’ [33]. Existing pilots may also feel the same, if not more strongly, regarding the ethos differences between manned and unmanned flight and this may make it even harder to recruit seasoned manned pilots into an unmanned role. However, RPAS pilots strongly still refer to themselves as true pilots; this has been seen in both literature as well as in the researcher’s personal experience.

Another potential factor of the use of manned pilots in an unmanned role is boredom and lack of sense of achievement. It can be inferred that a manned pilot will expect a certain level of stimulation from the platform he is currently flying, whether that is visual, physical, audible, cognitive or, in some cases, ethical [33]; the RPAS does provide this level of stimulation with only cognitive, visual and occasionally ethical stimulation being potentially available. This loss of stimulus, as well as the increase in platform automation, could potentially lead to pilot boredom on an RPAS; this ‘under load’ of pilot could lead to a drop in pilot performance and loss of situational awareness [34].

But how would this effect an RPAS pilot with no previous flight experience, and therefore, no expectation of system stimulation? It can be theorised that the lack of expected system stimulation would lead to better performance as well as situational awareness with an RPAS pilot with no expectation of stimuli being present. This, however, is a question that would require further research and is not within the remit of this thesis; this subject will be revisited in the Further Research chapter.

If it would be more desirable to, in terms of potential candidate pool, ethos, cost and comparative performance to train non-professional pilots as RPAS pilots then what should be the criteria by which potential candidates should be measured? This leads to the objective of this research, to identify whether an experienced video gamer would prove more effective at an RPAS based task than a non- gamer; the omission of commercial or private pilots from this study is both resources based as well as trying to mitigate previous potential flight experience as a factor within operator performance. The research is aimed at training an RPAS operator from scratch rather than adapting an already qualified operator, thus supporting a complete pilot training RPAS specific pipeline. Research has already indicated that an experienced gamer will perform better than a case group and will start to approach a qualified pilot in terms of flight control performance as well as cognitive tasks [35]; in figure 2.4, there is a brief graphical representation of McKinley's results from motion inference in which the video-gamer group proved to be more effective than the control group and also the pilot group at controlling divergence from a glide path

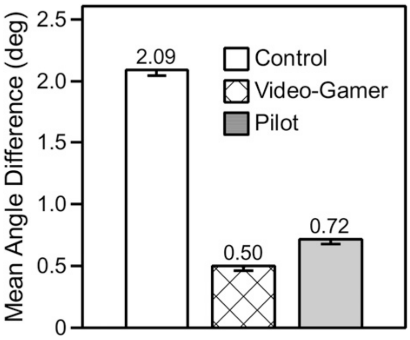


Figure 2.4 - **Motion Inference.** Mean absolute value of the angle difference. Error bars are standard error of the mean. [35: p.635-642]

These results indicate that it is justifiable to consider an experienced gamer more suited to potential RPAS candidacy than an inexperienced gamer, this is indicated by a better score for mean angle difference then existing pilots and far superior to the control group (figure 2.4).

Another element to be considered is the commercial sector’s unmanned pilots; often these types of platforms are much smaller than military designated RPAS; this could be associated within current global civil aerospace legislation regarding drone size and crash impact energy. It is more efficient for small commercial companies to produce small RPAS that conform to the relevant governing bodies’ legislation as this allows more freedom with research, development and especially testing.

The pilots of these systems seem to have wide ranging background with an often recurring theme of remote control (R/C) aircraft experience [36]. In fact, it has been found by the researcher with industry communication, that BAE systems small platform Herti predominantly trained already experienced R/C aircraft enthusiasts. This could well be due to the type of interface developed for the Herti system; it utilises a main single display and is controlled by an adapted R/C controller. In this case an R/C operator may adapt more quickly to this system due to the control setup. Theoretically it is likely that the design of the GCS interface will have a major bearing upon the adaptation rates of potential candidates from varying backgrounds to basic flight control.

Many smaller systems utilise a reduced, non-aircraft like, control interface and this could mean that a manned pilot’s performance upon these systems would in fact be reduced compared to that of other operators who have experience of this type of control interface. These smaller systems are likely to be more widely used within the commercial sector initially after structured licensing begins.

Licensing so far has not got as far as a regimented training program or even in terms of categorising and licensing of the systems themselves, let alone the pilots; the authors recent conversations with the Civil Aviation Authority (CAA) and other organisations reveal that a generic licensing system is still some way off completion as much of the current work is directed at the ethical and legal aspects of RPAS operation within civilian airspace.

2.2 Task Based Measurement

2.2.1 The Mission Essential Competency

The MEC or Mission Essential Competency is a relatively new approach to pilot training [37]. Developed in America it appears now to be a staple of the US Air Force training regime and its popularity as a training system performance measurement method is increasing, not only as a pilot training program, but for other branches of the military [38]. The MEC system is trademarked to the USAF and the AFRL.

As this research is targeted solely at pilot performance measurement, the MEC by system itself cannot be used. However, some of the structural concepts can be used to help identify a task-based pilot performance system which would aid in isolation of particular tasks relating to forthcoming experimentation. What follows is an inferred analysis (due to lack of detailed information available at the time of research) and an interpretation of a revised structure related to direct performance measurement. Initial research was, in fact, at error in the understanding of the true concept of a MEC, again due to lack of literature and verbal information and clarification. It was initially thought that the MEC system itself was used for pilot performance measurement rather than for training system gap and system analysis. This led to some of the following conceptual work being revised but still with the aim to use the concept of a task analysis approach with an aim to understand how the task can be analysed and incorporated into a performance measurement system.

The MEC is based on a tiered task analysis structure, which begins at the fundamental level of aircraft control and pilot performance and aggregates up to a more general level of competency [38]. This allows for a more focused approach to pilot training and has been shown to be more effective, both in time and performance, than conventional training.

A MEC is not, in itself, a full training programme but is used as a way in which to identify training gaps within current training, to this end it is related to evidence based training techniques; a MEC is then comprised of Supporting Competencies (SC) which are generalised areas in which the pilot must become proficient to complete the designated mission. At the root of the MEC and SC's are the Knowledge, Skills and Experiences (KSEs) which form the core of the pilots training and efficiency whilst completing the over- riding mission.

The basis for the MEC was originally derived as a job analysis method as far back as 1880 (Taylorism); Taylorism was developed by Frederick Taylor with the purpose of optimizing the work-rate of production-line workers by using time-management and trial-and-error methods to find the ‘one best way’ to complete a task.

This form of job analysis was a foundation for further research into a competency-based method.

The ethos behind the Mission Essential Competency is that every task has a start point and an end point with a variety of sub-tasks, knowledge, experiences and skills that lead to successful completion of the task. This type of method involved obtaining the views and opinions of the jobholder so that the way in which the job is carried out can be improved and streamlined.

This type of job analysis also leads to the identification of the best candidate to perform the job, in both a physical and psychological sense. The principle of this task analysis method led to many different applications in both the commercial world as well as the military. The MEC is just one of these applications.

The MEC seems to have a variety of different methods of analysis and creation ranging from workshops in which subject matter experts are used intensively to develop a MEC, to a survey based method in which multiple SMEs are asked to complete a survey and their input is then outputted into an MEC.

The MEC itself is an inferred tiered structure comprised of an inferred 3 levels:

The highest level is the MEC itself, which can be described as a phase of the mission that a pilot must complete. There can be several MECs contained in one mission which can be exclusive to one section of the mission but can also overlap with other MECs.

The mid-level is comprised of the supporting competencies, which not surprisingly, are designed to support in the completion of the higher level MEC. They are a generalised form of task analysis and contain areas of airmanship in which the pilot must be proficient.

The lowest level contains the Knowledge, Skills and Experiences that the pilot requires to complete various tasks during the mission. The KSEs are the key stone of the MEC. During the rest of this the KSE's shall be referred to as fundamental competencies. The following interpretations are inferred from existing works. [37]

Team Competencies

A MEC isn't purely about the pilot but also about the inter- and intra- team interactions depending on the type of mission being performed. For example, in a twin-seat aircraft the MECs are shared between the pilot and co- pilot; in a squadron situation the team competencies relate to how each crew relates and works with the other members of the flight.

Several of the Knowledge, Skills and Experiences mentioned in the forthcoming chapters relate to team competencies but are not team specific (i.e. could be carried out by either a single pilot or a team depending on the circumstance).

Knowledge

Along with Experiences, Knowledge and Skills are the fundamental parts of the MEC and its supporting competencies.

Knowledge can be defined as ‘the knowledge a pilot requires to complete a task or analyse a situation successfully’; as can be seen from the list this can range from knowledge of capabilities of enemy/allies to Comms standards. This type of fundamental competency is a learned competency and is affected by a pilot’s ability to learn, remember and apply.

Skills

Skills focus more around the more practical side of the fundamental competency; they are more likely to be learned through practical experience while using the system with a limited amount of knowledge acquired while 'in the classroom'.

Becoming competent at a skill-based task requires practise with the system which enhances not just functional knowledge of the system but physiological co-ordination which speeds completion of the task as well as the standards of accuracy.

It is in the this area of the fundamental competency that a pilot’s inferred skills and talents can truly be seen; things that can't be learned easily, such as hand/eye co- ordination, speed of analysis and interpretation, multi- tasking ability and adaptability, are key to the rapid and effective development of an operator and should be taken into account during the operator selection process.

Experiences

An experience can be seen as a pairing of knowledge and skills and can therefore be inferred as a slightly higher order than a fundamental competency; however, it is still not in the same order as a Supporting Competency, an Experience is based on a specific event or employment rather than a generalised fundamental competency set.

Experiences are specific tasks that an operator may encounter or have to complete while on a mission and in that sense they are very similar to knowledge and skills. Experiences are much more focused on external stimuli and circumstances and, although some experiences will be expected, even designated, to occur during a mission, other experiences will be based on random and uncontrollable events.

They also fit more with the phased mission structure of the MEC; some experiences may arise at random intervals while others could potentially continue throughout the mission, even with this, though, many experiences are likely to fall only in set MECs (or phases).

Supporting Competencies (SC)

By inference of the structure of the MEC it can be seen that a supporting competency is a lower order competency (when compared to a MEC) and entails a generalised form of pilot requirements, this can be seen in Revised MEC example task lists (Appendix C), where the SC list contains much more human factor oriented requirements while the MEC list is much more task/mission specific.

Unlike the MEC the supporting competency is not mission phase related and can appear at any time during the course of that mission; however, a supporting competency may have an enlarged or reduced role depending on the phase of the mission.

SCs such as Weapons Engagement Zone Management or Flight Battle Management may only appear during one Mission Essential Competency and, therefore, only during one phase of a mission.

Conversely SCs such as Communications and Timeline are likely to be involved in every phase of the mission and would appear in all MECs pertaining to that mission.

The supporting competency, as already, mentioned, is a lower order competency but it still sits above the fundamental competency (knowledge, skills and experiences) which forms the basis of the MEC.

The inferred structure of a MEC could, theoretically, be three tiered (see figure 2.5); this would allow for a simple, linear structure and hierarchy.

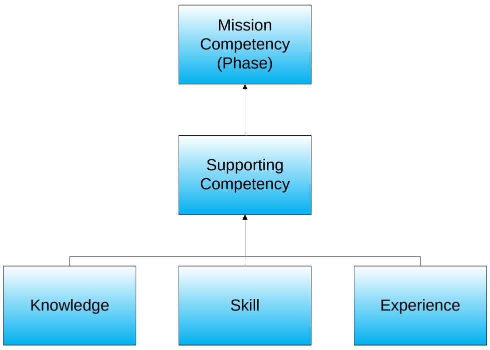


Figure 2.5 - Inferred Mission Essential Competency Structure

This, however, is not the case as can be seen when categorising the KSE into a supporting competency (Appendix C); the reason for this categorisation is for ease of performance measurement which is looked at in-depth in the next chapter. Below (figure 2.6) is a graphical representation of the non-linear relationship between KSE's, SC's and MEC's, it can be inferred from this that performance measurement with a MEC would be extremely challenging.

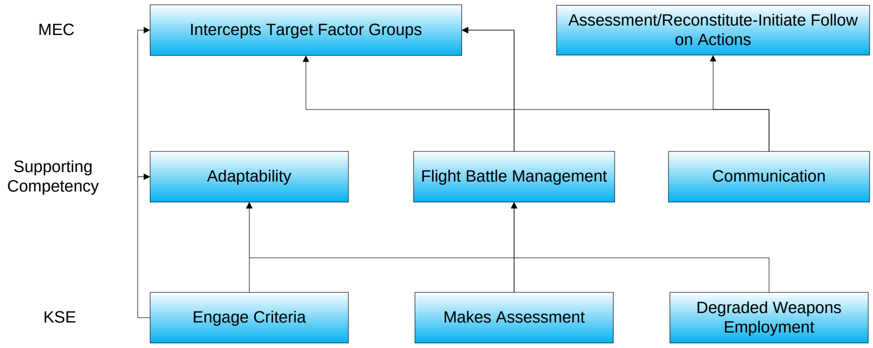


Figure 2.6 - Showing example MEC's, SC's and KSE's and their interactions

This problem occurs due to the nature of the different elements of the MEC; the MEC itself is a phase of a mission and is, therefore, time or task orientated; the Supporting Competency is a skill/knowledge set attributed to the pilot and is therefore human factors orientated; at the base level, the Knowledge, Skills and Experiences have a mixture of task, human factor and time orientation.

Below (figure 2.7) is a time based interpretation of the way in which tasks fit into the mission and MEC structure; some of the tasks (labelled alphabetically) may well be identical tasks, for example, task b may be the same as task e. This still does not identify the potential task overlap.

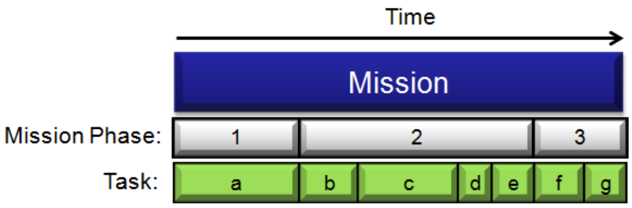


Figure 2.7 - Mission to MEC (Mission Phase) to Task Timeline

2.2.2 Evidence Based Training & Task Analysis

Evidence based training (EBT) is defined by Axiom Professional Health Learning as:

> "The definition of EBT is simply the conscientious and explicit use of the best current evidence in the design of training programs. The evidence is the outcomes from well-designed research studies in the fields of cognitive and educational psychology and instructional design. EBT isn't one specific instructional approach or one specific technique; rather it's a mind-set of never-ending curiosity over what variables lead to better learning outcomes." [39]

This definition is further supported by Richford in an non-peer reviewed online publication

> "Evidence-based practice refers to the use of research and scientific studies as a base for determining the best practices in a field." [40]

Many of the articles directed at evidence based training centre around the development of better practises within the health care industry but the concept of EBT is slowly spreading into other commercial and military sectors.

Using the concept of evidence based training as concept it is possible to apply this to pilot performance measurement and training [41].

Theoretical discussion from the investigation into evidence based training are contained within chapter 5.2.2 along with theoretical discussion for MEC development and adaptation.

2.2.3 Workload Measurement

Pilot workload, in this case of a mental tasking nature, is another subject area that has received much research attention and can either take an empirical or subjective form; an example of empirical form can be seen in Wilson’s paper [42]. Wilson uses multiple physiological methods to potentially measure increased pilot workload including heart rate, eye blinks, and brain activity to understand the pilot’s workload in a physical and empirical form. [42]

Lee, Yung-Hui and Liu, Bor-Shong's paper entitled "In-flight Workload Assessment: Comparison of Subjective and Physiological Measurements" [43] both subjective and physiological workload performance measures were considered. They concluded that:

> "The results of this study indicated that the cardiac variable (HR, HRV) and NASA TLX scale indices are sensitive enough to characterize workload for the different phases of flight. This suggests that the NASA TLX scale would be a practical tool to apply in operational environments, while the weighting and the magnitude of the ratings of the individual scales provide important diagnostic information about the specific source of loading within the task." [43: p.1083]

This, along with other similar studies supports the use of subjective workload measurement as a comparative tool.

During the course of later experimentation it will be necessary to measure the experimental participants’ interpretation of their own workload so that it can be compared with the empirical data; The NASA TLX: Task Load Index has been chosen to aid with this subjective performance measurement.

> "The NASA Task Load Index (NASA-TLX) is a subjective, multidimensional workload assessment tool. It was developed by the Human Performance Group at NASA Ames Research Centre over a three year development cycle that included more than 40 laboratory simulations. It is thought to be one of the most validated workload measurement tools in Human Factors Psychology and Engineering.”

> “The NASA-TLX rates perceived workload on six different scales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. According to Hart and Staveland (1988) [44], a participant should first rate the six scales according to how much they contributed to the workload required for the task being studied. After the ratings, each of the six scales is weighted. The weightings are achieved by answering 15 pair-wise comparisons and are designed to greatly enhance the sensitivity of the overall workload score while reducing between-rate variability.”

> “The pair-wise comparisons component is only needed after each distinctly different task. When completing similar tasks, it was found that the comparisons did not significantly add to the overall sensitivity of the tool. Some schools of thought go as far as to suggest that the pair-wise comparisons are never needed. This tool lets you select the method that best suits your preferences." [45]

The NASA TLX system provides an easy to use self-assessment tool which will aid in the understanding of the participants work load when compared to the other data received; as work load is an important factor of the later experiment, in term of overall performance, it is necessary to use an unbiased and tested tool set to aid data analysis.

The Subjective Workload Assessment Technique (SWAT) [46] was also considered to fill this role,

> "The Subjective Workload Assessment Technique (SWAT) has been developed in response to a need for a workload measure with known metric properties that is useful in operational or "realworld" environments.”

> &

> “This approach allows responses to be made in the operational setting using only three simple descriptors for each of three factors that have been used to operationally define workload. This approach also minimizes the amount of time required to make responses by keeping down the number and complexity of descriptors that an operator must memorize.” [46]

The NASA TLX tool was chosen based upon statistical findings [47]. An example NASA TLX sheet can be found in Appendix E. and the process of using this sheet is detailed more in the methodology section.

2.2.4 Information Set Usage Measurement

Eye tracking has been extensively used within the aviation community to better understand the way in which pilots interact with their systems. In fact there is a wealth of documentation supporting the use of scanning paths, object awareness etc. regarding pilot instrument usage as well as other cognitive based performance tasks.

Notably Andrew T. Duchowski summarises his 2002 report [48] that eye-tracking has become an integral part of analysing human interactions across multiple sectors involving visual analytics as well as being used as an interaction tool.

> "As the present review demonstrates, eye trackers have traditionally shown themselves to be valuable in diagnostic studies of reading and other information-processing tasks. The diagnostic use of an eye tracker, as exemplified by the research reviewed here, can be considered the eye tracker's mainstay application at the present time and probably in its near future." [48: p.13]

Another report by Hasse, Grasshof and Bruder [49] points to a more direct use of eye-tracking as a performance measurement tool; the experiment focused mainly on using an eye tracker as a potential performance predictor within an aircrafts automated systems interaction with a pilot, the study found that:

> "In summary, testing monitoring behaviour using dynamic simulations based on eye movements is an innovative approach that enables the development of new methods of personnel selection. We identified time sensitive eye-tracking parameters to serve as basis for identifying OMA (Operators Monitoring Appropriately) in future selection processes. In this regard, we have shown that eye tracking parameters are predictive of failure-detection performance. Thus, the monitoring test (MonT) can be introduced as an effective tool for investigating human performance in future ATM scenarios. "[49: p.143]

This again justifies the use of gaze as a potential performance measurement tool. However, this author did not have access to this kind of equipment, due to financial constraints, and a more cost effective solution had to be found. It was decided that a commercially available head tracker, given a large enough target area, could be potentially used to replicate the use of an eye-tracker for performance measurement; this could only feasibly work with increased information set (main visuals, map, instrument displays) size. The creation of the display for simulated use is detailed further within the methodology but it is worthwhile to mention at this point that the simulator had to be designed in conjunction with other research project needs, this constrained the potential size of the displays as well as their positioning and lead to further experimental development following the case study.

In determining the potential for eye movement and potential field of vision while using a head tracker the physiological aspect of eye movement had to be explored. Howard Reed states [50]:

> "The normal monocular visual field extends approximately 100 degrees laterally, 60 degrees medially, 60 degrees upwards and 75 degrees downwards" [50: p.177]

&

> "The temporal crescent is part of the field of binocular field (of vision) which is always monocular, i.e. the crescentic area situated temporally beyond 60 degrees.”
>
> [50: p.177]

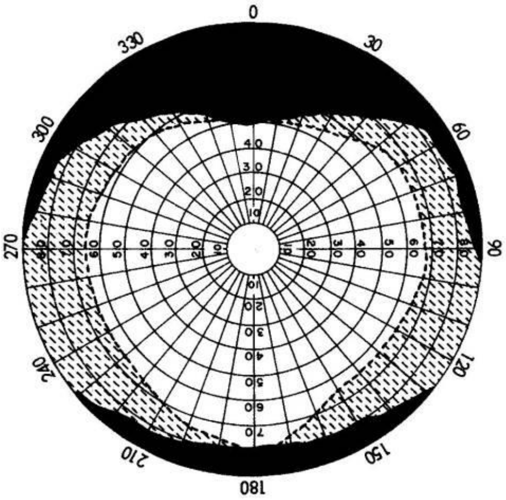


Figure 2.8 - Binocular field of view representation [51]

Figure 2.8 shows that an isolated field of view with static head movement would lead to a maximum 120 degrees of medial visual area and a 135 degrees of vertical visual area within the bounds of binocular vision, this does not however mean that a subject will necessarily utilise that entire field of vision and could well utilise head movement to aid in performance and comfort during the use of this visual area.

A report by the Environmental Protection Department of Hong Kong [52] suggests, based upon another study by J. Panero, that an optimum effective field of view is around 50 to 60 degrees around the x axis projected from the back of the head and between the eyes as can be seen in figure 2.9.

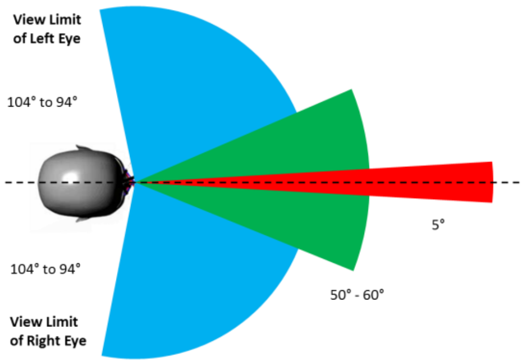


Figure 2.9 - Horizontal field of vision modified from [52]

2.2.5 Human Information Processing and Cognitive Factors

When making carrying out an action a person will base there action upon multiple forms of information and criteria; this process is highly based upon memory and the way in which both short term and long term memories are accessed for the resolution of the problem at hand.

Multiple theories exist, mostly based around Associationism memory, but these theories differ from each other with only some fundamental coherence between them. Many of the theories (such as Aristotle, British and American) do not share the same fundamentals as the others. [53]

One of the key fundamental areas agreed upon within Associationism are four key areas:

> "1. Ideas, sense data, memory nodes, or similar mental elements are associated together in the mind through experience. *Connectionistic*
>
> 2. The ideas can ultimately be decomposed into a basic stock of "simple ideas." *Redcutionisitic*
>
> 3. The simple ideas are to be identified with elementary, unstructured sensations." *Sensationalistic*
>
> 4. Simple, additive rules serve to predict the properties of complex associative configurations from the properties of the underlying simple ideas. *Mechanistic"* [53: p.10]

The commencement, analysis and completion of complex tasks is intrinsically linked to the use of memory, either by drawing on previous experience and/or by using short term memory as an 'information buffer'; with the above quotation indicating that memory itself has processes in which incident information is associated and analysed, either for storage or for application, it seems highly likely that task analysis will reflect that same form of structuring.

More recent research has since tried to apply the human approach to task analysis to automated systems [54]; these approaches consist of not just translating human motion for automation but also the observable and unobservable task sub-sections into an automated function. With the method presented by this research seemingly sound it should be theoretically possible to use objective data capture methods to measure human performance of operations by using the objective measurement of human reaction to incident information and by correlating to human output.

With both prior sources indicating that memory is key to the way in which information is processed and that tasks can be captured in an objective fashion that is applicable to automation it is justifiable in thinking that this can be applied to other forms of task capture and analysis. With memory being individual with an individual having different memories and prior experiences it is quite possible that the way in which an individual processes a task differing from that of another individual, however if the memories and experiences of a person are similar there could be a result in which two individuals complete the same task in a similar way. This is most likely for low complexity tasks while there may be greater divergence as task complexity, and routes in which to complete the task, become greater.

It may be that an optimum and robust method for task analysis and completion exists and it would be likely that this would become evident in a group with similar experiences, even with increasing workload level and complexity of task.

Depending on the complexity of the task a decision, or multiple decisions, are required to be made and are bounded by the environment in which they occur [55]; within dynamic decision making prior decisions and future responses are influenced by current decisions. These current decisions will be influenced by prior decisions and prior memories and experiences.

B. Nicholson & D. O'Hare investigated the ability to transfer this dynamic decision making ability between Gamers and Non-Gamers [56]; the results showed that the Gamer group were able to initially perform better under high and low workload tasks during practise and then perform the high workload objectives to a higher level than non-gamers with only a marginal ability increase over the non-gamers under low workload test conditions. This gives an indication that gamers were able to transfer prior task and decision making abilities from the practise sessions to a higher degree than the non-gamers; at a task level this would likely show that prior learning (memories and experience) have aided the gaming group initially and then further supplemented their ability to transfer task a decision making process to a new task more effectively than a non-gaming group. This would indicate that prior generic gaming experience and memories (processes) were adaptable to the new scenario. Although the experiment was a graphical based scenario it is likely that other forms of visual information (statistical, imagery and spatial) would be processed in a similar fashion by a gamer.

Patterns of information acquisition have been identified by a recent article in the Journal of Cognitive Engineering and Decision making[57]; these patterns are identified as a cyclical gathering of information by experienced pilots. This further suggest an experiential method of information collection and deployment in increasing workload level environment.

With the previous research in mind Wickens research into human factors and cognitive analysis can be applied. Cognitive Load Theory (CLT) [58] predicts that as learning and experience of a task increase the cognitive load upon the individual reduces allowing for increased complexity, or load, of task to be applied. This is potentially reflected in the ability of Gamers [56] to be able to process higher cognitive workloads than their non-game playing contemporaries due to a similarity of task (in terms of available information and application of that information encountered within a gaming environment.

2.3 Summary of Research Gaps

This chapter presents exploratory review of the ‘autonomous’ industrial and military field as well as reviewing operator licensing and performance measurement. This can be seen from the below category’s contained within this chapter:

- The history of RPAS
- The RPAS operator
- Mission Essential Competency and Evidence Based Training review
- Workload measurement
- Information set usage and field of vision
- Human Information Processing

Investigation of these areas lead to the identification of clear research gaps that experimentation could fill and, further to this, the formalisation of the research objectives that can be seen in chapter 1. In summary it was identified that there was limited published research within the following fields that could be supplemented by further investigation and experimentation:

- The objective performance measurement of RPAS operators
- The use of a head tracker as a performance analysis tool
- Investigation of information acquisition strategies of an RPAS operator
- The use of ‘information sets’ as a means to objectively performance measure an operator
- The use of ‘information sets’ as part of a task analysis structure
- Investigation of Gamers and Non-Gamers relating to their information acquisition strategies and how that relates to their task and flight performance

The following chapters are the result of the need to fill this research gaps.

Chapter 3 Methodology

This chapter details the process from case study through to final experimentation based upon concepts developed within the literature review. The key experimental concepts and issues were considered and the solution to these concepts and issues detailed.

This chapter can be seen as containing five fundamental areas:

- The development of a simulator for use with both the case study and the main experimentation
- The objectives of the case study
- The Case Study methodology
- Main experiment development
- Main experiment methodology including analytical methods

3.1 Simulator Development

At the beginning of the research it was identified that if empirical data were to be gathered a measurement medium would be required; this takes the form of a flight simulator that will allow for, not just flight measurement (measurement of flight variables such as altitude and speed), but also the use of a head tracker to meet the differing research objectives.

The simulator was developed in conjunction with two other colleagues, Mr. L. Le-Ngoc and Mr. C. Wright, as part of a multi-project simulation suite. This only included the functional hardware, all software and experimental setups were unique to each project and any software developed was solely used with this project and had no external contributions from any other parties.

Simulator development, mostly in terms of software, did not occur in a single instance but was an on-going process. Issues with the initial simulator setup were identified during the course Case Study and these issues had to be rectified with software changes and software creation.

3.1.1 Prototype Simulator Design

Requirements

The first stage of the actual simulator design was to identify the requirements of the simulator, figure 3.1 details this.

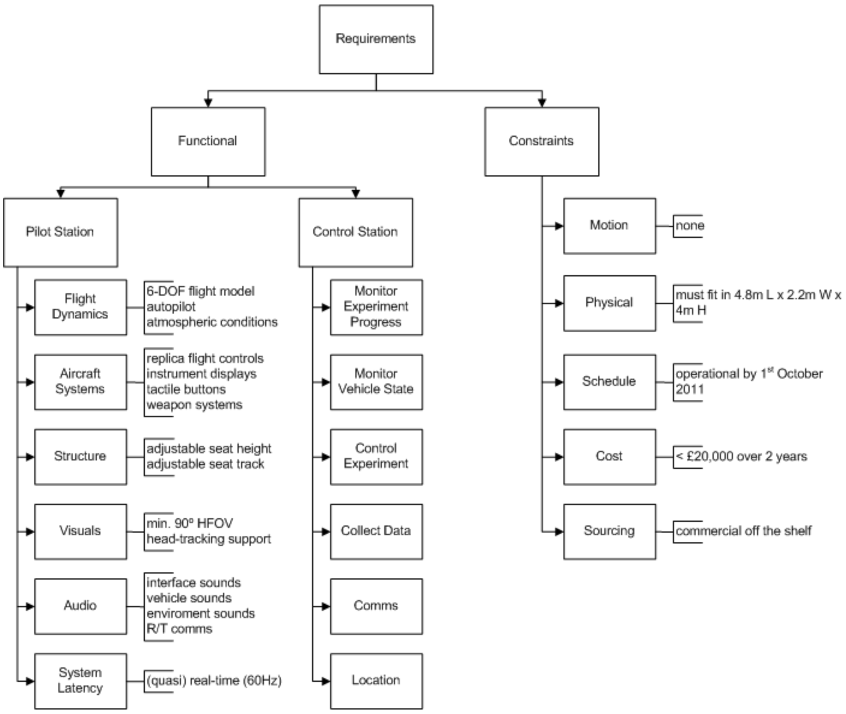


Figure 3.1 - Functional simulator requirements decided upon by G. Bedford, L. Le-Ngoc & C. Wright

The most important requirements are explored further below:

- To optimally use a head tracker as a means of data collection via information set usage observation a requirement was the use of a multi-screen display with each screen being of a large enough size to create large degrees of head movement to observe each individual information set.
- The simulation software must support some form of flight data collection ability as well as the ability to integrate and record head position data
- Simulation software must support an RPAS style flight model and have similar controls as well as information outputs.
- Flight controls must have an adequate level of fidelity to help promote participant simulation immersion.

What followed were discussions relating to the types of interface required, these ranged from a very specific replication of a fast jet type glass cockpit to a generic RPAS Ground Control Station based upon the Raytheon CGCS (Common Ground Control Station). This included the drafting of potential simulator configuration options, these can be seen below (figures 3.2-3.6).

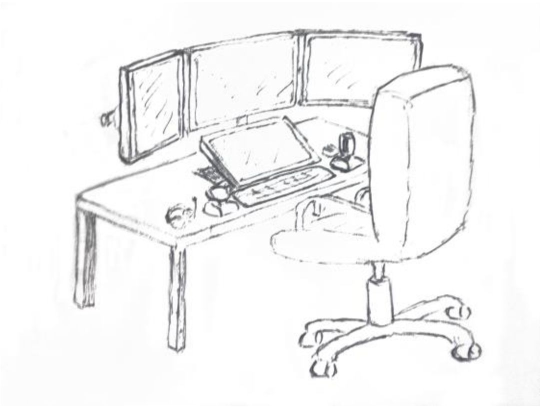


Figure 3.2 - Desk mounted quad screen interface

Figure 3.3 - Wide triple projection screen interface

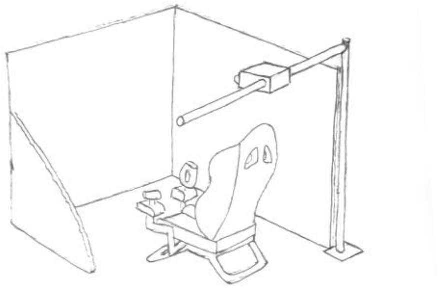


Figure 3.4 - Box projection screen interface

Figure 3.5 - Desktop hepta-screen interface

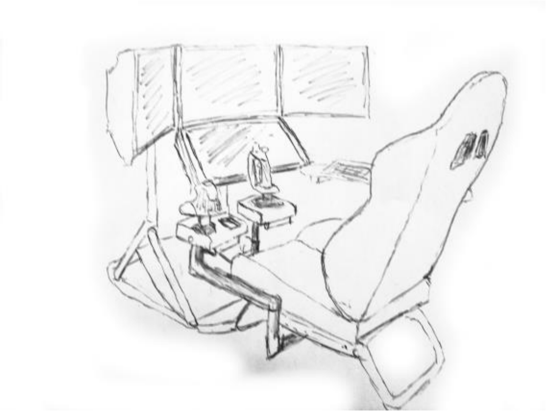


Figure 3.6 - Raytheon type CGCS quad screen interface

Key requirements were refined to four core fundamental requirements:

- High definition visual output to aid information acquisition
- A minimum of four screens to allow for segregation of disparate types of information
- A similar observation and control interface to that of existing RPAS platforms
- Computer ability to multi-task with minimal impact on speed of concurrent processes
- Large potential field of vision to encourage large degrees of head movement during disparate types of information gathering

The conceptual designs and requirements aided the decision to utilise a high performance gaming computer with triple 27" screens and a fourth 22" touch screen (figure 3.6). With respect to the field of vision section, in chapter 2.2.4, it was decided to use this screen configuration after calculating potential field of view for a single screen at a set head-to-screen distance (80 cm); to obtain a field of view for a single screen of 50° (deemed adequate to promote head movement during testing) the screen size required was 27”. The use of three of these screens, each angled to be perpendicular to the heads position, created a total field of regard of approximately 140° (when including screen bezels) in the horizontal plane; other configurations were calculated with differing screen-to-head distances but this was found to be the most viable, for calculations for differing fields of view please see table 3.1 below.

Table 3.1 - Screen Size Calculation (not including bezel correction)

Variables			**Output**	A	B	C	
Total Viewing Angle				120	135	140	deg HFOV
Viewing Distance	80	cm	**3xLandscape (curved)**				
Resolution	16:	9	Viewing Angle Per Screen (horizontal)	40.00	45.00	46.67	deg
			Viewing Angle (Vertical)	**23.63**	**26.95**	**28.08**	deg VFOV
			Screen Size (Imperial)	**26.73**	**30.42**	**31.67**	in
			3xPortrait (for same size screens as landscape)				
			Viewing Angle Per Screen (horizontal)	23.63	26.95	28.08	deg
			Viewing Angle (Vertical)	**40.00**	**45.00**	**46.67**	deg VFOV
			Total Viewing Angle horizontal (for portrait)	**70.88**	**80.84**	**84.25**	deg HFOV
			Screen Size (Imperial)	26.73	30.42	31.67	in
			3xPortrait (Flat)				
			Viewing Angle Peripheral Screen	19.74	21.48	22.01	deg HFOV
			Viewing Angle Centre Screen	23.63	26.95	28.08	deg HFOV
			Total Viewing Angle horizontal (for portrait)	40.00	45.00	46.67	deg HFOV
			Total Viewing Angle (horizontal)	**63.12**	**69.91**	**72.10**	deg HFOV

The 22” screen was decided upon due to its potential for use as an interactive unit (although this was, eventually, not to be a requirement); in terms of field of view it did not have the same requirements as the three 27” displays as it would be on its own, singular, vertical plane. The degree of head movement required to move from the 27” displays to the 22” display was calculated, with the use of the previous macro, and found to be adequate to promote head movement.

This was designated the GCS simulator. Only the GCS simulator was used during the course of this research.

The GCS simulator design was formalised as a similar layout to the cockpit but using a COTS home racing setup as can be seen in figure 3.7. The use of a ‘home racing’ setup provided a cost effective and adaptable solution for simulator construction and use.

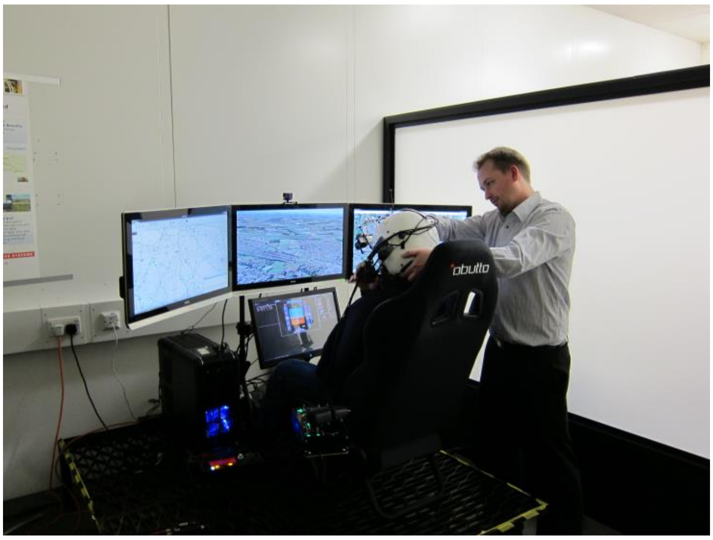


Figure 3.7 - GCS simulator final

As the simulator was required to run high performance graphical applications it was required to be a high performance gaming machine; the selection of the hardware is detailed in the following section.

PC Specification

It was deemed that experimentation had to be carried out on a high specification PC with high definition monitors (to aid with user immersion). A high degree of immersion is required to increase the level of realism (to a point suitable to the SME) and to aid participant involvement with the experimentation. The high specification of the PC was also required to make sure all the necessary software would run without any large diminishing of software performance.

The level of realism could only be tested by interaction with the participants and the feedback given; the feedback from the participant indicated that he was satisfied with the realistic nature of the simulator.

With these requirements in mind a custom made PC was ordered. This utilised:

- An intel i7 quad core processor to allow for a high level for processing power as well as multi-tasking ability
- Twin Nvidia 580 GTX graphics cards to allow for four high-definition and high-quality outputs to the three 27” and single 22” screen
- Single SSD hotswap bay to allow for inclusion of a master SSD 126GB drive; the SSD drive allowed for much faster hard drive operation

The above elements were considered to be the most important part of the PC itself but the PC, of course, includes other ancillary functions such as Ethernet/LAN support, Bluetooth (for wireless peripheral operation) and USB 3.0 support (to allow for faster peripheral data transfer).

Head tracking

Head tracking utilised a commercial-off-the-shelf TrackIR system, developed by Naturalpoint, and augmented with Naturalpoint's software developer kit (SDK) to use a revised version of the vector tracking algorithm. Development of this was carried out using Microsoft’s Visual Basic 2008. The flight data was analysed for screen usage posthoc using bespoke macros and formulae created within Microsoft Excel; once analysed and the flight transcribed and synced the flight data and head tacking data can be incorporated. These macros and development will be further detailed in the following section.

Display Screens

The display of the simulator had four information sets, each on its own screen; these are defined in appendix D along with each screen’s label.

One of the requirements of the simulator identified earlier within this section is that the screens are of a large order of size and that the participant sat at a set distance from these screens; this is to encourage head movement for observation of certain information sets as, once configured and oriented properly, they create an approximate 140° field of view coverage which allows for maximal use of simulator space and participant head movement.

The information sets were as follows:

- BenQ1 (IS1) = MFD map, Plan-G Flight Planner and information display
- BenQ2 (IS2) = Forward facing camera
- BenQ3 (IS3) = Google Earth real-time link
- IIyama (IS4) = Instrument Display and Autopilot Controls

Background lighting was also a factor; although a better contrast is obtained for the displays with no background lighting this also places greater strain on the participants eyes as well as causing potential health and safety issues while under experimentation (potential to trip or knock objects not visible in low light conditions). For experimental purposes a small desk lamp was utilised to provide enough light for health and safety while not diminishing the resolution of the main displays. Full specifications for the hardware can be found in appendix B.

Software (Case Study)

A full software specification list is available in appendix D.

Several requirements had to be met in the configuration of this simulator; these requirements were informed by currently available Ground Control Station designs and software. The first main requirement was the ability to display multiple information sets on multiple screens, this led to the first decision to use multiple screens and inferred the choice of cost effective simulator software, in this case Microsoft's Flight Simulator X (FSX). FSX was tested prior to simulator build for the multi-screen capability and was found to be acceptable and was also found to be acceptable for display quality (with the potential for inclusion of satellite imagery) as well as frame per second rate. The simulation software was further augmented with a generic RPAS Predator model purchased from Abacus which was modified to allow for a ‘first person’ camera view and a delete of the Heads-Up Display (HUD); this was then supplemented with VFR Real Scenery, which is satellite terrain imagery overlay for the simulator developed by horizon, to allow the operator direct visual cueing from the Google Earth display to the main camera visual display. (Latency issues of RPAS operation (operator input having a delay to corresponding platform output) were also considered when choosing software; the latency issues (when controlling over large distances) can be severe with potential for loss of the platform but, in terms of the task analysis nature of the experimentation, they were not considered as a governing factor when performing a task or competency and are more a fine point of specific platform control.)

The simulation software was also chosen for its developer support network; this network had already developed data loggers/outputers as well as an approximate RPAS Predator model. The software was then easily configurable for display and simulation purposes. The predator model was chosen for this experiment as it is one of the most widely known, and used, RPAS systems worldwide with both the USAF and the RAF using it in a military theatre.

Another software choice in Google Earth was informed by an SME contact as already being in use with the RAF, as well as worldwide with other military organisations, as a planning and navigational tool; due to certain security acts it is not possible to substantiate this. Other potential map outputs (such as the FSX in built map as well as other FSX plug-ins) were found to be lacking in detail and fidelity and so were excluded as candidates for experimental use.

It was therefore desirable to be able to incorporate a Google Earth display into the simulation, again the developer network for Flight Sim X had already created an effective solution thus reducing developmental schedules. Pictures can also be layered upon Google earth, this allows for the potential for “spotting” of desired images, shapes or places; the “spotting” could potentially be incorporated into experimentation as a demonstration of information acquisition.

During experimentation three pieces of data logging software were used, one for specific aircraft performance, one as a co- ordinate output to Google Earth and the third logging head movements.

The Google Earth Logger (Blackbox) not only provides a real-time link to Google Earth but also records the flight in a re-playable KML file.

> “Keyhole Markup Language (KML) is an XML notation for expressing geographic annotation and visualization within Internet-based, two-dimensional maps and three-dimensional Earth browsers. KML was developed for use with Google Earth, which was originally named Keyhole Earth Viewer. It was created by Keyhole Inc. which was acquired by Google in 2004. KML became an international standard of the Open Geospatial Consortium in 2008. Google Earth was the first program able to view and graphically edit KML files. Other projects such as Marble have also started to develop KML support.” – [Keyhole Markup Language, Wikipedia, March 2014]

The systems logger (Vivendobyte) outputs data into a CSV format which can be accessed through Microsoft Excel for analysis. The analysis of the raw data can be carried out manually or with Vivendobyte's bespoke Excel Macro.

A further software module called Plan-G was used for mission planning and as a means to change automatic flight plans; information received from the RAF has led to the need for 95-98% of the platform control to be handled by the simulator, this includes almost all flight control (pitch, yaw and roll) with the exception of take-off and landing. This can be done using existing autopilot software built into FSX but the autopilot must have an updateable flight plan; Plan-G allows for flight planning without the use of the simulation software and for flight plan changes to be made during the course of simulation. However Plan-G is not a dynamic flight planner and the updated flight plan must be uploaded to FSX and then the desired waypoint selected.

This did not, however, affect the experimental procedure because using/setting the autopilot is a task in itself.

During the course of the experimentation the participant was asked to fly a pre- planned basic way point mission to the best of his ability. The way points for this mission were marked on both aircraft map and Google earth and bearing from previous way point to new way point will also be given verbally.

Pilot seat and display screen mounts

The Obutto racing platform (figure 3.8) was chosen as an adequate (comfortable and usable in terms of a GCS type arrangement) seating and mounting platform and enables the mounting of all four screens with some minor modification and the addition of an extra screen clamp for the fitment of the fourth screen below screen 2. It was seen that this configuration would be acceptable being similar in configuration to the Raytheon CGCS. The final simulator configuration can be seen in figure 3.7. Full specification for software can be found in appendix D.

Figure 3.8 – Obutto Racing Platform

Controls

The flight control interface for the case study utilised a Saitek stick, throttle and rudder configuration (later changed to Thrustmaster HOTAS stick and throttle for the main experiment due to the need for a more intuitive and realistic control setup). The layout of these controls again mimicked that of the Raytheon CGCS as well as the Predator GCS in which the stick is placed to the right of the participant, the throttle to the left and the rudder at the participant's feet. This positioning is comfortable for extended periods of operation.

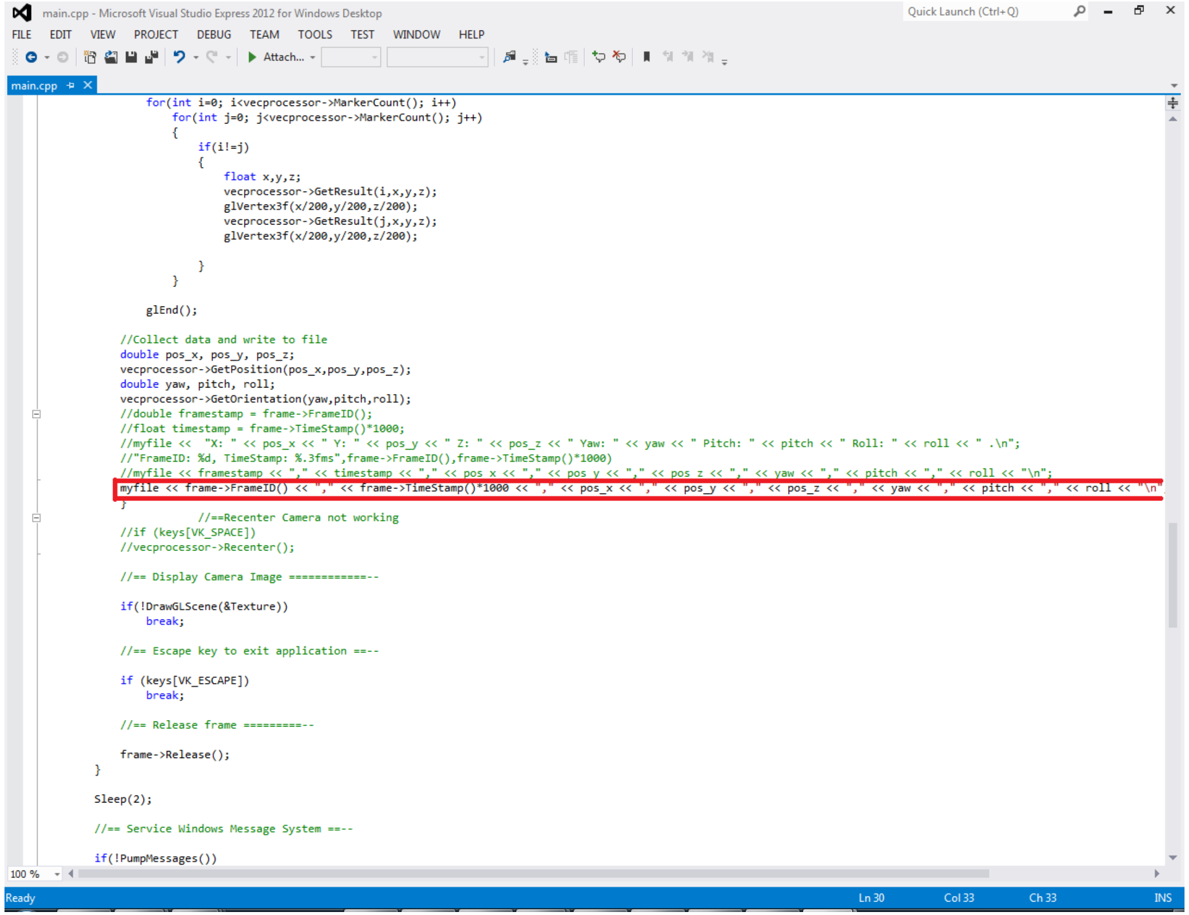
3.1.2 Analysis Software Development (Case Study)

To allow measurement of head tracking using COTS hardware as well as measurement of this data relating to the transcript and flight data it was necessary, firstly, to adapt an existing program, in this case Naturalpoint's Vector Tracking project, to allow for head position data output and, secondly, to perform functions upon this data after syncing it with the flight transcript. The flight transcript was a running commentary from the participant which involved identification of participant acknowledged screen changes and key task points (including identification of task switching and task process implementation); the head data to transcript sync involved creating a time stamped spreadsheet for the transcript and then applying the time stamped constraints, created within the transcript, to the head tracking and flight data. Head and flight data can then be processed between those constraints and the processed values applied to the relevant section of the transcript.

The program used for acquiring the head position data was part of a SDK package released by Naturalpoint (www.naturalpoint.com), the creators of the TrackIR hardware and released under their development subsidiary name of Optitrak. TrackIR has two forms, normal and track clip, the first form utilises a single combined sensor and emitter which projects infrared light into the area in front of the sensor, it then detects the position and orientation of three reflective surfaces (mounted on a cap or other head mounted item) and uses this orientation to determine the head position; this head position could be determined to 1/100th of a degree shift and would capture a value every tenth of a second for, not only time, but also movement in the x, y, z axis and pitch, yaw and roll (6 degrees of freedom).

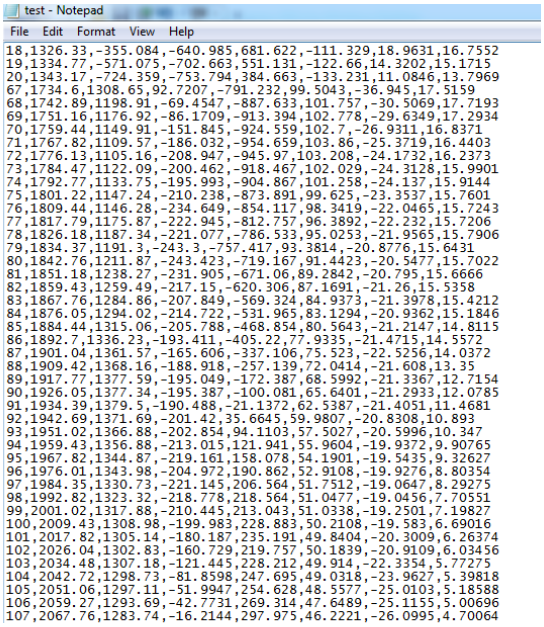
The vector tracking program (Naturalpoint), which only works with the normal reflective method, displays a graphical representation of the positioning of the reflecting sensors; by modifying the vector tracking program, by the insertion of code (figure 3.9) that prints the head tracker variables and time stamp to a CSV (Comma Separated Variable) file, it is possible to use this positioning to produce a basic CSV output containing time as well as pitch, yaw, roll and x, y, z axis movement. It is this output that was used during the initial experimentation to determine screen usage. Figure 3.9 shows the small section of code input to allow data output to file while 3.10 shows the output data.

The highlighted code in figure 3.9 shows the variables captured during the vector tracking (frame, time, positions x, y, z and pitch, yaw and roll) process being printed to “myfile” which is the output CSV file.



```
            for(int i=0; i<vecprocessor->MarkerCount(); i++)
                for(int j=0; j<vecprocessor->MarkerCount(); j++)
                {
                    if(i!=j)
                    {
                        float x,y,z;
                        vecprocessor->GetResult(i,x,y,z);
                        glVertex3f(x/200,y/200,z/200);
                        vecprocessor->GetResult(j,x,y,z);
                        glVertex3f(x/200,y/200,z/200);

                    }
                }

            glEnd();

        //Collect data and write to file
        double pos_x, pos_y, pos_z;
        vecprocessor->GetPosition(pos_x,pos_y,pos_z);
        double yaw, pitch, roll;
        vecprocessor->GetOrientation(yaw,pitch,roll);
        //double framestamp = frame->FrameID();
        //float timestamp = frame->TimeStamp()*1000;
        //myfile <<  "X: " << pos_x << " Y: " << pos_y << " Z: " << pos_z << " Yaw: " << yaw << " Pitch: " << pitch << " Roll: " << roll << " .\n";
        //"FrameID: %d, TimeStamp: %.3fms",frame->FrameID(),frame->TimeStamp()*1000)
        //myfile << framestamp << "," << timestamp << "," << pos_x << "," << pos_y << "," << pos_z << "," << yaw << "," << pitch << "," << roll << "\n";
        myfile << frame->FrameID() << "," << frame->TimeStamp()*1000 << "," << pos_x << "," << pos_y << "," << pos_z << "," << yaw << "," << pitch << "," << roll << "\n"
        }
                    //==Recenter Camera not working
        //if (keys[VK_SPACE])
        //vecprocessor->Recenter();

        //== Display Camera Image ============--

        if(!DrawGLScene(&Texture))
            break;

        //== Escape key to exit application ==--

        if (keys[VK_ESCAPE])
            break;

        //== Release frame =========--

        frame->Release();
    }

    Sleep(2);

    //== Service Windows Message System ==--

    if(!PumpMessages())
```

Figure 3.9 - Vector tracking data output code insertion



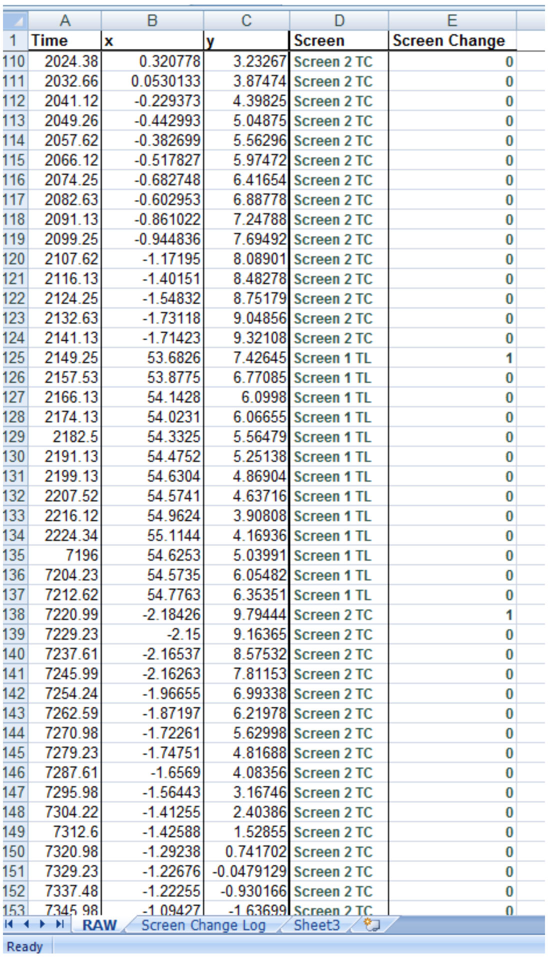
```
18,1326.33,-355.084,-640.985,681.622,-111.329,18.9631,16.7552
19,1334.77,-571.075,-702.663,551.131,-122.66,14.3202,15.1715
20,1343.17,-724.359,-753.794,384.663,-133.231,11.0846,13.7969
67,1734.6,1308.65,92.7207,-791.232,99.5043,-36.945,17.5159
68,1742.89,1198.91,-69.4547,-887.633,101.757,-30.5069,17.7193
69,1751.16,1176.92,-86.1709,-913.394,102.778,-29.6349,17.2934
70,1759.44,1149.91,-151.845,-924.559,102.7,-26.9311,16.8371
71,1767.82,1109.57,-186.032,-954.659,103.86,-25.3719,16.4403
72,1776.13,1105.16,-208.947,-945.97,103.208,-24.1732,16.2373
73,1784.47,1122.09,-200.462,-918.467,102.029,-24.3128,15.9901
74,1792.77,1133.75,-195.993,-904.867,101.258,-24.137,15.9144
75,1801.22,1147.24,-210.238,-873.891,99.625,-23.3537,15.7601
76,1809.44,1146.28,-234.649,-854.117,98.3419,-22.0465,15.7243
77,1817.79,1175.87,-222.945,-812.757,96.3892,-22.232,15.7206
78,1826.18,1187.34,-221.077,-786.533,95.0253,-21.9565,15.7906
79,1834.37,1191.3,-243.3,-757.417,93.3814,-20.8776,15.6431
80,1842.76,1211.87,-243.423,-719.167,91.4423,-20.5477,15.7022
81,1851.18,1238.27,-231.905,-671.06,89.2842,-20.795,15.6666
82,1859.43,1259.49,-217.15,-620.306,87.1691,-21.26,15.5358
83,1867.76,1284.86,-207.849,-569.324,84.9373,-21.3978,15.4212
84,1876.05,1294.02,-214.722,-531.965,83.1294,-20.9362,15.1846
85,1884.44,1315.06,-205.788,-468.854,80.5643,-21.2147,14.8115
86,1892.7,1336.23,-193.411,-405.22,77.9335,-21.4715,14.5572
87,1901.04,1361.57,-165.606,-337.106,75.523,-22.5256,14.0372
88,1909.42,1368.16,-188.918,-257.139,72.0414,-21.608,13.35
89,1917.77,1377.59,-195.049,-172.387,68.5992,-21.3367,12.7154
90,1926.05,1377.34,-195.387,-100.081,65.6401,-21.2933,12.0785
91,1934.39,1379.5,-190.488,-21.1372,62.5387,-21.4051,11.4681
92,1942.69,1371.69,-201.42,35.6645,59.9807,-20.8308,10.893
93,1951.02,1366.88,-202.854,94.1103,57.5027,-20.5996,10.347
94,1959.43,1356.88,-213.015,121.941,55.9604,-19.9372,9.90765
95,1967.82,1344.87,-219.161,158.078,54.1901,-19.5435,9.32627
96,1976.01,1343.98,-204.972,190.862,52.9108,-19.9276,8.80354
97,1984.35,1330.73,-221.145,206.564,51.7512,-19.0647,8.29275
98,1992.82,1323.32,-218.778,218.564,51.0477,-19.0456,7.70551
99,2001.02,1317.88,-210.445,213.043,51.0338,-19.2501,7.19827
100,2009.43,1308.98,-199.983,228.883,50.2108,-19.583,6.69016
101,2017.82,1305.14,-180.187,235.191,49.8404,-20.3009,6.26374
102,2026.04,1302.83,-160.729,219.757,50.1839,-20.9109,6.03456
103,2034.48,1307.18,-121.445,228.212,49.914,-22.3354,5.77275
104,2042.72,1298.73,-81.8598,247.695,49.0318,-23.9627,5.39818
105,2051.06,1297.11,-51.9947,254.628,48.5577,-25.0103,5.18588
106,2059.27,1293.69,-42.7731,269.314,47.6489,-25.1155,5.00696
107,2067.76,1283.74,-16.2144,297.975,46.2221,-26.0995,4.70064
```

Figure 3.10 - Vector tracking data output (time, frame number, x, y, z, pitch, yaw, roll)

These data sets were then imported, sorted, filtered and ordered within Microsoft Excel using custom macros; these macros were created specifically for this data analysis by the researcher. The initial import process involved a basic CSV data import with pre- defined criteria for the import (specific columns and rows targeted for import, in this case time, pitch and yaw) and then a subsequent automated analysis of the data using predefined degree boundaries to identify which screen the head position indicated; the code associated with this screen identification is shown below, it is based on a cascading logical analysis of the x (pitch) and y (yaw) variables.

```
=IF(AND(B2>13, B2<60, C2>-24), "Screen 1 TL", IF(AND(B2<12, B2>-10, C2>-30.6), "Screen 2 TC", IF(AND(B2<-11, B2>-65, C2>-30), "Screen 3 TR", IF(AND(B2>29, B2<70, C2<-25), "Screen 4 BC", "Void"))))
```

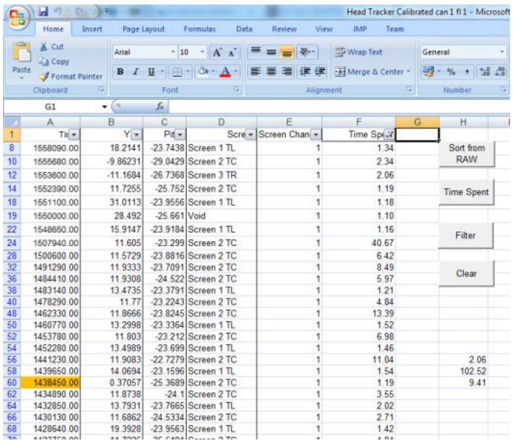
What followed was another part of the automated process which determined if there had been a screen change between one row of data and the next, if a screen change occurred (movement from screen 1 to screen 2 for example) the screen change column would return a 1 and, if not, a 0. As only the time and head yaw and pitch were to be considered these were the only data columns imported. This was further supplemented by the inclusion of 'void' zones between screens, as can be seen in the code on the previous page; the final output for the screen cells was 'Screen 1 (Top Left)', 'Screen 2 (Top Centre)', 'Screen 3 (Top Right)', 'Screen 4 (Bottom Centre)' and 'Void' as can be seen in figure 3.11.



	A	B	C	D	E
1	Time	x	y	Screen	Screen Change
110	2024.38	0.320778	3.23267	Screen 2 TC	0
111	2032.66	0.0530133	3.87474	Screen 2 TC	0
112	2041.12	-0.229373	4.39825	Screen 2 TC	0
113	2049.26	-0.442993	5.04875	Screen 2 TC	0
114	2057.62	-0.382699	5.56296	Screen 2 TC	0
115	2066.12	-0.517827	5.97472	Screen 2 TC	0
116	2074.25	-0.682748	6.41654	Screen 2 TC	0
117	2082.63	-0.602953	6.88778	Screen 2 TC	0
118	2091.13	-0.861022	7.24788	Screen 2 TC	0
119	2099.25	-0.944836	7.69492	Screen 2 TC	0
120	2107.62	-1.17195	8.08901	Screen 2 TC	0
121	2116.13	-1.40151	8.48278	Screen 2 TC	0
122	2124.25	-1.54832	8.75179	Screen 2 TC	0
123	2132.63	-1.73118	9.04856	Screen 2 TC	0
124	2141.13	-1.71423	9.32108	Screen 2 TC	0
125	2149.25	53.6826	7.42645	Screen 1 TL	1
126	2157.53	53.8775	6.77085	Screen 1 TL	0
127	2166.13	54.1428	6.0998	Screen 1 TL	0
128	2174.13	54.0231	6.06655	Screen 1 TL	0
129	2182.5	54.3325	5.56479	Screen 1 TL	0
130	2191.13	54.4752	5.25138	Screen 1 TL	0
131	2199.13	54.6304	4.86904	Screen 1 TL	0
132	2207.52	54.5741	4.63716	Screen 1 TL	0
133	2216.12	54.9624	3.90808	Screen 1 TL	0
134	2224.34	55.1144	4.16936	Screen 1 TL	0
135	7196	54.6253	5.03991	Screen 1 TL	0
136	7204.23	54.5735	6.05482	Screen 1 TL	0
137	7212.62	54.7763	6.35351	Screen 1 TL	0
138	7220.99	-2.18426	9.79444	Screen 2 TC	1
139	7229.23	-2.15	9.16365	Screen 2 TC	0
140	7237.61	-2.16537	8.57532	Screen 2 TC	0
141	7245.99	-2.16263	7.81153	Screen 2 TC	0
142	7254.24	-1.96655	6.99338	Screen 2 TC	0
143	7262.59	-1.87197	6.21978	Screen 2 TC	0
144	7270.98	-1.72261	5.62998	Screen 2 TC	0
145	7279.23	-1.74751	4.81688	Screen 2 TC	0
146	7287.61	-1.6569	4.08356	Screen 2 TC	0
147	7295.98	-1.56443	3.16746	Screen 2 TC	0
148	7304.22	-1.41255	2.40386	Screen 2 TC	0
149	7312.6	-1.42588	1.52855	Screen 2 TC	0
150	7320.98	-1.29238	0.741702	Screen 2 TC	0
151	7329.23	-1.22676	-0.0479129	Screen 2 TC	0
152	7337.48	-1.22255	-0.930166	Screen 2 TC	0
153	7345.98	-1.09427	-1.63699	Screen 2 TC	0

Figure 3.11 - Head tracking output data demonstrating success of screen identification and identification of screen change

These "raw" data sets were then further filtered by another macro to import only those rows of data with a positive result for screen change into a new sheet, these results were then analysed by subtracting the initial rows time value with that of the successive row, this gave amount of time spent on each screen before a screen change occurred. These data sets were further filtered by removing data with time values of less than an inferred data acquisition time frame, in this case 0.5 seconds; this value was derived through simulator testing with the 0.5 seconds including time to move the head into a position in which the information could be observed adequately (roughly around 0.3 seconds) and then acquired, this left only 0.2 seconds for information acquisition which was felt to be the absolute minimum for any meaningful information acquisition. This was to remove any screens that were looked at that the pilot could not possibly have retrieved data from. This time frame was purely a subjective interpretation by the researcher through his own experiences and that of discussions with fellow researchers and supervisors and was not based on any other prior research. Figure 3.12 shows this filtered and processed data; it can be seen that the screen change column only display's 1's indicating that only a transition from one screen to another has been logged with the respective screen changed to shown in the 'Screen' column. The time column represents when this changed occurred with 'Time Spent' the duration of this screen usage; the 'x' and 'y' values were a redundant check of head position to check that the screen position identifier was working correctly.



	A	B	C	D	E	F	G	H
1	Ti	Y	Pit	Scre	Screen Chan	Time Spe		
8	1558090.00	18.2141	-23.7438	Screen 1 TL	1	1.34		
10	1555680.00	-9.86231	-29.0429	Screen 2 TC	1	2.34		
12	1553600.00	-11.1684	-26.7368	Screen 3 TR	1	2.06		
14	1552390.00	11.7255	-25.752	Screen 2 TC	1	1.19		
18	1551100.00	31.0113	-23.9556	Screen 1 TL	1	1.18		
19	1550000.00	28.492	-25.661	Void	1	1.10		
22	1548650.00	15.9147	-23.9184	Screen 1 TL	1	1.16		
24	1507940.00	11.605	-23.299	Screen 2 TC	1	40.67		
28	1500600.00	11.5729	-23.8816	Screen 2 TC	1	6.42		
32	1491290.00	11.9333	-23.7091	Screen 2 TC	1	8.49		
36	1484410.00	11.9308	-24.522	Screen 2 TC	1	5.97		
38	1483140.00	13.4735	-23.3791	Screen 1 TL	1	1.21		
40	1478290.00	11.77	-23.2243	Screen 2 TC	1	4.84		
48	1462330.00	11.8666	-23.8245	Screen 2 TC	1	13.39		
50	1460770.00	13.2998	-23.3364	Screen 1 TL	1	1.52		
52	1453780.00	11.803	-23.212	Screen 2 TC	1	6.98		
54	1452280.00	13.4989	-23.699	Screen 1 TL	1	1.46		
56	1441230.00	11.9083	-22.7279	Screen 2 TC	1	11.04		2.06
58	1439650.00	14.0694	-23.1596	Screen 1 TL	1	1.54		102.52
60	1438450.00	0.37057	-25.3689	Screen 2 TC	1	1.19		9.41
62	1434890.00	11.8738	-24.1	Screen 2 TC	1	3.55		
64	1432850.00	13.7931	-23.7665	Screen 1 TL	1	2.02		
66	1430130.00	11.6862	-24.5334	Screen 2 TC	1	2.71		
68	1428640.00	19.3928	-23.9563	Screen 1 TL	1	1.42		

Figure 3.12 - Filtered data with ‘time’ in milliseconds and ‘time spent’ in seconds

The final output data from the experimentation required the manual sync of the screen, transcript and flight data) so that all three data sources could be analysed concurrently; in future experimentation two of these data sources (flight and screen) would be automatically sync’d by the simulation software itself with the transcript data becoming redundant for the revised main experiment.

The flight data would be logged using a pre-designed logging tool called VivendoByte (www.fsxlogger.vivendobyte.net), this would output multiple parameters of flight data and eventually be aligned with this head tracking data using a transcript of the flight.

Figure 3.13 shows a systems architecture for the data collection and analysis.

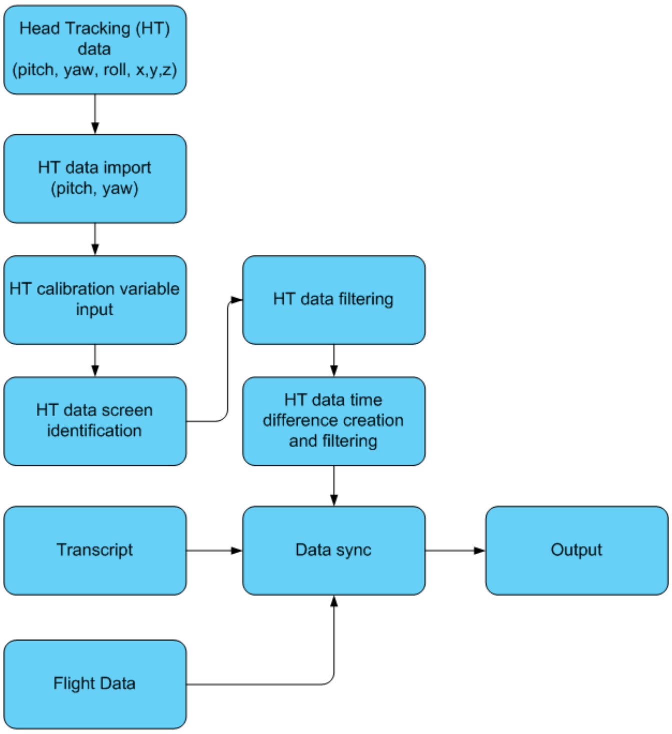


Figure 3.13 - Data handling architecture

What follows is the full experimental plan for the case study with a brief overview of pertinent findings associated with the research objectives (see chapter 1.2).

3.2 Case Study

The problem statement for case study: Is it possible to measure pilot performance objectively while undergoing RPAS flight simulation in terms of a task based metric and what development is required to support robust data collection and analysis?

The aims are to

- Measure the usage of independent information sets (i.e. flight data, positional data and main visual data) with the use of a commercial-off-the-shelf head tracker augmented by custom post-processing metrics
- Create a task-phase breakdown using pilot transcripts
- Apply the measured head position (information set usage) and flight data to the task phase breakdown and identify any performance markers
- Identify any issues leading to potential problems associated with this type of measurement (either software or hardware) or any other problems that could lead to non-relevant results

The case study was developed to aid research progress and to derive a better understanding of the problems and issues faced with future experimentation; there does not seem to have been a similar use of commercial-off-the-shelf (COTS) hardware available to the researcher for this type of experimentation with most similar research utilising much more expensive eye-tracking equipment rather than head tracking (see chapter 2.2.4) so it was necessary to evaluate this hardware, for this purpose, in an experimental environment. Also to be understood were the way in which the tasks concepts, developed from information contained in the literature review and detailed within the 'Findings and Analysis' chapter, could be applied to this experimental configuration.

An original research aim was to create a task list which could then be applied to the revised theoretical Mission Competency structure developed within the literature review. These tasks would consist of both researcher, operator and data observation of the pilot's flight and would be sub-analysed for supporting Knowledge and Skills required of the operator. Further to this the data would be analysed to try and provide a way in which to measure either the supporting knowledge and skills or the task itself; it was hoped that this analysis would help form a baseline for performance measurement of later experimentation to identify similarities between pilots and their task data and, therefore, justify the creation and use of a task based RPAS evaluation tool.

The case study became more of a problem-finding study to help develop and inform later experimentation and research objectives rather than as a research basis on which to prove hypotheses surrounding the original research aims.

3.2.1 Study Objectives

The following study was designed to create an understanding of the tasks that current RAF Predator operators must use and complete during the course of a mission, this was derived through questionnaires and discussion with subject matter experts (RAF Reaper pilots). The experimentation explores tasks of the most basic level mostly concerning basic flight control, take-off and landing and navigation.

It was suggested, by a subject matter expert [59] , that up to 95-98% of the fundamental flight control is automated with only take-off and landing being purely manual (at this time) on the Predator platform, although it would have been ideal to have multiple SME inputs for this it was impossible to contact further SME's due to availability. The automatic operations were likely to initially have a human input from the operator so, even though the operator is not operating the vehicle, it was possible to analyse the tasks required to prepare the system for automatic operation.

There are two main aims of this study; the first being to create a database of tasks (including knowledge and skills) required of the operator for basic flight control and the second being the understanding of the way in which the operator uses information and then applies the correct response.

The creation of the task database is key to this experiment and to future experimentation and was achieved mainly through a dynamic iterative process involving questionnaires and a basic mission walkthrough by the SME; a basic navigation mission was then presented to the participant to verify his questionnaire and interview responses. The use of only a single participant for experimentation is further explained in chapter 3.3.3.

Each discovered task included information regarding:

- stage of mission task started
- information observed for task completion
- the decision process based upon available information
- action taken for task completion
- stage of mission task completed at

The second aim was to verify how the operator uses available information, both consciously and sub-consciously. This was done by presenting the operator with a new navigational challenge but denying him access to certain data sets; these data sets, consisting of a main visual output (or virtual world, screen 2), an flight data information/control display (heads down display, screen 4) and a 2D planning map display (screen 1, not using satellite imagery) and a 2D dynamic map display (using satellite imagery, screen 3). The removal of a certain data set allowed for examination of the candidates strategies to compensate for the loss of information by using other data sets to regain (if possible) the loss of that potentially key data; for instance the loss of the planning display during flight was compensated by increased use of the dynamic map display. Which data set is used for compensation is directly related to the type of task being performed by the candidate at the time. Scores taken from this phase of experimentation were compared to previous scores as well as operator decisions regarding necessary data sets.

This initial experimentation led on to further experimentation with non-expert participants and was used to put two sets of participants through a task learning cycle as well as being used as a gauge as to how well the non-expert participants perform and use information in comparison to the SME's.

3.2.2 Study Objectives Overview

The first research objective was to investigate whether a basic task list could be created with relevant data based upon the information, decisions and actions for each respective task as well as for performance scores.

The second objective was to determine whether denial of information sets (turning off a screen with an information set), required by a specific task, will lower the performance of the operator within that given task while tasks that do not require that information sets will be relatively unaffected.

The third objective was to determine whether all of the defined tasks, that were generated as part of the first objective, would be able to be grouped into both phase of mission and type groups to allow for a generalised performance measurement.

With the required groups and task lists being generated within the previous objectives this lead on to a basis for more rigorous testing of the basic task list with non-expert participants who will be male and of the age group 18-35 which adheres to the RAF recruitment age group.

Due to the constraints of further experimentation only one task could be examined in further experimentation.

3.2.3 Subject Matter Expert and Pre-flight Analysis

The acquisition of subject matter experts for this study proved extremely challenging as the target group, RAF Predator/Reaper pilots, are of limited numbers and, until recently, were solely based at Creech Air force base in North America. Although it was hoped that many participants would become available it transpired that only a single participant could be found for experimentation.

What follows is a brief description of this pilot and his background; the description demonstrates the use of already professional military pilots being transferred to RPAS flight. Only information relevant to the experimentation will be discussed here to protect the pilot's anonymity.

Other target groups, such as commercial sector operators, were deemed unacceptable for experimentation mainly due to operators of corresponding medium to large RPAS platforms both being extremely few in number and also lacking experience; smaller platform operators were deemed unacceptable due to their lack of experience with the experimental platforms level of automation and it's increased amount of information displays over the commercial small platform norm.

The participant's responses to the initial questionnaire (see Appendix D - Research Output, Questionnaire A) identified him as an experienced military pilot with a background in operation of Hercules C130 transport aircraft with several thousand hours of manned flight experience as well as over one thousand hours of unmanned flight experience.

A point of note is the comparatively small amount of time spent in simulated unmanned flight training compared to real unmanned flight training, in fact simulated unmanned flight only represents 6.25% of his total training flight time. This could be due to his previous experience with manned flight but it also demonstrates the large degree of confidence that the trainers have in their trainees in terms of performance using an Unmanned Aerial Vehicle.

From a budgetary aspect this is surprising as an RPAS system more closely resembles a normal flight simulator than that of manned simulator versus manned flight; it would have been thought that much more training could have been carried out using an RPAS simulator for training as this would be much more cost effective and give lower potential risk of the destruction of an expensive piece of equipment and collateral damage.

The actual re-training of the pilot (identified by the participant in questionnaire A) closely resembles the theoretical task based training timeline model that is further investigated in section:

- ‘Effects of Controls’ representing ‘Basic Control’
- ‘ISR (intelligence gathering)’ representing ‘Interpretation’
- ‘Weapons Employment’ representing ‘Employment’

Communications were not a likely training focus as the pilot should already have a high degree of relevant communications experience which would not be platform, manned or unmanned specific. The ‘close air support’ module would represent the next level of training in which the training no longer relates specifically to the platform but to the inter-team competencies. This type of training demonstrates the potential effectiveness of an evidence-based modular system.

As this questionnaire was submitted some time before the experimentation it was possible to adjust the experimental methodology and equipment configuration to better replicate the design and operation of the Predator GCS.

3.2.4 Case Study Methodology

Questionnaire

Before experimentation could be carried out there had to be an understanding of the participant’s background and data were taken of the participant’s relevant experience (see Appendix D). This consisted of gathering information regarding the participant’s experience of both manned and unmanned aircraft and included approximations for time logged for both formats of flight and what type of platform was flown.

This led on to further set of questions regarding RPAS operation and mission experience. This set of questions started the process of task identification and formed a basis for the observed section of the experimentation.

The next set of questions was in regard to the simulator setup and Ground Control Station (GCS) operation. The participant was introduced to a 'on runway' display of the simulator with information sets displayed as they would be during flight. This was to identify the participants interpretations of the fidelity, as well as comfort, of the simulator and how this could be related to a current RPAS GCS.

The participant was then allowed to test fly the system for approximately 10 minutes. Once the participant had test flown the system he was asked a further set of questions relating to the simulator’s respective performance and realism compared to a real world Ground Control Station.

Due to the potential security restrictions of questioning an active military RPAS operator the questions where of a general nature and where not directed at anything that may have been covered by the official secrets act; the operator however was free to decline to answer if he felt uncomfortable.

Observation

The second phase of the experiment required the participant to fly a pre-planned route of take-off, 2 waypoints and landing. Prior to commencing the simulation the participant was asked to notify the experimenter verbally of any tasks he was about to implement, the information set(s) that led him to identify the need for the task (if any), the information sets he required to complete the task, the decision he made to complete the task and a notification when he felt the task has been completed successfully. An audio recording was made of all proceedings with the transcript from the audio recording logged, by the experimenter, into an excel spreadsheet. There was potential for the participants natural cognitive functions to be affected by this process but, due to the nature of the experiment, it was not possible to identify whether this was the case; this issue was to be rectified in later experimentation where an audio log of actions was no longer required.

With the current configuration of the simulator certain head movement implied the participant's use of the relevant information sets so a log was kept of the participant's head movements with the use of head tracking hardware; this potentially showed the sub-conscious use of data sets by the participant although was possible that an information set was not observed while the head position showed a certain information set usage.

Before the experiment commenced the participant was given 5 minutes to familiarise himself (at that time only one male RAF participant was available) with the information set position and controls and then a further 5 minutes of live flight (10 minutes total) so that they could understand the flight dynamics of the simulated Predator flight model. A crib sheet of controls was also provided for reference; use of the crib sheet itself was not logged and would only show as "void" when associated with the head positioning data.

This was followed by a questionnaire.

Experiment Timeline:

- 5 minutes of familiarisation with simulator display and controls
- 5 minutes of live (manually controlled) flight to familiarise with the flight model
- Questionnaire

- Simulator reset and flight plan loaded
- Audio recording started
- Participant given instructions regarding experimentation
  - Please verbally identify the following
    - Task start and type
    - Information sets observed to cue task start (if any)
    - Information sets used to monitor task or make task decision
    - Decision made
    - Task Completion with success or fail
- Participant asked to configure Autopilot/Flight Plan for two waypoints and return (if applicable)
- Data loggers started and participant asked to input throttle response for sync purposes
- Participant asked to move head in set sequence to calibrate time stamp for head tracker
- Participant asked to commence take off
- Participant asked to fly to waypoint 1 (if applicable)
- Participant asked to create new waypoint (for set location) using the Plan-G software, upload to the simulator and configure autopilot accordingly (if applicable)
- Participant asked to identify target using IS3 display (if applicable)
- Participant asked to fly to waypoint 2 (if applicable)
- Participant asked to approach airfield and land
- Debrief

Once completed the pilot was asked to review the created task list and rectify any task data as necessary.

A further description of each task was required of the participant in which he must specify

- Needs of task

- Objective of task

This was added to the task list spreadsheet.

Human Interaction

This phase of experimentation was used to determine the participant's use of available data sets. It is likely that the previously used head tracker may not fully identify actual participant use of the available information sets and, therefore, experimentation must prove that these information sets are critical to task completion.

Some of the displays were necessary for task completion such as the flight planner, MFD and autopilot controls. For display IS4, if the display is being denied, then only the instrument output was removed and the autopilot controls left in place. Display IS1 is critical for control of the platform but also supplies an information set; in this case the display will only be made available at the request of the participant to complete the assigned task after which the participant was again denied access to the display until requested.

The analysis was achieved by repeating the observation experimentation with a new set of way points; this is to eliminate any learned behaviour from the previous phase of experimentation.

The flights below may include a random task, such as object spotting or flight route re-planning, this was determined during the course of these flights by observation of the flights progression by the experimenter. The previous initial flight did include both object spotting and re-planning tasks.

Human interaction:

- Flight 1 (HIF1)

The participant was asked to repeat the process in the previous Observation experiment. Two new waypoints and one dynamic waypoint were supplied.

- Break

Once repeated the participant was allowed a 15 minute break and was then asked to count down from 90 in sets of three to help eliminate short term memory of the experiment. Short term memory can last between 15 to 30 seconds, according to Atkinson and Shiffrin [60], so a 90 second count down was deemed adequate to remove short term memory.

- Flight 2 (HIF2)

The participant was asked to repeat the process in the previous Observation experiment. The waypoints were the same as in HIF1. One random information set was removed for the duration of the flight.

- Break – as previous
- Flight 3 (HIF3)

As HIF2 but with a different information set removed while the previously removed is re-instated.

- Break – as previous
- Flight 3 (HIF4)

As HIF2 but with the final information set removed.

- Debrief

The participant was taken through the task list for HF1 and HF2, which should be identical to each other, and asked to explain any inconsistencies between the task lists.

The participant was then taken through each individual recorded task in HIF2 and asked

- Was the removed information set used during this task? 1..10
- Do you feel that the removed information set impaired task identification? 1..10
- Do you feel that the removed information set impaired task performance? 1..10
- Did you complete the task successfully? y/n
- How effectively did you complete this task in comparison to HF1 (if applicable)? 1..10

Overall Debriefing

On review of the experimental data the configuration points for the head tracker were defined by the participants initial setup and then used to evaluate the data; the derived tasks were then e-mailed to the participant (subject matter expert) to check that he agreed with the researcher’s interpretation of the flights. If the participant had disagreed with the researchers interpretations of the flights and the pilots actions then the researcher would ask for clarification as to why and corrected accordingly; this, fortunately, did not occur.

Conclusion

The case study identified concerns regarding data processing and collection as well as minor hardware issues (such as the need for field of vision restriction) and software replacement with more viable programs (such as X-Plane 9), this then passed to further development within the main experiment and is detailed in the following section.

Rather than supplying any meaning full or justifiable results the case study produced data that could not be validated in terms of statistical analysis but it is still considered any interesting finding and, as such, has been included within this thesis in section 4.1.

3.3 Main Experiment Development

With problems having been identified with the equipment and data acquisition processes in the cast study, this chapter also contains the solutions found to these potential issues

The aims of this experiment were to:

- Collect and compare participant information set (type of visual data available i.e. map, virtual world and platform instrumentation) usage to see if usage differentials can be observed using advanced statistical analysis
- Correlate information set usage to required task performance to identify performance differentials between participants
- Identify any potential information gathering process differences between the participants
- Identify if this research can support the hypothesis that a participant with large amounts of computer games based experience would be a more viable candidate for RPAS training than a participant with no computer games based experience

The development section of this chapter details the modifications made to the case study hardware and software configuration which allows much more accurate data gathering and operation of the head mounted hardware systems; included are the changes of simulation software that occurred and the reasoning behind these changes.

Due to the analysis of the case study data being inefficient it was decided that a software solution had to be created to streamline and speed up the analysis process; this software, which was originally designed by the author as a performance measurement tool, has been adapted to be purely a data capture and singular operator analysis suite. The data output by this software (created in Matlab) are then be incorporated into other statistical analysis tools. This software was dubbed SimPACT (**Sim**ulation **P**rocessing, **A**nalysis and **C**orrelation **T**ools) and its development and function is detailed within section 4.2.

The software is designed within a Graphical Users Interface (GUI) environment and should allow for ease of use as well as much better ordering and storage of data; some of the head tracker data analysis is incorporated as well as a database of tasks which can be configured for performance scoring and independent variables associated with the scoring metrics. These scoring metrics, however, are not utilised for later experimentation but the software solution still retains the ability to output relevant statistical data.

3.3.1 Head Tracker Accuracy

One of the key issues identified during the case study was the inaccuracy of the previous head tracker setup; this issue was identified as being caused by two key problems, the first being a subjective interpretation of head position calibration, this was caused by the experimenter having to individually analyse each line of output head tracking data and manually sync the potential head position to verbal notification of actual head positioning, once a sync point had been identified the time stamps from the flight data, transcript and head tracking data could be aligned. The second problem was due to that of eye movement versus head movement.

The concept of the use of this large visual area was tested during the case study; the four displays available gave a total visual area in excess of 120 degrees laterally and 46.72 vertically (see chapter 3.1 and figure 3.14) with a distance of approximately 80 cm from each screen. Although the data gathered during testing was not statistically viable it did indicate, along with researcher observations, that the full binocular field of vision wasn't fully utilised and head movement was indeed a large supplement to information gathering.

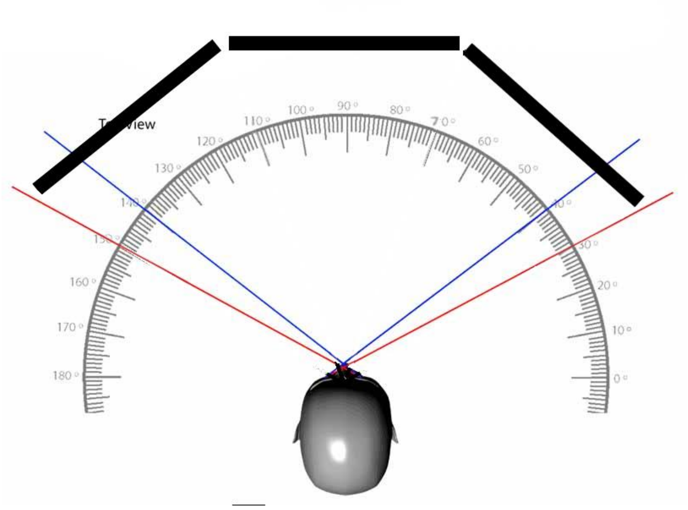


Figure 3.14 - Displays associated with angle and field of view (modified from www.best- 3dtvs.com)

However, the data and observations also indicated that measurement without restricting the binocular field of view would not be viable due to head position not being a completely valid indicator of information set usage. High load tasks were observed to produce much less head movement between adjacent and relevant information sets and this lead to either the head tracker indicating the wrong information set identified as being used or the head position being in a 'void' area (see figure 3.15); the void area represents an area between screens that contains no relevant information and only contains the screens bezels.

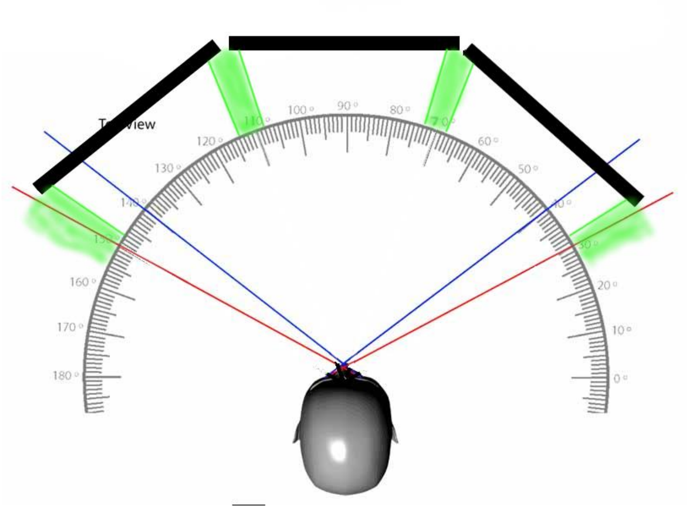


Figure 3.15 - As 3.14 with void areas marked in green

it was decided that effective field of view had to be constricted and a cost effective ‘blinker’ system has been created to restrict field of view to approximately 40 degrees laterally and 30 degrees vertically (see figure 3.16). The lateral restriction approximately equates to the total visible area of a single display; it is still to be tested whether this will allow for increased accuracy results.

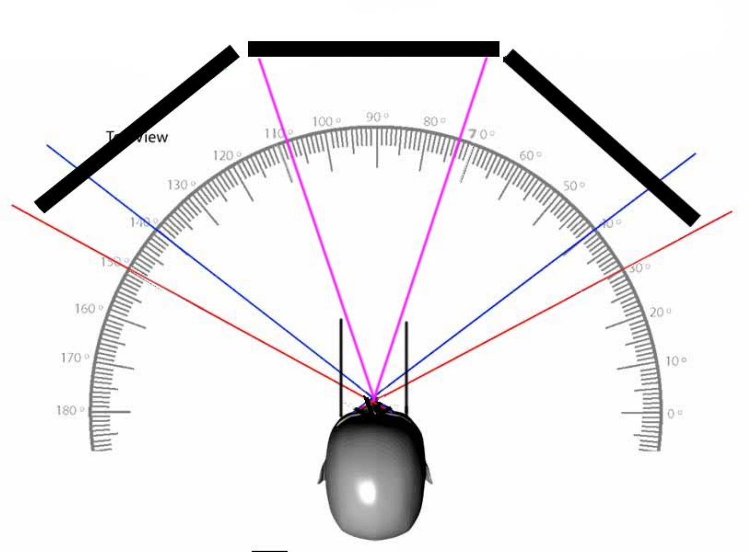


Figure 3.16 - As 3.14 but including field of view restrictors (black lines) and new approximate effective field of view

A concern raised by limiting field of view is a disturbance of normal cognitive behaviour; a head movement to view an object is much more demanding than an eye movement and, when compared to normal operation, this may affect changes in the way in which the equipment is used; it is expected, though, that this concern may not, in fact, be detrimental to the experimentation due to the removal of the 'blinker' device only allowing improvements in speed of information gathering and information set switching. The 'Blinker' may allow the participant to perform better but the participant is still expected to use the information available in a similar way with or without the field of view restriction. As this study does not focus on the behavioural differences from actual RPAS flight, a cognitive behavioural change will not be needed to be taken into account. Any potential participants that may have previously encountered this system in its pilot state will not have their behaviour compared to these previous experiences, it is simply a comparison between two groups with differing skill sets and experiences with no previous definition of behaviour.

Figure 3.17 shows the design of the ‘blinker’ device while figures 3.18 & 3.19 show the positioning and effect of the ‘blinker’ device.

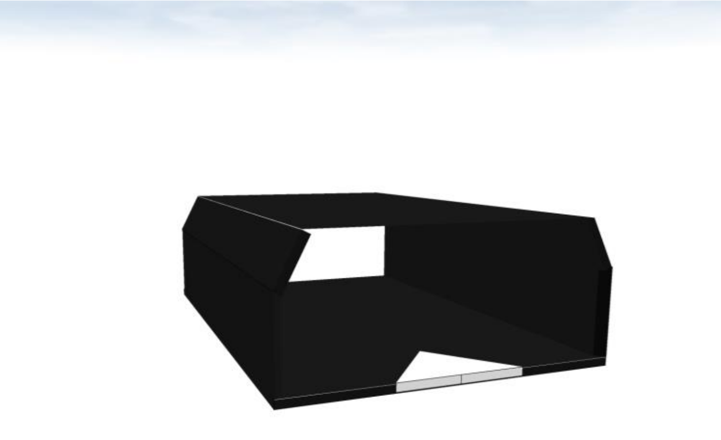


Figure 3.17 - 'Blinker' type eye movement restriction solution

Figure 3.18 - 'Blinker' type eye movement restriction solution with cap for head mounting

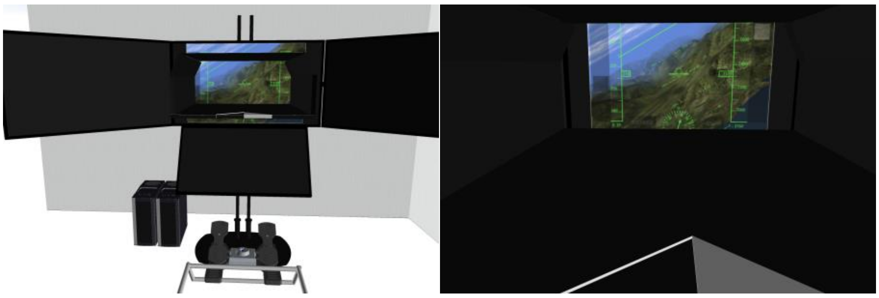


Figure 3.19 - Field of view restriction using blinker type device; at distance then actual view

This restriction effectively limits the pilot’s field of vision to approximately 40 degrees horizontally and 20 degrees vertically, 40 degrees is also approximately the same coverage as that of one of the single 27" displays, this was calculated using basic geometric mathematics with the full experimental rig; the distance from the screens was measured as well as a development participants indication of maximum view range in both x and y axis, trigonometric calculations were carried out on these values and returned the previously stated view ranges.

α = Screen Ratio = 16:9

μ = Screen Size (27 inch)

z = Distance from Screen to Eye (80 cm)

Θ = Viewing Angle (FoV - Horizontal) per Screen

$$\Theta = 2 * \left(arctan\left(\frac{cos(\arctan(\alpha)) * \ \mu * 2.5}{2 * z} \right) \right) = 40.37°$$

β = Viewing Angle (FoV - Horizontal) per Screen

$$\beta = 2 * \left(arctan\left(\frac{sin(\arctan(\alpha)) * \ \mu * 2.5}{2 * z} \right) \right) = 23.37°$$

If the head orientation is positioned between two screens this only allows a maximum of approximately 40% of each screens display being visible which is not an effective working percentage of screen availability for effective screen use. This, therefore, promotes head movement in the acquisition of information from the differing sets.

3.3.2 Data Correlation

With the case study there were three different data sets that required manual correlation using a standard time base, these being pilot transcript, flight data and head tracking data; to solve this issue it was decided to firstly remove the need for transcription and task identification.

Instead of reviewing an entire flight in terms of performance, it was decided to concentrate solely on a single task with very definite start and end points. The participant's performance would then be measured comparatively with other participants from differing backgrounds, in this case experienced gamers versus non-gamers. The decision to omit actual pilots will be discussed further in chapter 4.1.3.

By removing the transcript and using a workload analysis tool supplement, by a questionnaire relating to cognitive ability in one task, removed the need for a transcript directly associated with flight and head tracking data.

The next step was to try and amalgamate the head tracking data and flight data; the solution to this came in the form of a switch of simulation software from Microsoft Flight Sim X to Laminar Research's X-Plane 9, another COTS simulation solution. X-Plane allows for much easier data collection and includes its own data logger as well as network support.

X-Plane allows for UDP (User Datagram Protocol) from the primary simulator to a secondary simulator, effectively replicating the primary flight simulator on a secondary PC also running x-plane; this allowed for data collection to be carried out on the secondary machine. A key data output available was in-cockpit camera position which generated angular in-cockpit head movement data which can be linked to real world head movement, this position would be part of the flight data and be defined by a compass co-ordinate system; using the camera position, though does not give physical head position thus further calculations had to be performed to relate the in-cockpit camera position to the real world head position. This means that the original Optitrack vector program would no longer be necessary for head orientation data collection.

However, this now presented a new problem as the participant is required to operate the in-cockpit camera independently (for object spotting purposes) of their own head position, the in-cockpit camera cannot simultaneously be used to collect head position information and allow the participant to interact with the virtual world; this was solved by only allowing head movement related to camera positioning with the secondary data collection computer. This allowed for the participant to operate the main visual camera independently from his own head movement, in effect creating two separate and independent main visual cameras with the head tracking camera related data being captured by the secondary computer while the primary computer allowed the participant to interact with the virtual world. Figure 3.20 shows the new system data collection architecture.

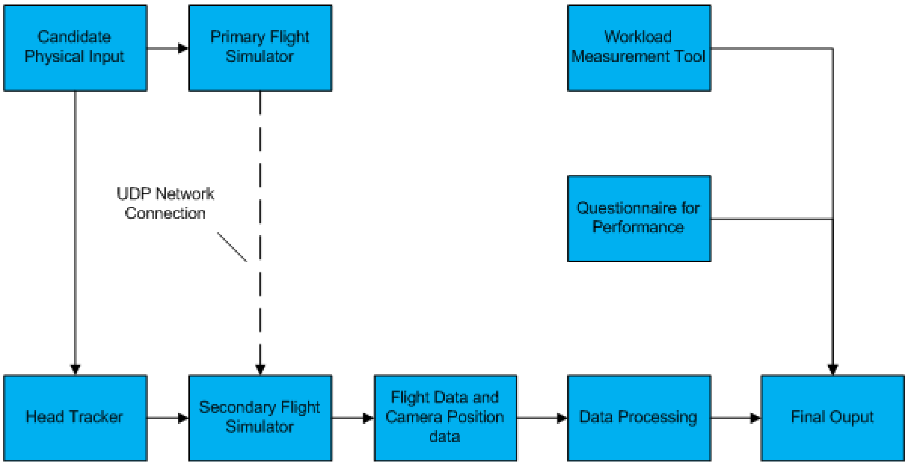


Figure 3.20 - Revised data collection system architecture

X-Plane still supports the use of Google Earth as a participant position orientation tool as well as having developer support which allows the use of this secondary camera and separate screen flight instrumentation availability. The revised systems physical architecture is shown in figure 3.21, it is worthwhile to note that the flight planning display has now been removed as flight planning is no longer required; this display has been replaced with an object identification reference.

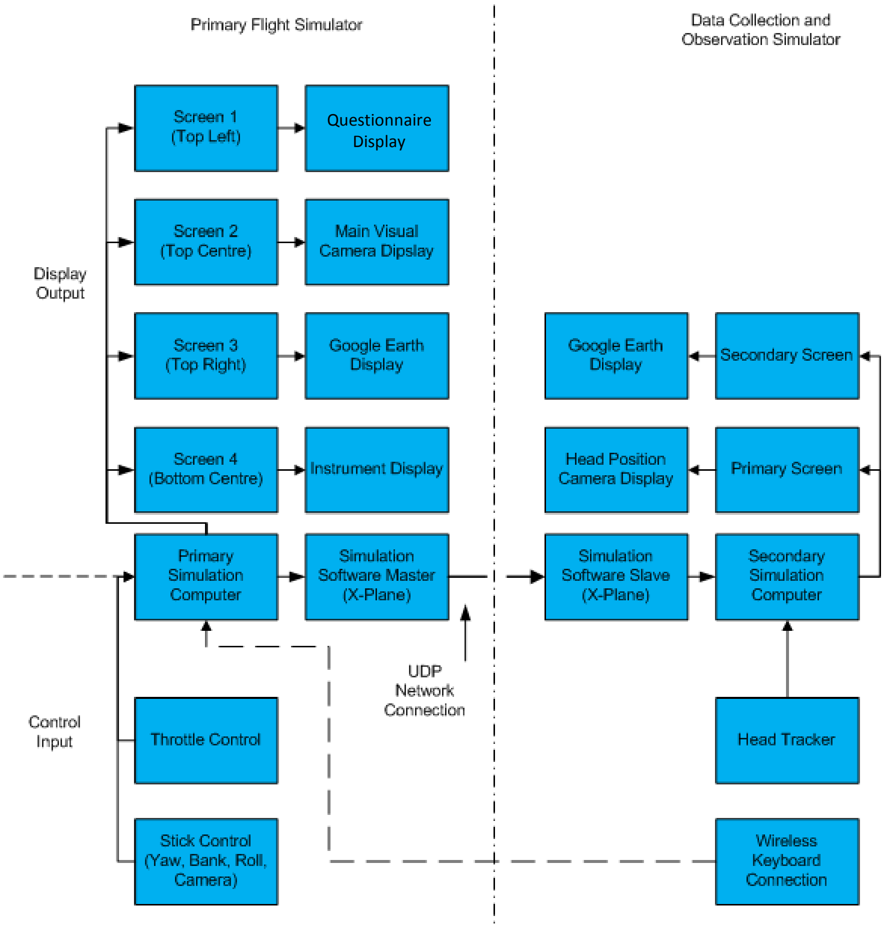


Figure 3.21 - Physical system architecture

The main visual camera position is measured, as already mentioned, using a compass based axis, or 'bearing' so, at this point, the camera position does not relate to the real world head position; the following chapter details the creation of Graphical User Interfaces or GUI's that aid in the logging and processing of data.

3.3.3 Participant Availability

During the case study it was identified that there would be a high risk of obtaining experienced RPAS pilots for further experimentation; this would lead to an extremely small potential data set and a lack of statistically viable output. It was decided that to overcome this, experienced pilots as well as more specifically RPAS pilots, would no longer be used as part of the experimentation. The removal of RPAS pilots as participants means that a direct correlation of this research to actual RPAS pilots, their training and performance levels can no longer be made; this prevented the study from drawing an empirical task-based analysis of professional RPAS operation.

Instead, a comparison between experienced games players versus non-games players would be carried out; this allowed for a much broader participant pool and, therefore, potential data resource that would be statistically more valid. As has already been identified in chapter 2.1.3 some research has already been carried out into RPAS pilot selection with the use of pilots, gamers and non-gamers and this research aims to support these initial findings using a different approach.

3.4 Experimental Design

With the use of the base simulator detailed within this chapter, the insights derived from the case study and the subsequent development work. it was then possible to proceed with an experiment which hopes to answer the objectives stated in chapter 1.2.

This experiment focused on participant ability to manage increasing workload, generated by loss of flight automation, as well as find, identify and remember objects spotted while in flight and under experimental conditions. Effectiveness is measured by analysing the flight data concurrently with a post-flight participant subjective analysis of their own performance related to workload; this is supplemented by analysing correctness of information regarding objects available for spotting during the flight. Further to this the participant's information set usage is analysed to try to identify key differences in simulator operation relating to increased or degraded performance.

Overall these data were used to identify key differences between gamers and non-gamers in terms of their ability to adapt to increasing workload and to try to identify key differences between gamer and non-gamer information set usage. A full specification for the experimental setup can be found in appendix B.

3.4.1 Participants

Participants for this experimentation were required to have the following characteristics, these characteristics were determined using a questionnaire prior to experimentation which was reviewed before the participant is either accepted or declined. An invitation to partake in this experiment was sent by email to those showing interest, this e-mail will contain the initial questionnaire to determine whether the potential participant meets the experimental criteria. This questionnaire, once returned, was evaluated and then an acceptance e-mail sent back (if applicable).

The criteria for selection are as follows:

Age: 18 - 35

Gender: Male

Gaming Experience:

Either -

Experienced gamer: At least three hours of games play per week and over one year's experience (this approximation was based upon [61]

or

Inexperienced games player: Less than one hour of game play per week and less than one years' experience. No experience would be ideal

Criteria for relevant game experience: First Person Shooter (FPS), Simulation (Driving, Flying or other), Role play (either massive multi- player online or single player), Platform, Action (3rd person type)

Disabilities: must have none (due to control interface or movement in/out of the simulator) including colour blindness, mobility impairment (full body), visual impairment beyond the capabilities of the system to allow

Availability: Must be available during weekdays (9am - 5pm) or Saturdays (9am - 5pm) for a maximum of 2 hours

Diction: Must have good verbal and written standard of English due to the need to complete questionnaires

On acceptance as an experimental participant, the participant was presented with an information sheet containing a project brief which defined the purpose of the experiment. This brief included the experimental steps that the participant was required to carry out and also included an expected time frame for experiment completion; if the participant decided to continue with experimentation then he was asked to sign an ethical consent form attached to the e-mail containing the project brief. Also attached to this e-mail was a questionnaire pertaining more directly to work and education background as well as more detailed information relating to their gaming experience (if applicable). All participants remain anonymous.

3.4.2 Questionnaire Design

As mentioned in the previous section Questionnaire A (Appendix E) was designed to identify the suitability and correct grouping of future participants. This questionnaire was based on two key variables, age and gaming experience, with additional information relating to level of education and employment status. This additional information was gathered primarily as an additional set of variables if further, and deeper, analysis was required to identify differences between the Gamer and Non-Gamer groups.

Further information was gathered to eliminate participants that would not be suitable due to testing availability as well as limiting disabilities.

As previously mentioned suitable ages fall between the range of 18 to 35 while gaming experience is defined by at least 3 hours of play per week and one years' experience of gaming for a potential gaming group participant while the non-gaming group required less than an hours play per week and less than a years experience.

The overall design of the questionnaire was intended to allow for quick analysis and selection of potential candidates by presenting questions relating to the key deciders within 13 short questions. It was hoped that participants were more likely to respond to a questionnaire that would not require in-depth responses thereby expanding the potential participant pool with an increased number of responses.

Questionnaire B was designed to gather additional information directly relating to the experimental setup to make sure the participant was comfortable with using the experimental setup; this level of comfort would impact directly on the participants performance during the experimentation so any issues identified by a participant during this questionnaire (reviewed immediately and stored for reference) could be corrected or the experimentation could be aborted if the issue was too severe to proceed.

Questionnaire C, part of the raw data output for experimentation, was designed to gather performance data relating to the objects spotted. The data gathered from this was easily converted to an excel output and added to the statistical analysis.

Questionnaire D was designed to gather additional supporting data if required and to identify if any problems occurred during experimentation. The aim of experimentation was to try to objectively measure participant performance so, due to the nature of the subjective questionnaire D output, was not included in the statistical analysis.

The nature of the questions of questionnaire D can be partitioned into 3 distinct sets:

- Own performance analysis (including physical and mental stress)
- Simulator fidelity and usability
- Experimentation fidelity and usability

Any indication of a performance inhibiting issue, such as motion sickness or visual impairment would have led to the participant and related data being discarded.

3.4.3 Hardware

The simulation hardware, created for experimentation, has not been significantly modified from the configuration seen in section 3.1 with the computer system specs remaining unchanged. What follows is a description of each simulator part and its relevance to experimentation.

Screens

As with the case study three 27" screens and a fourth 22" touch screen, were used during experimentation but have a revised information set display; these information sets are as follows:

- Screen 1 - Questionnaire Screen (not used during flight)

- Screen 2 - Main Visual Display
- Screen 3 - Google Earth map display
- Screen 4 - Flight Instrumentation display

The change to available information sets occurred due to a revision in the format of experimentation in which the participants were no longer required to navigate using an interactive navigation tool and 2D map display as the experiment focused solely on the completion of an object spotting task and a concurrent flight stability task. The three remaining, in-flight, information sets were still be used in a similar fashion to the initial experimentation but with Google Earth now used to locate areas of interest for object spotting within the virtual world; the flight data information set was still used in the same fashion as that of the initial experimentation.

Figure 3.22 shows information sets as they appear on each screen and also shows the final configuration of the simulator for experimentation.

Figure 3.22 - Top left - Questionnaire Screen, Top Middle - Main Flight Visual, Top Right - Google Earth with flight path and rough object location, Bottom - Flight Instrumentation & final configuration

Controls

For primary flight control the Thrustmaster HOTAS Warthog 'stick and throttle' (see figure 3.23) is used with the 'stick' controlling pitch, yaw and roll and the throttle controlling engine power. Further to this the stick's thumb 'hat' control controlled axial movement of the main visual camera and a further switch controlling visual incremental zoom.

Figure 3.23 - Thrustmaster HOTAS Warthog 'stick and throttle' with hat control and zoom control highlighted (http://brain.pan.e-merchant.com)

A rudder control was not necessary as it was deemed that the addition of this type of control would be too problematic for Non-Gamers and would lead to potentially high levels of degraded performance for those with no flight simulator experience while giving those with large amounts of experience a significant advantage. The rudder controls are controlled solely by the simulator software.

Head Tracking

The head tracker used was the TrackIR created by Naturalpoint (see figure 3.24); this has a good degree of accuracy (with accuracy of 100ths of a degree of rotational movement in a 6 axis environment), especially when using the track clip pro feature rather than the original reflective cap mounted arrangement. This hardware is also supported by the flight simulation software. The track IR receiver is mounted centrally above screen 2 and allows for full data regarding head movement beyond the experimental requirements with no loss of data; the track clip itself is mounted on top of the head to prevent interference from the 'blinker' adaptation.

Figure 3.24- TrackIR with track clip pro by Naturalpoint (http://www.vrconcepts.co.nz/)

Computers

This experiment utilises two computers (unlike the case study which only used a single computer), as can be seen in section 3.1, to carry out the simulation and to record the head tacking and flight data. The primary computer is used for the participant's interaction with the simulated flight while the second computer, in this case a mid-range laptop, is used to replicate this flight, collect the flight data and unprocessed head position data and process the flights data output

3.4.4 Software

Flight Simulation Software

Originally, for the case study, Microsoft's Flight Sim X (FSX) was used but this software proved inadequate, as detailed earlier in this chapter and this prompted a switch to Laminar Research's X-Plane 9 which allowed for much better data output and correlation.

This software was then supplemented with satellite generated terrain imagery so that the Google earth display, located on screen 3, could be used as an aid to positional awareness and relation to object spotting areas and flight paths. A free to download Global Atomics X-Plane Reaper model was downloaded to perform as the simulation platform, this model was modified slightly to provide slightly better flight performance and also to remove the HUD (Heads-Up-Display) from the main visual (camera) screen; the removal of the HUD allows for partitioning of different information sets rather than the previous configuration which combined two differing information sets, a comparison can be seen in figure 3.26.

Further to this the simulation software allows for object insertion (see figure 3.25) which is a key requirement of experimentation.

Figure 3.25 - Inserted object in X-Plane 9 with additional photorealistic scenery

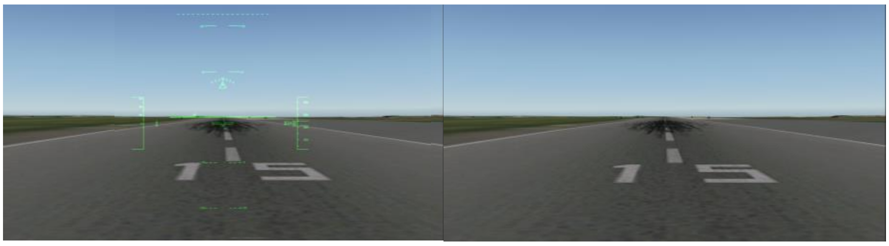


Figure 3.26 - HUD/no-HUD comparison

Flight Simulator Software add-ons

Additional plug-ins had to be utilised for X-Plane 9 to enable experimentation to proceed.

The two primary plugins used are Virtual Camera Plugin created by Barbour [62] and external High-fidelity Simulator Instruments version 2 (XHSI2) which is an open source project [64]. One of the flaws of X- Plane 9, which led to the use of FSX in the case study, was the lack of support for multiple screen outputs when using a single PC system; the simulator is required to output two directly associated software outputs which must allow for direct user interaction, this being the main visuals and flight instrumentation. The use of the XHSI2 plugin and the removal of the need for the participant to physically interact with buttons associated with the flight instrumentation allows for flight instrumentation to be displayed upon a second screen.

The virtual camera plugin is utilised by the second, data logging, computer and allows for camera independence between the primary computer's simulation and the replicated simulation upon the secondary computer thus allowing head tracker independent camera control from the primary simulator.

A third stand-alone application, called Black Box flight logger [64] is a free to use application which allows GPS data (Global Positioning System) to be streamed from the flight simulator and integrated into Google Earth as updateable position vector. This allows for a full flight trace to be displayed from the simulator within Google Earth in either a 2D (Top Down) or 3D display. More importantly for experimentation it gives the participant current position of the platform within the simulated world and allows for positional awareness related to geographical features observed in both the main visual display and the Google Earth display.

Google Earth

Google earth, produced by Google Incorporated, is a globally used satellite imagery tool; its uses are wide ranging with its architecture allowing privately developed programs and apps to interact with it (such as the Black Box application used as part of this experimentation). The Google Earth display allows for either a 2D or 3D view (see figure 3.27) which is fully rotatable; this display can also be static or dynamic thus allowing tracking of an object with constantly updating GPS co-ordinates or visualising an object's progress through the virtual world using a static area. The latter is used for experimentation.

Figure 3.27 - Google Earth 2D view (left) & 3D view (right), red line representing flight path, yellow box representing search area

SimPACT Analysis Software (Matlab)

This development of the SimPACT software is detailed in the 'Findings' chapter. This software was created in response to the lengthily analysis times encountered within the Case Study.

Questionnaires

The questionnaires were accessible via Google Forms, which allow for a much more automated and structured approach to questionnaire output. Google forms allow for the participant to not only manually list answers (where applicable) but also to select from multiple choice and scalar answers (where applicable); the questionnaire responses are compiled into a downloadable excel spreadsheet which were incorporated into the participants main flight data.

3.4.1 Ethics

This experiment followed standard university ethical guidelines, stipulated by the Ethics Approvals (Human Participants) Sub-Committee, within the Code of Practice on Investigations Involving Human Participants (see appendix E).

To this end the ethical implications of this experimentation were considered and evaluated using the Ethical Clearance Checklist a copy of which is included in appendix E); the ethical checklist did not show any requirements for experimentation to be reviewed by an ethical board. A full experimental plan with ethical forms and sign-off sheets were also presented to the Head of Department for further evaluation.

Participant data are treated as stipulated by the Data Protection Act 1988.

A potential participant had the option of withdrawing from the experiment at any time.

3.5 Experimental Procedure

3.5.1 Pre-study

The pre-study was designed to evaluate potential participants so that any participant(s) not meeting the experimental criteria can be identified and removed from the potential participant pool. Questionnaire A, located in appendix E is designed to identify participant acceptability for experimentation and was attached, as a Google Form link, to an initial contact e-mail notifying them of the experiment and asking if they wished to take part. Some of this pre-study data is also examined during the experimental analysis.

Questionnaire A aims to identify the gamer's gender, age and gaming experience; this is performed by asking multiple choice questions and in some cases asking the potential participant to elaborate on his/her experience. Participant availability during the experimental time frame was also determined so that those who are unlikely to be able to participate could be removed from the participant pool.

3.5.2 Experiment

The experimental procedure consisted of three parts; there was a 10 minute break between each of these phases in which the participant is not allowed to interact with the system but may ask pertinent questions regarding their performance:

- Simulator Familiarisation
- Main Experiment
- Post-experiment Questionnaire

Simulator Familiarisation

In this phase of experimentation the participant was asked to perform two repetitive tasks for ten minutes or until the participant felt confident in the tasks' completion.

The first task entailed the participant operating the camera control, under automated flight, and asked to focus and identify multiple objects in separate locations along the pre- defined flight path with the use of the Google Earth display to help locate the targets and the reference guide to identify the targets. At any time the participant could indicate that they felt they had adequate task performance and were asked to move to the second task.

If the participant did not feel they were performing acceptably they were given until the allotted time of ten minutes and then the researcher would terminate the task.

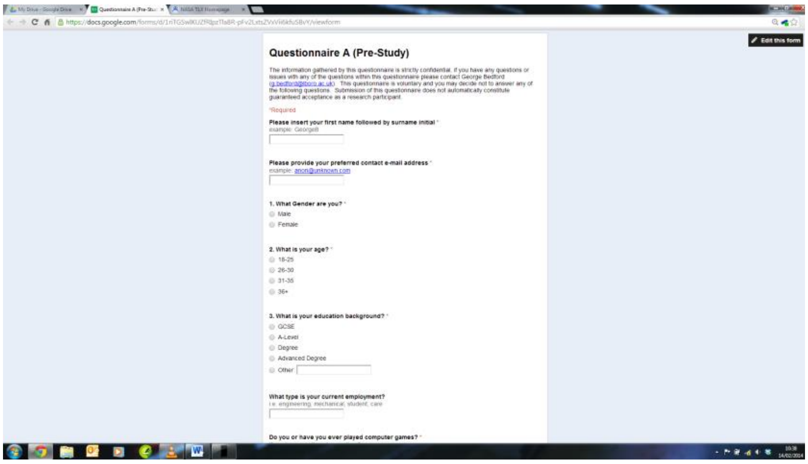
The second task helped to familiarise the participant with the flight control interface and the platform dynamics; this was carried out by asking the participant to manually fly the pre-defined route as in task one. This route did not require any changes in altitude, bearing or speed and was purely to enable fundamental participant control of maintaining consistent flight variables. Similarly to task one the participant was allowed to indicate that he felt confident in his ability to control the platform and that the task could be terminated; the task was also terminated after 10 minutes even if the participant felt that they weren't confident in flight control.

The participant was then asked to fill out questionnaire B to support their decisions and to give feedback as to their interaction with the simulator.

Main Experiment

The participant was then asked to perform a total of nine five minute flights under experimental conditions, the participant was not allowed to interact with the researcher and any questions directed at the researcher were ignored unless they related to a simulator or participant issue.

Each flight consisted of a level and linear flight plan passing close to an area in which multiple objects were located with no less than four objects being present and no more than eight. Figure 3.28, figure 3.29, figure 3.30 and figure 3.31 show the respective information sets displayed on each screen.



Questionnaire A (Pre-Study)

The information gathered by this questionnaire is strictly confidential. If you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions. Submission of this questionnaire does not automatically constitute guaranteed acceptance as a research participant.

*Required

Please insert your first name followed by surname initial *
example: GeorgeB

Please provide your preferred contact e-mail address *
example: anon@unknown.com

1. What Gender are you? *
○ Male
○ Female

2. What is your age? *
○ 18-25
○ 26-30
○ 31-35
○ 36+

3. What is your education background? *
○ GCSE
○ A-Level
○ Degree
○ Advanced Degree
○ Other:

What type is your current employment?
i.e. engineering, mechanical, student, care

Do you or have you ever played computer games? *

Figure 3.28 – Questionnaire Display

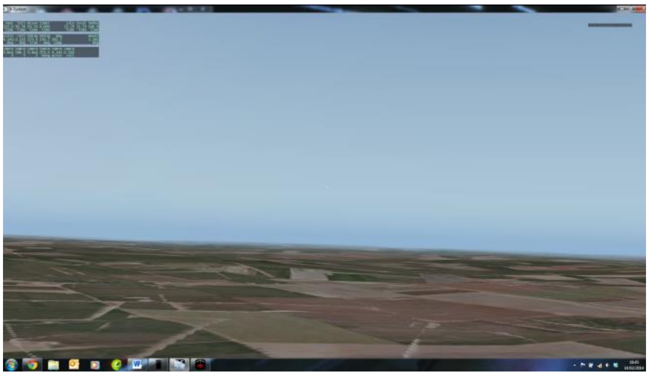


Figure 3.29 - Main visual (camera) Display

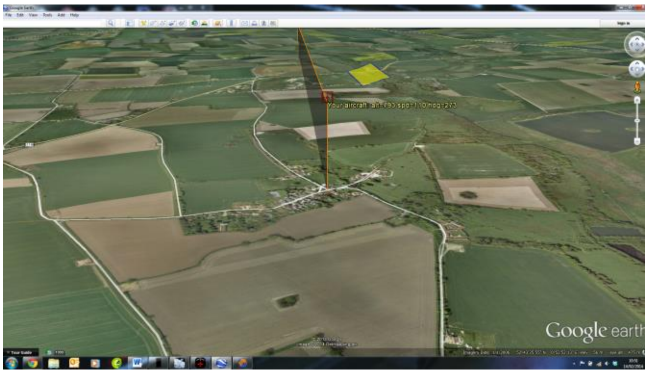


Figure 3.30 - Google Earth Display with object location area

Figure 3.31 - Flight Instrumentation Display

Before the flight began the participant's primary goal was to locate the objects by using the Google Earth display as an aid, orient the camera so that the objects are viewable, identify and remember the number and type of these objects. Questionnaire C was issued post flight to determine the participant’s object analysis along with the NASA TLX workload form.

Before each flight the participant was also asked to calibrate his head position by orienting on the centre of screen 1 and announcing that this orientation has been achieved at which point the researcher saved this calibration data; the participant was then asked to carry out the same process for each of the subsequent screens.

The nine flights consisted of three repetitions depending on degree of workload. The first three flights represent a low degree of workload imposed by the simulator, in this case the flight was fully automated and would not require the participant to perform any flight related interactions other than operate the camera.

The second batch of flights constituted a medium level of workload; in this case the participant had manual control of bank and pitch with the throttle still automated. The participant was asked to fly along the designated flight path and keep their altitude as close to a pre-determined value as possible.

The third batch of flights constituted a high level of workload; in this case the participant had manual control of bank, pitch and throttle and was asked to maintain the flight path and constant altitude and additionally keep the vehicle’s speed to a pre-defined constant as well.

Each of these individual flights was logged and saved within a specific participant folder using the measurement software; the questionnaires were also added to their relevant flight along with the relevant NASA TLX data. The NASA TLX survey incorporates three main steps (which can be seen in appendix E); the participant was asked to input his first name and surname initial (to be later filed and renamed to protect anonymity) and input the flight just undertaken (to be supplied by the investigator), he was then asked to input a scalar interpretation of his Mental Demand, Physical Demand, Effort, Temporal Demand, Performance and Frustration. Finally, he was asked multiple pair-wise questions to increase the accuracy of the workload measurement; the final data output is a single TLX score relating to workload.

There was a break of 5 minutes between each set of flights.

Post experiment questionnaire

On completion of the prescribed number of flights the participant was asked to complete questionnaire D (see appendix E); completion of this questionnaire ended the experimentation. Figure 3.32 shows a graphical representation of the experimental procedure.

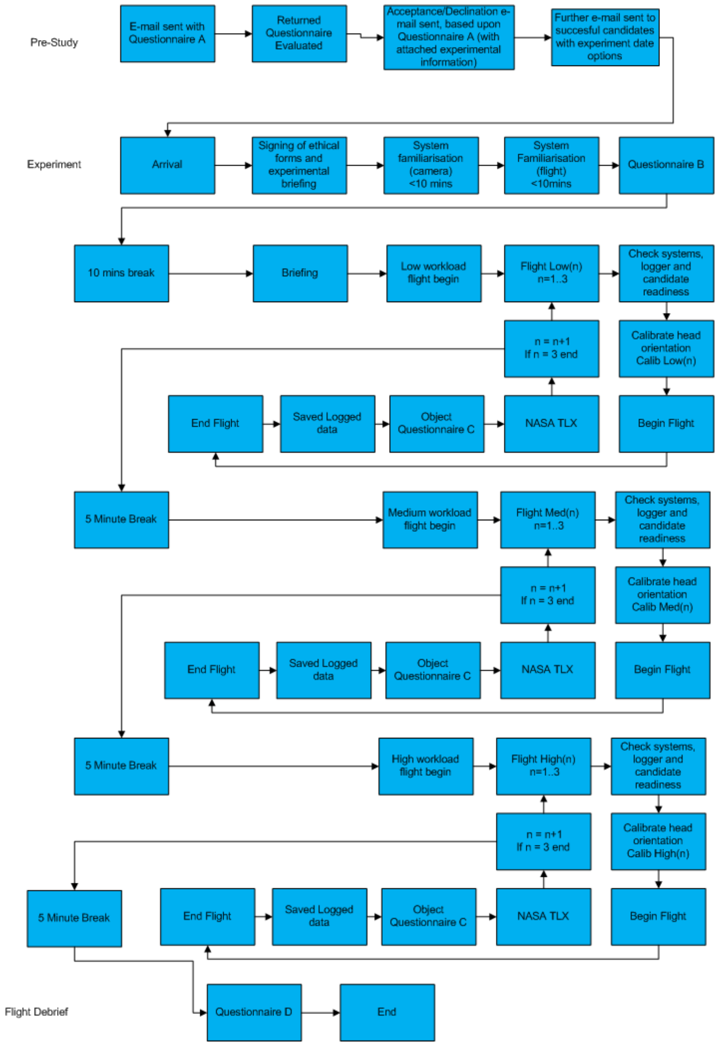


Figure 3.32 - Graphical representation of experimental procedure

3.5.3 Statistical Analysis

Statistical analysis is composed of sections relating directly to flight control workload levels, detailed previously within this chapter; to analyse the data in relation to these levels several methods are used including direct comparison of data as well as comparison of variance, means (information set usage), standard deviation and also includes MANOVA (Multivariate Analysis of Variance), ANOVA (Analysis of Variance) and Spearman's Rank Correlations to identify significance of data and normal distribution test for intergroup uniformity. These tests were chosen over more traditional T-Tests so that the potential for error, when comparing multiple variables, would be reduced; this is due to the analysis requiring multiple T-Tests to be carried out simultaneously to arrive at usable findings. The MANOVA was chosen over multiple T-test's to allow group to group comparison while using multiple variables and it followed that any more detailed comparison would be utilised by the MANOVA's ANOVAs. This method reduces the potential statistical error.

All non-subjective data are collected directly from the flight simulator with an approximate collection rate of a line of data (one entry for each variable per line) every 10th of a second, creating ten outputs of all non-subjective data, per second.

Table 3.2 shows the different type of processed, non-subjective, data output for final statistical analysis; it is to be noted that any calculations relating to the flight data, in terms of variance and standard deviation, are related to a pre-defined mean. This mean represents the desired parameter that the researcher wishes the participant to achieve rather than taking a dynamic mean based solely upon their actual data, for example, the researcher would wish the participant to maintain an altitude of 3000ft therefore this would be a pre-defined mean of 3000. This shows their variance from the desired parameter and, therefore, giving a better representation of their flight performance with the known parameters in mind. The parameters were derived by a need to balance difficulty and realism and the optimums were found through trial and error.

Table 3.2 - Data outputs

	Data Type										
	Flight Data						Screen Used				Workload (sj)
Data output type	Altitude	Vertical Speed	Throttle	Speed	Bank	Time	Screen 1	Screen 2	Screen 3	Screen 4	NASA TLX
Total time	×	×	×	×	×	✓	✓	✓	✓	✓	×
Percentage of time	×	×	×	×	×	×	✓	✓	✓	✓	×
Variance	✓	✓	✓	✓	✓	×	×	×	×	×	×
Standard Deviation	✓	✓	✓	✓	✓	×	×	×	×	×	×
Single Score	×	×	×	×	×	✓	×	×	×	×	✓
Screen Change Count	×	×	×	×	×	×	✓	✓	✓	✓	×
Dwell Time	×	×	×	×	×	×	✓	✓	✓	✓	×
Paired Screen Path Percentage*	×	×	×	×	×	×	✓	✓	✓	✓	×

*Percentage of transition from one screen to another (i.e. Screen 4 to Screen 2) during the flight

The process of analysis for each workload level can be seen below.

Initially a MANOVA test was carried out, over all the available variables, to identify whether, globally, both groups are similar or significantly different; before reviewing the MANOVA results a prior Levene's test for homogeneity of variance was carried out and analysed so that non-homogeneous data can be excluded from the overall MANOVA. Levene's test is used to identify the equality of variance for a single variable between two groups; if there is a lack of equality (less than 0.05) the two data sets would not be suitable for comparison using ANOVA or MANOVA techniques.

The independent ANOVA results from the previous MANOVA are then be reviewed to identify whether there are independent significant differences within the overall MANOVA and the potential significant differences in relation to the available data type (i.e. screen usage, flight control, performance score).

Each data type is then analysed; this includes a type-related MANOVA if available but then moved to descriptive statistics followed by data correlation relating specifically to the same data set. Any differences or significant correlations are identified and the reason for these results explored, cross referencing with other statistical test data if possible. A list of data types can be seen below:

- Screen Data
- Flight Data (not used for low level workload)
- NASA TLX Data
- Object Spotting Data

The individual workload analysis concludes with a cross-data type correlative analysis to try and identify disparate variables potential links to one another; the potential reasoning for these results are also drawn here with adequate explanations where possible.

3.5.4 Data Significance, Sample Sizes and Non-Normality

When looking for significance within the data, the sample size and effect size (f^2) must be determined for MANOVA tests.

MANOVA tests require a certain level of power; the power represents the probability that the test will reject the null hypothesis; the higher the power the more sensitive the test is likely to be.

The effect size represents the strength between the effects predictor and dependant variable; if the large effect size is reached then there is a strong relationship whereas if only the medium effect size is reached there is a tentative relationship. If the medium effect size isn't reached then it is unlikely there is a relationship present.

The effect size is dependent on the power and sample number of the test. The effect size, when using small groups, was calculated by using GPOWER 3.1 software.

For a standard error probability (α) of 0.05 with a Power of 0.95 and a large and medium effect size ($f^2 = 0.35$ & $f^2 = 0.15$) with 2 groups and 6 dependent variables (for example screen usage and dwell times) or DV's the sample size would need to be 68 and 146 respectively. With a potentially limited number of participants and limited experimentation time frame this sample size was not attainable; instead, computing the MANOVA for a 30 sample group with the same error probability and power, the effect size becomes $f^2 = 0.91$.

By reducing the power value to 0.8 the required effect size value for a group of 30 samples becomes $f^2 = 0.59$. Making sure the effect size reaches this level for the MANOVA test is crucial to significance.

When using MANOVA to test for significance between data sets it is necessary to check DV data sets for normal distribution of data; if the data are not normally distributed (Shapiro-Wilk significance < 0.05) then a MANOVA test cannot be performed using that variable. If a DV set is not normal it is, therefore, not be included in the relevant MANOVA test but will instead examined using descriptive statistics and related to the previous MANOVA where possible.

When considering Spearman's correlations (rho or ρ), which are used to identify connections between data sets and to help understand disparate data set relationships, a significant level of correlation must first be obtained; this can be in two forms, a weak correlation (significance < 0.05) or a strong correlation (significance < 0.01). A weak correlation may suggest two variables affecting each other, whereas a strong correlation heavily suggests this effect; the correlation value (ρ) can indicate either a positive or negative correlation. A positive correlation would mean both variables increased/decreased together while a negative correlation would mean that as one variable increased the other would decrease and vice versa. The Spearman's correlation is also non-parametric which means that all data can be analysed, not just the normalised data sets.

Chapter 4 Findings and Analysis

This chapter investigates the results from the main experiment and draws inferences from these results as well as potential new hypotheses and conclusions. It begins with a brief description of the two participating groups.

A short pre-analysis section identifying any problematic variables in terms of normality etc. and the solution relating to that variables analysis is then provided.

Following this, each workload condition is investigated with the analysis format from the previous chapter.

4.1 Case Study Findings

4.1.1 Brief Analysis

This analysis, although intended to be used as an empirical result, was purely subjective (no results or conclusions could be verified with the use of statistics) due to the minimal size of the participant population. The analysis of the data required just for this single participant required an extremely large time frame, with data collation and transcription processes taking almost 1 month for just five experimental flights. The very long analysis time frame arose due to the manual processes (such as transcription, task/phase analysis of transcription and data syncing with task/phase analysis) that had to occur before the flight and head tracking data could be understood in a task/phase context; only a small part of the data collation processed was, in fact, automated (this being the screen/head position processing macro).

This large time frame indicated that a large change to experimental methodology and data collection and interpretation was required. This development is detailed in the following chapter, Main Methodology - Development.

Data output format

The data output and combination was confined to Microsoft Excel. Each flight transcription, once fully transcribed, was analysed for pilot identification and observer inference of task type and task start and stop points, these tasks were further categorised into mission phase elements. The transcription process allowed for an output to an Excel spreadsheet with transcribed text, and identification of the task and mission phase, this can be seen in table 4.1.

Table 4.1 - Example task transcription

Start Time	End Time	Transcribed Vocalisation	Task	Mission Phase
1.517	1.944	Cool		
1.944	114.084		HT from 18	
114.084	115.523	no, it's not BG is it		
115.523	132.676			
132.676	135.524	there we go, cool		
135.524	165.576			
165.576	167.176	Keyworth, am I being dim?	Pre-Flight Planning	Pre-Flight Planning
167.176	174.719			
174.719	175.402	got it		
175.402	176.875	there we go		
176.875	177.903	******* miles away		
177.903	180.909	(instructor)		
180.909	182.313	that’s alright, we'll go there		
182.313	188.651	(instructor)		
188.651	196.365			
196.365	197.538	ok, that’s saved		
197.538	200.276	(instructor)		
200.276	210.996			
210.996	212.09	ok, pre-flight		
212.09	214.085			
214.085	215.258	we're gonna go #nav#		
215.258	227.511			
227.511	228.162	#nav#		
228.162	228.81	#alt#	Configuration	Systems Check
228.81	237.33			
237.33	238.661	and #GPS#		
238.661	240.561			
240.561	241.765	right, ready for flight		
241.765	245.116			

Once completed the three data sets were sync'd; this was done for screen to transcript, by using a known screen usage pattern and aligning the screen patterns time stamp to that of the relevant transcript point (where this pattern was identified within the transcript). Flight data were sync'd to the transcript by noting the transcript time stamp for the autopilot engaged 'beep' and aligning the transcript time stamp to the flight data notification of autopilot engagement. The data were then manually sorted into a task based structure; the flight data and head tracking datas were then manually processed and assigned to that task, an example of this partitioned data can be seen in table 4.2 where, on the left hand side, partitioned screen usage data (displaying time and percentage) is tabulated and, on the right, the partitioned flight data (displaying the mean, variance, mean per second, variance per second, standard deviation and polynomial regression scores). For many of the tasks the simulator was being flown automatically so the flight data associated with these phases of flight was automatically disregarded as it is a test of the systems control and not the pilot's performance. This processed datas were then averaged, so that individual flights tasks could be compared to a global result, for similar tasks.

Table 4.2 – Task screen data output (left) example including flight data (right)

Take Off

HT

	Time (s)	%
Screen 1	0	0
Screen 2	17.34	0.369502
Screen 3	0	0
Screen 4	0	0
Void/Fluc	29.588	0.630498
Total	46.928	

	Throttle (%)	Alt (m)	V.Speed (m/s)	Bank (°)	Speed (kts)
MEAN	90.20244	953.0859	1684.963063	10.88681	86.63250736
VAR	465.0104	196839.7	1555588.53	310.9414	1161.981497
MEANps	1.920018	20.28706	35.86553987	0.231733	1.844029531
VARps	9.909018	4194.505	33148.40883	6.625925	24.76094224
StdDev	21.5641	443.6663	1247.232348	17.63353	34.0878497
Poly Reg	**0.8943**	0.9955	0.9086	**0.6511**	0.9678
order	4	4	4	2	3

Task related head tracking comparison

What follows in this section is a brief comparison of the averaged tasks associated with each flight; below is a combined table (table 4.3) of the task averages for each flight. This table omits data for flight 4 as the head tracking data proved impossible to process due to the tracker becoming uncalibrated during experimentation.

It was often observed by the experimenter that screen 1 and screen 4 proved problematic in exact screen identification; at the time of experimentation screen 1 was situated above screen 4 due to equipment allowing the mounting of screen 4 centrally being unavailable.

It was observed by the experimenter (by observing the participants eye movement closely) that the participant would often, when prioritising both of these information sets, site his head in vertically central position between these two screens and use eye movement to observe both information sets with the minimum of head movement, this creates the potential for a large degree of error in the use of screen 1 and 4.

Head tracker task comparison

All percentages are in decimal format

					Sc 1 Off			SC 4 Off			Sc 2 Off	
		Flight 1			Flight 2			Flight 3			Flight 5	
Task	Screen	Time	Percentage	Screen	Time	Percentage	Screen	Time	Percentage	Screen	Time	Percentage
Pre-flight Planning	Screen 1	24.79	0.47843289	Screen 1	102.69	0.568051	Screen 1	142.26	0.737114	Screen 1	63.69	0.744607
	Screen 2	17.13	0.33059925	Screen 2	14.05	0.07772	Screen 2	19.37	0.100365	Screen 2	10.68	0.124861
	Screen 3	0	0	Screen 3	0	0	Screen 3	6.6	0.034198	Screen 3	1.62	0.01894
	Screen 4	4.3	0.08298755	Screen 4	14.13	0.078163	Screen 4	9.22	0.047773	Screen 4	1.74	0.020343
	Void/Fluc	5.595	0.10798031	Void/Fluc	49.906	0.276065	Void/Fluc	15.546	0.080551	Void/Fluc	7.805	0.091249
	Total	51.815		Total	180.776		Total	192.996		Total	85.535	
		Time			Time			Time			Time	
Configuration	Screen 1	0	0	Screen 1	4.32	0.24393	Screen 1	2.78	0.081477	Screen 1	3.2	0.127964
	Screen 2	0	0	Screen 2	0	0	Screen 2	1.93	0.056565	Screen 2	1.12	0.044787
	Screen 3	0	0	Screen 3	0	0	Screen 3	0	0	Screen 3	0	0
	Screen 4	23.14	0.72548282	Screen 4	8.93	0.504235	Screen 4	19.62	0.575029	Screen 4	5.11	0.204343
	Void/Fluc	8.756	0.27451718	Void/Fluc	4.46	0.251835	Void/Fluc	9.79	0.286928	Void/Fluc	15.577	0.622906
	Total	31.896		Total	17.71		Total	34.12		Total	25.007	
		Time						Time			Time	

Take Off	Screen 1	0	0	Screen 1	2.79	0.092253	Screen 1	0	0	Screen 1	22.19	0.515891
	Screen 2	20.14	0.50104488	Screen 2	16.81	0.555831	Screen 2	17.34	0.369502	Screen 2	8.44	0.19622
	Screen 3	0	0	Screen 3	0	0	Screen 3	0	0	Screen 3	1.13	0.026271
	Screen 4	9.57	0.23808339	Screen 4	0	0	Screen 4	0	0	Screen 4	10.97	0.255039
	Void/Fluc	10.486	0.26087173	Void/Fluc	10.643	0.351916	Void/Fluc	29.588	0.630498	Void/Fluc	0.283	0.006579
	Total	40.196		Total	30.243		Total	46.928		Total	43.013	
Autopilot Management		Time			Time			Time				
	Screen 1	0	0	Screen 1	0	0	Screen 1	0	0		6.47	0.399951
	Screen 2	0	0	Screen 2	0	0	Screen 2	0.5	0.067376		0	0
	Screen 3	0	0	Screen 3	0	0	Screen 3	0	0		0	0
	Screen 4	2.58	0.95768374	Screen 4	1.65	0.993976	Screen 4	2.86	0.385393		2.91	0.179885
	Void/Fluc	0.114	0.04231626	Void/Fluc	0.01	0.006024	Void/Fluc	4.061	0.547231		6.797	0.420164
	Total	2.694		Total	1.66		Total	7.421			16.177	
Take Off Monitor		Time			Time			Time			Time	
	Screen 1	0	0	Screen 1	16.08	0.219714	Screen 1	16.64	0.272104	Screen 1	56.54	0.794257
	Screen 2	6	0.29756001	Screen 2	13.93	0.190337	Screen 2	6.68	0.109234	Screen 2	2.06	0.028938
	Screen 3	0	0	Screen 3	21.13	0.288716	Screen 3	12.34	0.201789	Screen 3	1.69	0.023741
	Screen 4	10.74	0.53263241	Screen 4	1.98	0.027054	Screen 4	4.06	0.066391	Screen 4	0	0
	Void/Fluc	3.424	0.16980758	Void/Fluc	20.066	0.274178	Void/Fluc	21.433	0.350482	Void/Fluc	10.896	0.153064
	Total	20.164		Total	73.186		Total	61.153		Total	71.186	

											Time	
Monitor AVG	Screen 1	182.02	0.3976667	Screen 1	17.12	0.140297		192.94	0.415537	Screen 1	457.92	0.740179
	Screen 2	89.7	0.19597134	Screen 2	20.85	0.170864		86.33	0.18593	Screen 2	0	0
	Screen 3	36.33	0.07937167	Screen 3	54.91	0.449982		99.98	0.215328	Screen 3	122.82	0.198526
	Screen 4	17.65	0.03856069	Screen 4	0	0		2.66	0.005729	Screen 4	11.428	0.018472
	Void/Fluc	132.02	0.28842961	Void/Fluc	29.147	0.238857		82.405	0.177476	Void/Fluc	26.493	0.042823
	Total	457.72		Total	122.027			464.315		Total	618.661	
Positional Awareness	Screen 1	85.39	0.30507215	Screen 1	7.84	0.031318	Screen 1	14.51	0.298143	Screen 1	-	-
	Screen 2	53.71	0.19188928	Screen 2	85.4	0.341143	Screen 2	16.79	0.344991	Screen 2	-	-
	Screen 3	58.97	0.21068163	Screen 3	113.25	0.452394	Screen 3	11.79	0.242254	Screen 3	-	-
	Screen 4	7.58	0.027081	Screen 4	0	0	Screen 4	0	0	Screen 4	-	-
	Void/Fluc	74.251	0.26527594	Void/Fluc	43.845	0.175145	Void/Fluc	5.578	0.114613	Void/Fluc	-	-
	Total	279.901		Total	250.335		Total	48.668		Total	-	
Object Recognition	Screen 1	11.13	0.12161542	Screen 1	0	0	Screen 1	-	-	Screen 1	-	-
	Screen 2	53.58	0.5854586	Screen 2	3.46	0.706122	Screen 2	-	-	Screen 2	-	-
	Screen 3	2.8	0.03059507	Screen 3	1.4	0.285714	Screen 3	-	-	Screen 3	-	-
	Screen 4	0	0	Screen 4	0	0	Screen 4	-	-	Screen 4	-	-

	Void/Fluc	24.008	0.26233091	Void/Fluc	0.04	0.008163	Void/Fluc	-	-	Void/Fluc	-	-
	Total	91.518		Total	4.9		Total	-		Total	-	
Re-Plan		Time										
	Screen 1	15.49	0.39052061	Screen 1	-	-	Screen 1	-	-	Screen 1	-	-
	Screen 2	12.81	0.32295475	Screen 2	-	-	Screen 2	-	-	Screen 2	-	-
	Screen 3	0	0	Screen 3	-	-	Screen 3	-	-	Screen 3	-	-
	Screen 4	0	0	Screen 4	-	-	Screen 4	-	-	Screen 4	-	-
	Void/Fluc	11.365	0.28652464	Void/Fluc	-	-	Void/Fluc	-	-	Void/Fluc	-	-
	Total	39.665		Total	-	-	Total	-	-	Total	-	
Re-Task												
	Screen 1	21.6	0.23879277	Screen 1	-	-	Screen 1	-	-	Screen 1	-	-
	Screen 2	14.89	0.16461224	Screen 2	-	-	Screen 2	-	-	Screen 2	-	-
	Screen 3	1.05	0.01160798	Screen 3	-	-	Screen 3	-	-	Screen 3	-	-
	Screen 4	23.6	0.26090321	Screen 4	-	-	Screen 4	-	-	Screen 4	-	-
	Void/Fluc	29.315	0.3240838	Void/Fluc	-	-	Void/Fluc	-	-	Void/Fluc	-	-
	Total	90.455		Total	-	-	Total	-	-	Total	-	
Waypoint Management		Time										
	Screen 1	36.39	0.63874603	Screen 1	-	-	Screen 1	-	-	Screen 1	-	-
	Screen 2	0	0	Screen 2	-	-	Screen 2	-	-	Screen 2	-	-
	Screen 3	0	0	Screen 3	-	-	Screen 3	-	-	Screen 3	-	-

	Screen 4	7	0.12286953	Screen 4	-	-	Screen 4	-	-	Screen 4	-	-
	Void/Fluc	13.581	0.23838444	Void/Fluc	-	-	Void/Fluc	-	-	Void/Fluc	-	-
	Total	56.971		Total	-	-	Total	-	-	Total	-	
Expedite		Time									Time	
	Screen 1	10.29	0.53123387	Screen 1	-	-	Screen 1	-	-	Screen 1	7.36	0.4362
	Screen 2	3.02	0.1559112	Screen 2	-	-	Screen 2	-	-	Screen 2	0	0
	Screen 3	0	0	Screen 3	-	-	Screen 3	-	-	Screen 3	4.4	0.260772
	Screen 4	1.16	0.05988642	Screen 4	-	-	Screen 4	-	-	Screen 4	1.23	0.072898
	Void/Fluc	4.9	0.25296851	Void/Fluc	-	-	Void/Fluc	-	-	Void/Fluc	3.883	0.230131
	Total	19.37		Total	-	-	Total	-	-	Total	16.873	
Autopilot Management		Time										
	Screen 1	1.78	0.14586577	Screen 1	0	0	Screen 1	-	-	Screen 1	-	-
	Screen 2	4.05	0.3318856	Screen 2	0	0	Screen 2	-	-	Screen 2	-	-
	Screen 3	0	0	Screen 3	3.02	0.33634	Screen 3	-	-	Screen 3	-	-
	Screen 4	2.58	0.21142342	Screen 4	1.65	0.183762	Screen 4	-	-	Screen 4	-	-
	Void/Fluc	3.793	0.31082521	Void/Fluc	4.309	0.479898	Void/Fluc	-	-	Void/Fluc	-	-
	Total	12.203		Total	8.979		Total	-	-	Total	-	
		Time						Time			Time	
	Screen 1	14.05	0.16142375	Screen 1	3.195	0.069325	Screen 1	2.97	0.026357	Screen 1	172.26	0.617676
	Screen 2	37.3	0.42854845	Screen 2	6.81	0.147764	Screen 2	92.74	0.823002	Screen 2	0	0

Approach Control	Screen 3	4.46	0.05124199	Screen 3	7.77	0.168594	Screen 3	1.95	0.017305	Screen 3	80.28	0.287862
	Screen 4	0	0	Screen 4	4.64	0.100679	Screen 4	1.07	0.009495	Screen 4	7.45	0.026714
	Void/Fluc	31.228	0.35878582	Void/Fluc	23.672	0.513637	Void/Fluc	13.955	0.123841	Void/Fluc	18.894	0.067749
	Total	87.038		Total	46.087		Total	112.685		Total	278.884	
		Time						Time			Time	
	Screen 1	9.41	0.07213381	Screen 1	25.4	0.155378	Screen 1	10.68	0.124166	Screen 1	50.03	0.636126
	Screen 2	102.52	0.78588293	Screen 2	100.72	0.61613	Screen 2	73.98	0.860093	Screen 2	0	0
Landing Control	Screen 3	2.06	0.01579125	Screen 3	10.93	0.066862	Screen 3	0	0	Screen 3	25.94	0.329824
(Sc 1=Sc 4)	Screen 4	0	0	Screen 4	1.42	0.008687	Screen 4	0	0	Screen 4	0	0
	Void/Fluc	16.462	0.12619201	Void/Fluc	25.002	0.152944	Void/Fluc	1.354	0.015742	Void/Fluc	2.678	0.03405
	Total	130.452		Total	163.472		Total	86.014		Total	78.648	

Table 4.3 - Flight task average comparison, it can be seen above the flight reference an indicator of which screen had been removed for each flight. A dash in a data box indicates that a task was not performed or observed during that flight

It is only relevant at this point to compare tasks from specific flights where the tasks appear or can be inferred to not be directly affected by the loss of a screen. Due to the long list of tasks and the lack of repeated data for many only a few tasks are examined to see if there is potentially relevant interaction between the head data and the task type.

Expected Screen usage tasks

Some task that were performed were expected to have a certain amount of screen usage, an example of this would be flight planning in which screen 1 was predominantly used to create the flight plan with a small usage of screen 2 by which the pilot would upload the flight plan into the simulator. This sort of pre- disposed behaviour was also expected with tasks such as re-planning, landing, approach control and configuration.

Where data can be compared to multiple task sets it can be seen that a potential expected pattern of screen usage does, in fact, occur. Using the Landing Control task as an example it can be seen that the screen 2 (outside visual) usage is extremely high with the first 3 flights and this is to be expected as the pilot was using external visual cues while attempting landing. The screen 1 (flight planner/map) use is reasonably low but also still maintains a similar amount of usage indicating that there may well be a certain pattern to an experienced pilot's information set usage; to further support this, with the loss of external visuals, it can be seen that the pilot's information set usage changes to try and compensate for the external visual information set loss with a much higher use of the map display as well as a much higher use of the Google Earth display.

Unknown Usage tasks

These tasks represent phases of the flight in which there is no expectation of information set usage; a perfect example of this type of task is the monitor task in which the pilot has no apparent goal other than maintaining awareness of the simulators current state.

Although the non-aggregated monitor task showed a large amount of fluctuation in terms of their individual information set usage it can be seen from the averaged task data (all monitor task during flight combined and averaged using a weighting based upon amount of time spent on each task) that there is some potential for a generic amount of information set usage. For example, the external visual screen usage during the first three flights appears to not fluctuate by more than 2.5% indicating that there could be a pattern pertaining to that information sets usage; the most striking potential inference is of the pilot’s trust in the automated system to effectively fly the aircraft. The information set usage during the monitor task shows a maximum usage of 3.8% and a minimum of 0%; although this is possibly interesting, this percentage cannot be wholly trusted due to the potential of the head trackers misidentification of screen use due to the use of eye movement which impacted upon the degree of head movement required to view a certain information set.

Tasks and void data

As previously mentioned within the problem identification section the head trackers accuracy in terms of screen identification was extremely poor and the output data not to be wholly trusted; amounts of void areas recorded ranged from a very acceptable 0.6% to a completely unacceptable 63%. With an average void (not viewing any information) percentage over all the tasks of 22.7% it is not possible to categorically state that any of the potential, previously mentioned, interpretations are viable as these void areas could represent an information set usage that has not been logged due to the pilots head position being in a void location and would, therefore, skew any potential interpretations drawn from the data showing information set usage. This large amount of void appearance is likely due to the pilot’s usage of eye movement as opposed to head movement, with the pilot often viewing a data set with a large degree of eye deflection as opposed to head deflection leaving the head in a void area. This problem was rectified in later experimentation by limiting the pilot’s use of eye movement with a ‘blinker’ type device.

Flight performance

Although again not statistically viable due to the small experimental participant pool, the flight data did indicate certain potential characteristics relating to experienced pilot performance. As already stated the majority of the flight did not require the operator to have any direct input over the flight control of the simulator but during the phases where direct pilot input was required there appeared to be certain characteristics relating to adequate platform control. The table (table 4.4) below shows the pilot’s flight data for several variables during take-off; below that, in figures 4.1 & 4.2 the trace of the altitude and vertical speed can also be seen.

	Throttle	Alt	V.Speed	Bank	Speed
MEAN	90.20244	953.0859	1684.963063	10.88681	86.63250736
VAR	465.0104	196839.7	1555588.53	310.9414	1161.981497
MEANps	1.920018	20.28706	35.86553987	0.231733	1.844029531
VARps	9.909018	4194.505	33148.40883	6.625925	24.76094224
StdDev	21.5641	443.6663	1247.232348	17.63353	34.0878497
Poly Reg	0.8943	0.9955	0.9086	0.6511	0.9678
order	4	4	4	2	3

Table 4.4 - Experimental participant flight data for take-off, flight 3

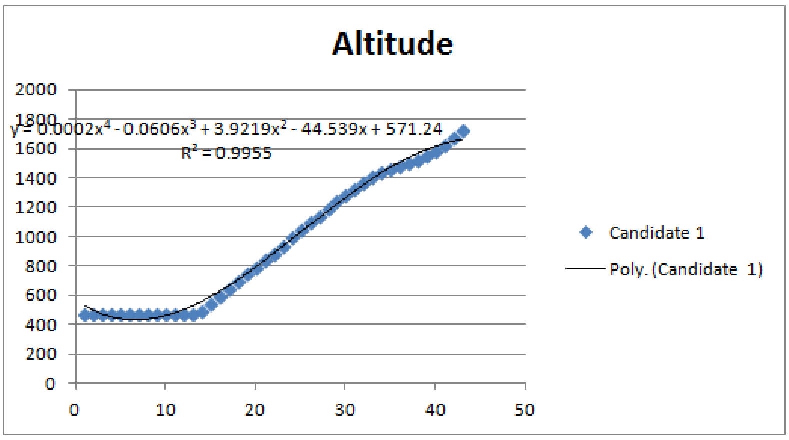


Figure 4.1 - Experimental participant altitude trace for take-off, flight 3 (time vs. feet) including polynomial regression root mean squared score

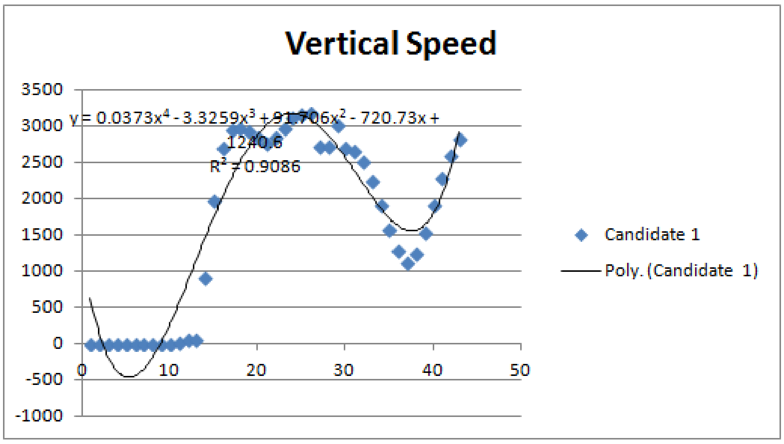


Figure 4.2 - Experimental participant vertical speed trace for take-off, flight 3 (time vs. feet per second) including polynomial regression root mean squared score

This task was identified as being completed when the autopilot was engaged at approximately 1800 feet. The root mean squared scores for these data show a high degree of curve fitting correlation, what this may indicate relates not to set optimums for indicators such as rate of climb or speed but to the pilot operating the platform smoothly and putting as little stress upon the platform as possible. Although this may be the case, limits to some of the indicators would still have to be applied, for example, an extremely high climb rate would potentially lead to platform stress as well as stalling both of which could lead to the destruction of the platform.

By comparison a test participant's data (novice simulator operator) showed a much less smooth operation of the platform (see figure 4.3 & 4.4) with much lower scores for polynomial regression occurring in the vertical speed trace, this could potentially justify a test for smoothness as a performance measure.

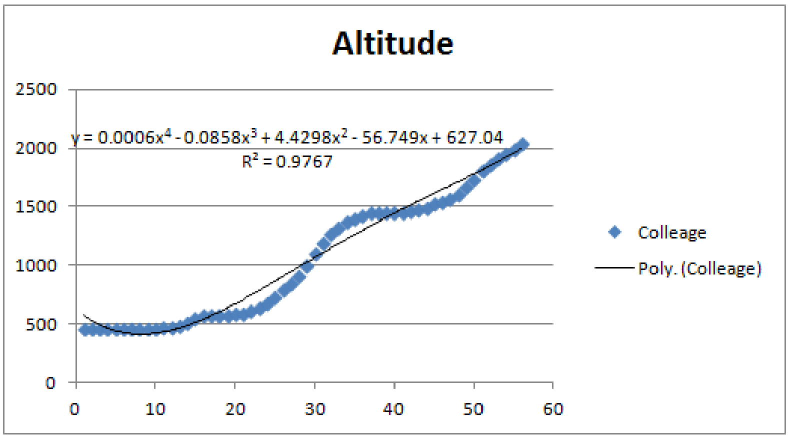


Figure 4.3 - Test participant altitude trace for take-off, (time vs. feet per second) including polynomial regression root mean squared score

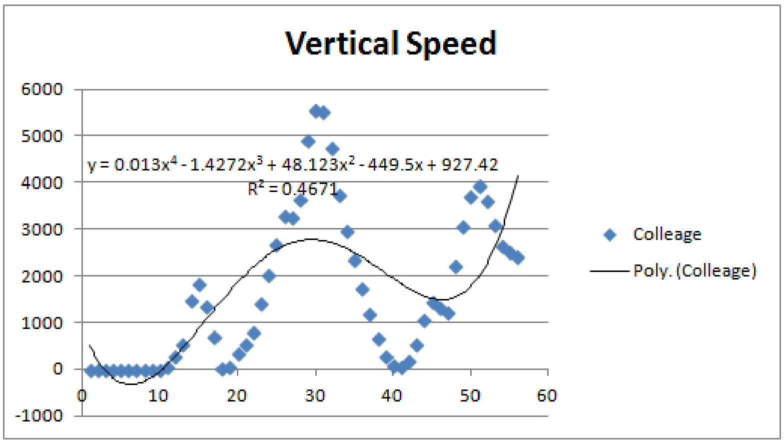


Figure 4.4 - Test participant vertical speed trace for take-off, (time vs. feet per second) including polynomial regression root mean squared score

4.1.2 Conclusion

Although interesting as a development experiment the case study did not lead to any empirically viable results due to the issues relating to head tracking accuracy as well as experimental participant availability. However, as a case study to determine issues that could possibly be present during the main experimentation it worked extremely well, identifying the head tracking faults as well as identifying potential participant pool issues and performance data collection and correlation.

The rectification of these issues will be dealt with in chapter 4.1.

4.1.3 Problem Identification

Head tracking

During the course of experimentation several issues arose concerning head tracking calibration, operation and data collection; this included the application of head tracking data to the transcript and flight data. It was found very difficult to identify exact time match points between the transcript and head tracking data, originally the head movement pattern would be used for calibration and identification of the correlation of the head tracking data and the transcript, but this proved to be ineffective and lead to large amounts of inferred data and the use of their potential reference points between the two data sets.

Another issue, which did not help the collation of data, was the large degree of observed error with the use of the head tracker; it was often seen, even after calibration, that there would be a large 'void' aspect to the head tracking measurements which indicated the head position to not be in a relevant calibrated screen partition. Eye movement without restriction can account for a large amount of this issue and it was observed, especially during high workload intensity tasks that the pilot would focus on the relevant data sets and position his head in such a way that allowed minimal head movement between data sets and maximal eye movement. Understandably this skewed results but also showed the pilots reliance on eye movement rather than head movement; unfortunately, the optimal point for minimal head movement between data sets occurred within the void areas between screens which accounts for the large amount of "void" areas used during experimentation. Another upshot of this minimal head movement strategy is that the void areas had to be redefined as not being exactly at the same angles demonstrated in figure 3.15; the pilot often supplemented head movement with eye movement to look at the outer screens and this lead to the pilot, even when looking at screen 1 (for example) being shown as looking at screen 2. The most seemingly often occurrence of this would be the use of screen 1 and screen 4 and occurred across the horizontal void area.

Overall this lead to the head tracking data not being statistically viable as a measurement but the output will still be considered as a potential indication of information set usages.

Data Collection

Multiple recording solutions were being used concurrently for this experiment, these being the head tacking output, VivendoByte flight data logger and audio recording software. In some cases the flight logger and head tracking software failed to record relevant data, this was only discovered after experimentation. It was realised a better solution to combined data logging had to be produced, this will be explored within the following chapter.

Transcription

Although no error can be associated with the transcription, the use of transcribed flights and their application to the empirical data, led to large data combination and evaluation times scales, with a 20 minute flight taking several hours to transcribe and analyse fully before the data could be applied. Future experimentation would require much faster analysis times which would not be allowed by flight transcription.

4.2 SimPACT Software

The SimPACT software suite (**Sim**ulation **P**rocessing, **A**nalysis and **C**orrelation **T**ools) was created solely by the author to address the issue of lengthy data analysis time frames, as identified by the case study. This set of tools are designed to:

a) Capture and store flight and head tracking data from the simulator with respect to the flight number and participant identification
b) Analyse and process the flight and head tracking data to give usable outputs that can then be further analysed using a statistical software suite

Matlab was chosen for the creation of this software as the suite incorporated pre-defined mathematical functions as well as an intuitive Graphical User Interface (GUI) creation interface.

Matlab is a mathematically based engineering and programming software with a myriad of functions relating to engineering, statistics and mathematics:

> "MATLAB® is a high-level language and interactive environment for numerical computation, visualization, and programming. Using MATLAB, you can analyse data, develop algorithms, and create models and applications. The language, tools, and built-in math functions enable you to explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages, such as C/C++ or Java™."[65]

To aid with ease of data collection and processing it was decided to use Graphical User Interfaces; three of these interfaces were constructed by the author to aid data collection and performance measurement but with a revision of objectives it became only necessary for the use of two of these interfaces.

This section only covers the structural function of the software and the design of the GUI's; the programming, rather than being described and documented within the thesis, can be found within the attached disc. It must be noted that this software was designed for a specific experimental format and would have to be modified for use with another simulator or host computer by altering the file paths within both ‘Collection’ and ‘Processing’ GUI ‘.m’ files.

4.2.1 Data Collection Interface

Figure 4.5 below is the GUI for data collection, this GUI, and supporting data collection programming, was designed to aid the collection and storage of data from the flight simulator by exporting the raw data directly from the simulator folder and storing it in a user and flight specific folder for later processing.

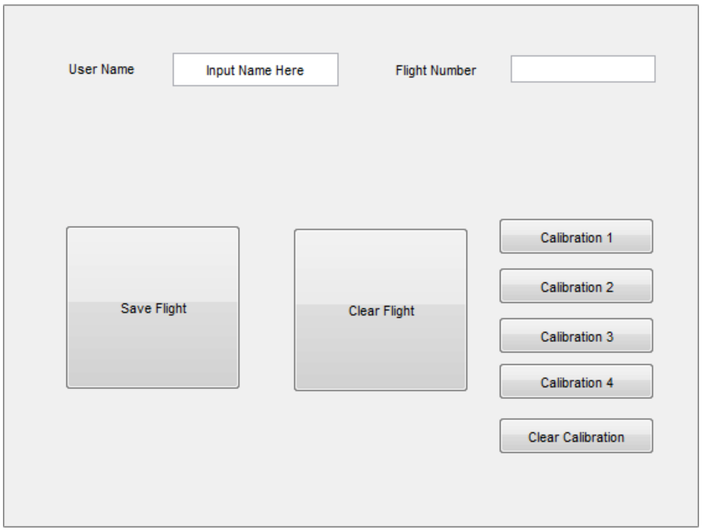


Figure 4.5 - Data Collection User Interface

Its secondary purpose is to allow for a semi-automated calibration of the screens rather than a retroactive calibration as previously performed. This calibration is also flight specific removing any chance of calibration discrepancy caused by head tracker relocation between flights. The calibration data consists of the head orientation with regards to each screen i.e. ‘calibration 1’ = Screen 1; this calibration data are output into a 'calibration' folder contained within each flight folder and each screen receives its own text file with the positional relevant information.

The flight data output by the simulator is of a delimited format with the delimiter being a ‘|’, Matlab reads these data sets (as long as no text is present) and keep the partition based on the ‘|’ delimiter. This type of file, unsurprisingly, is called a delimited file (DLM); the output saved data sets are of a different format and is known as a Comma Separated Variable (CSV) file, instead of using a ‘|’ as the delimiter it uses a ‘,’, hence the comma separated variable name. The CSV file is much more easily read by Matlab.

Figure 4.6 demonstrates the data collection software architecture.

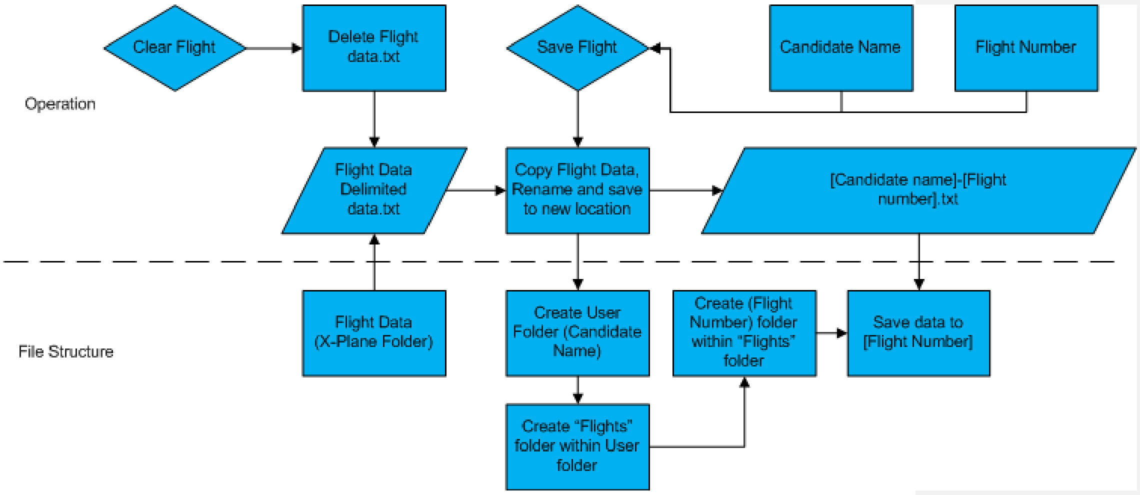


Figure 4.6 - Data collection software collection architecture

With the data saved (both raw and calibration) it was necessary to process the data using a second GUI which also required development.

4.2.2 Post Processing GUI

This GUI is much more complex than the Data Collection GUI and performs multiple operations upon the raw data. Figure 4.7 shows the main interface. Originally designed to run through an entire flight, using a transcript as a base, the user can select participants and their respective flights from the 'Select User' and 'Load Flight' boxes respectively. These are auto-populated within the opening function of the program.

It is then possible to define the type of task being carried out at a certain stage of experimentation, this task type option allows for various variables and metric parameters to be pre-set to allow for different outputs of performance measurement, for example if the task does not include manual flight then the flight data is omitted. Another example would be the type of metric applied to the data, again an example would be using the flight data; if the task is not a landing or take-off task but does require manual control the vertical speed and trace would be measured using variance and standard deviation rather than a polynomial regression root mean squared. The configuration of these parameters is handled by a third GUI called, unsurprisingly, 'Configuration', however, as only one single task was used during the course of this experimentation, the 'Configure' program will not be examined.

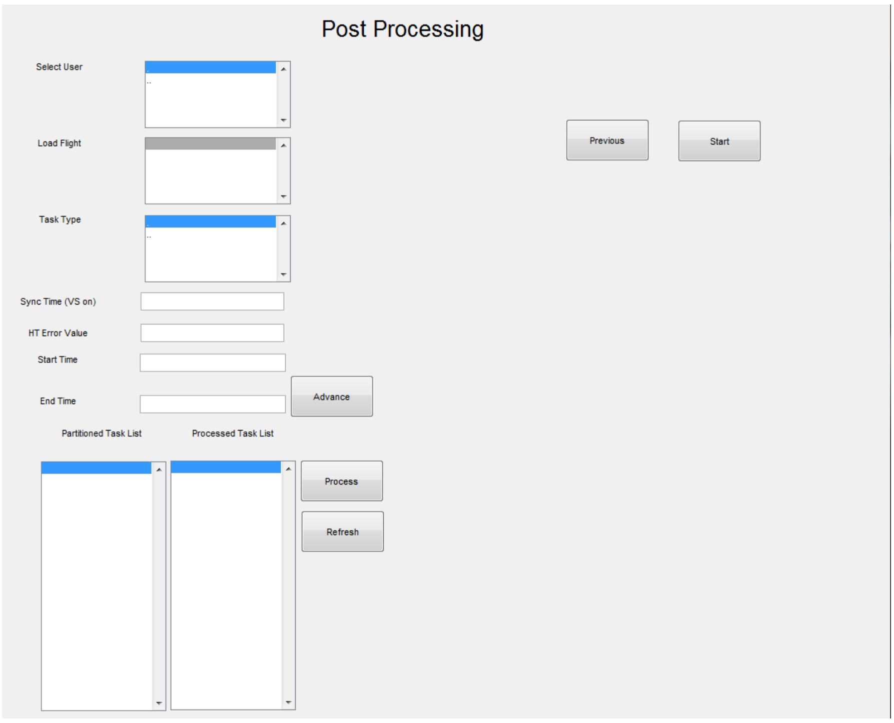


Figure 4.7 - Post processing user interface

'Sync time' is also associated with a transcript of a flight and allows the user to designate when auto pilot was engaged, by identifying the auto pilot engagement noise, within the transcript; the time of this engagement is input into the program and the program will then find the flight data time of autopilot engagement and calculate the differential. This time differential is then applied to all transcripts to data processes. Transcripts were not taken as part of the main experiment and this makes ‘Sync time’ a redundant feature.

'HT Error' allows for the input of a parameter by which all information set usage times below this value were classed as 'void', this is to give a modifiable variable by which to eliminate screen fluctuations where the pilot cannot be thought to realistically be observing any information.

'Start/End times' are also associated with the transcript based experiment; they allowed for the user to input the start and finish times associated with a specific task within the transcript and, using the time differential from the 'Sync', would copy the data between the respective time boundaries and output the task into a time and task labelled 'Partitioned task list' folder on pressing the 'Advance' button. Again, this type of process is not required for the revised experimentation and the start time was set to a 0 default and the end time was left blank so that all the data was processed.

'Partitioned task list' is a database of a single flight’s total task’s, at this point each task represents only raw data which has been partitioned based upon the transcript.

'Processed task list' contains the processed task from the previous list all in chronological order with also the averages of identical tasks taken and output as 'Global Tasks'. The 'Process' button begins the processing of the partitioned tasks so that they may reach this final state.

As each flight is now a single task much of this interface is redundant, but, in case of further research into the field with transcription the interface will remain unchanged.

The data processing architecture is relatively complex so only key functions within processing are now detailed. Figure 4.8 shows an interpretation of basic information set processing operation while figure 4.9 shows basic flight data processing.

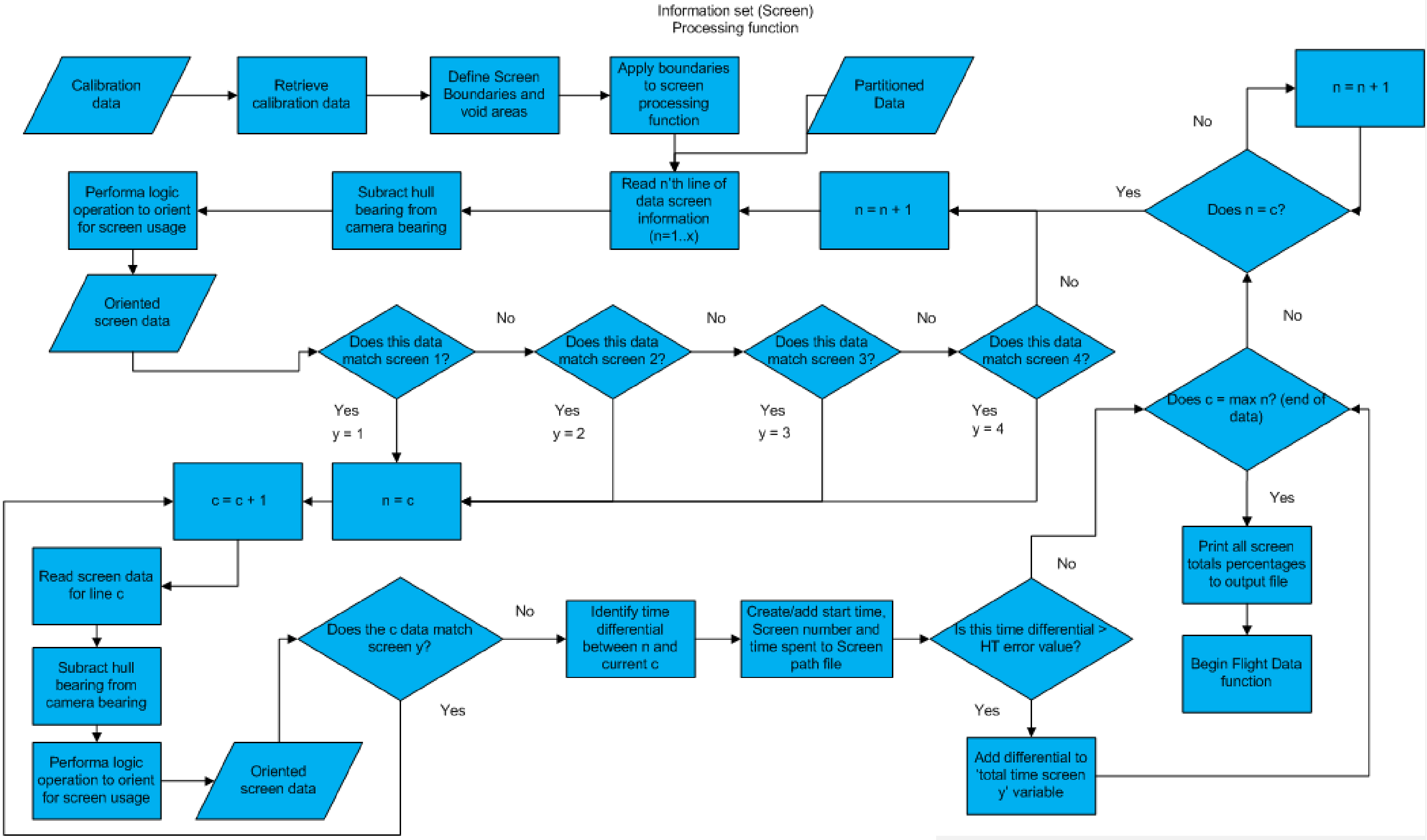


Figure 4.8 - Basic screen processing function architecture

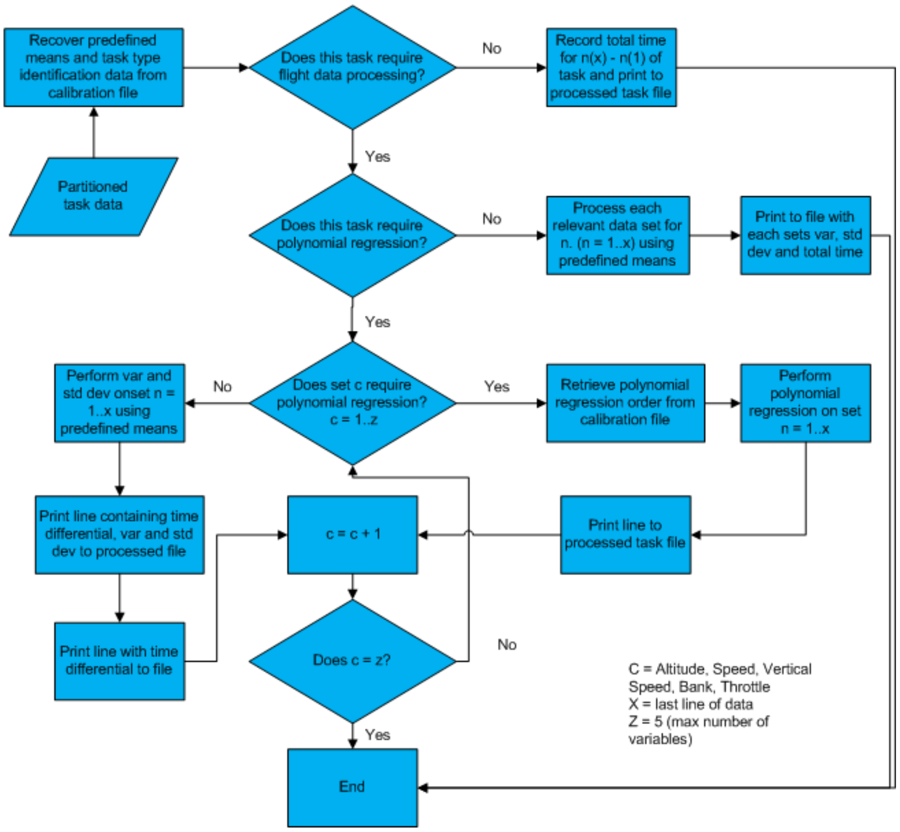
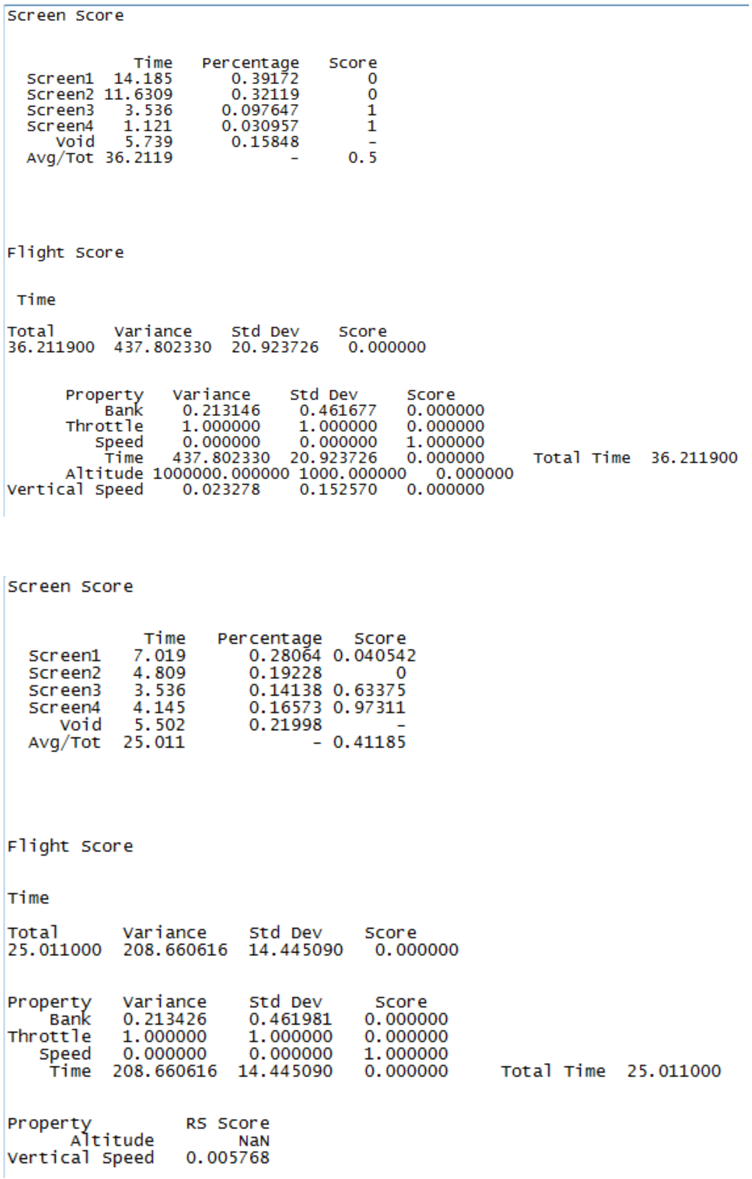


Figure 4.9 - Basic flight data processing function architecture

The final output per task type can be seen below, in figure 4.10, as a comparison between task type outputs. This comparison is no longer of particular interest as only one type of task was output by the forthcoming experimentation; the score columns should also be ignored as these related to the performance measurement metric which is no longer being used.



```
Screen Score


            Time   Percentage   Score
  Screen1  14.185     0.39172       0
  Screen2 11.6309     0.32119       0
  Screen3   3.536    0.097647       1
  Screen4   1.121    0.030957       1
     Void   5.739     0.15848       -
  Avg/Tot 36.2119           -     0.5




Flight Score


 Time

Total      Variance    Std Dev    Score
36.211900  437.802330  20.923726   0.000000


      Property   Variance    Std Dev     Score
          Bank    0.213146    0.461677   0.000000
      Throttle    1.000000    1.000000   0.000000
         Speed    0.000000    0.000000   1.000000
          Time   437.802330  20.923726   0.000000     Total Time  36.211900
      Altitude 1000000.000000 1000.000000   0.000000
Vertical Speed    0.023278    0.152570   0.000000
```

```
Screen Score


            Time   Percentage   Score
  Screen1   7.019      0.28064 0.040542
  Screen2   4.809      0.19228        0
  Screen3   3.536      0.14138 0.63375
  Screen4   4.145      0.16573 0.97311
     Void   5.502      0.21998        -
  Avg/Tot  25.011            - 0.41185




Flight Score


Time

Total      Variance    Std Dev    Score
25.011000  208.660616  14.445090   0.000000


Property   Variance    Std Dev     Score
    Bank   0.213426    0.461981   0.000000
Throttle   1.000000    1.000000   0.000000
   Speed   0.000000    0.000000   1.000000
    Time  208.660616  14.445090   0.000000     Total Time  25.011000


Property        RS Score
      Altitude       NaN
Vertical Speed  0.005768
```

Figure 4.10 - Differing task type output comparison

4.2.3 SimPACT Software Validation

The SimPACT software developed within the previous sections was validated by a ‘trial and error’ process in which either the researcher or a willing volunteer would be asked to perform certain basic tasks. An example of this is testing the screen calibration and identification process in which the volunteer would be asked to look at each screen in turn and the time of screen switch noted; the data from this would be processed and compared to the timing values gathered by the researcher. If an error was present the software would be adjusted and then the process attempted again.

The post processing functions were tested by using ‘dummy’ data taken from an example test flight with known parameters; this allowed for any anomalies in ‘post-processing’ to be identified and corrected. The post-processing function was run numerous times upon the ‘dummy’ data before all the anomalies were rectified.

4.3 Participants Analysis

4.3.1 Description

This study was composed of a total of 30 valid participants, all males between the age range of 18 to 35. Age ranges are not explored during the analysis due to the very small group sizes.

The 30 participants were split into 2 groups, Gamers and Non-gamers; the criterion for being in either group was based upon previous and current gaming experience with

1) gamers (G) being categorised as playing over 3 hours of games per week, on average, for the previous year
2) non-gamers (NG) were classified as playing less than this 3 hour threshold.

This resulted in 2 groups of 15 valid participants; group A (see Appendix F - Participants) had an approximate mean age of 25.3 years and a standard deviation of 3.886 years while group B (see Appendix F - Participants) had an approximate mean age of 27.17 years and a standard deviation of 4.636. Overall the approximate mean age and standard deviation were 26.23 and 4.309 respectively; this shows that the Gamer group tended to be slightly younger with slightly less variation within age range although there is no significant difference or similarity between approximate age ranges (p-value of 0.242). As detailed in chapter 3.5.4 two groups of 15 participants is acceptable for MANOVA comparison with effect size (f^2) of 0.91 for a power of 0.95 and 0.59 for a power of 0.8; correlative statistical analysis is also valid for these group sizes.

Exclusions and Interesting Cases

A total of 34 participants underwent experimentation with 4 of this number being excluded; 3 of these participants were allocated to the NG group but did, in fact, have large amounts of prior gaming experience, however this gaming experience was not detected effectively by questionnaire A and was only identified prior to and during experimentation; once this issue was identified all participants, including those that had already taken part, were asked if they had prior gaming experience. The 3 NG invalid participants all had played computer games excessively when in their teenage years but had stopped prior to, or around, their early twenties; when their data were reviewed in comparison to the Gaming group they had similarities in both flight performance (high degree of flight control), spotting performance (high levels of accuracy and observance) and screen usage (very similar percentage uses) suggesting that although they have no recent games experience this did not diminish their ability to perform well in a virtual environment, possibly due to learned information acquisition/employment processes still being in affect. Due to the small group size their data is not analysed further as a potential subset.

A further Non-Gaming participant, who had almost no experience of any computer-based game, showed the lowest performance scores (in terms of both flight control and spotting score). This should not be a surprise but he indicated, prior to experimentation, that he was a trainee pilot and had real world experience of the control types and information displays; with this in mind his terrible performance was, indeed, unexpected and immediate comparative analysis identified that one aspect of his screen usage (dwell time) was much lower than the norm for both groups. This could indicate that his information processing ability was lacking (not spending long enough on relevant information sets) and this was confirmed by the participant himself when he notified the researcher of a failure on his previous licensing test due to poor visual frame of reference ability. Another possibility is the lack of realism (i.e. motion cues and fidelity) negatively impacting on his potential performance; this is less likely though as another candidate, who displayed the highest ability to control the platform as well as the a perfect object spotting score, also had flight experience.

A single participant had experimentation terminated, part-way through experimentation, due to visual impairment directly effecting his ability to perform adequately; this problem arose due to a visual aid device (large glasses) conflicting with the experimental setup (blinkers). This participant’s data were not used during analysis.

4.3.2 Gaming Experience

When examining participant groups gaming experience, both within group and combined, it was found that the combined groups mean and standard deviation for total years, since they started playing games, were 16.04 years and 6.5 years respectively; only 1 participant had only recently started playing computer games or simulations with only 0.2 years of experience, this participants also turned out to be a special case and is discussed further after the main analysis. The amount of hours spent playing games per week (within the last year) ranged from 0 to 25 with a mean of 7.63 hours showing that most participants do not play excessively and that high levels of game time is not the norm when groups are combined.

Reviewing the pre-study questionnaire (Questionnaire A) it is often difficult to determine what type of simulation is played (unless categorically stated) so it is not viable to compare participants past involvement with flight simulation which may have led to a slight advantage when carrying out the experiment; the responses do seem to indicate that flight simulation does not seem to be a common gaming pastime, with only 10% of the total candidates having tried flight simulation; the most common form of simulation was indicated as driving simulation (40%). The vast majority showed involvement with First Person Shooter (FPS) at 60%, Role Play (RPG) at 50%, Football simulation/management at 20%, Strategy or Real-Time Strategy (RTS) at 43.3% and action/simulation (often Grand Theft Auto) at 26.7%; group G showed more variation in types of game played but it seemed that grouping did not affect the likelihood to play a specific type of game.

When examining the groups individually the G group showed a mean and range for play estimation of 14.66 and 5-25 hours per week respectively; the NG groups play estimation mean and range were 0.61 and 0-3 hours per week, this shows that the groups have been separated successfully based upon their relevant experience. The total length of both groups total gaming experience is, in fact, remarkably similar with group G having a mean of 16.533 years while group NG had a mean of 15.547 years with standard deviations being very similar (5.35 and 7.64 years respectively).

4.3.3 Data Normality Tests & Evaluation

Before any significant data analysis was performed all data sets underwent tests for normal distributions, this is due to a MANOVA test requiring all variables to display a normal distribution. If a variable does not display a normal distribution it must be excluded for all subsequent MANOVA tests; this does not mean however that the variable is not viable for analysis, it can still be subjected to further non-parametric tests and comparison to check for a non-normal distribution. To test for normality the Shapiro-Wilks (S-W) is used, rather than the Kilmolgorov-Smirnoff test as the S-W test relates more to small data sets (less than 100 samples) than larger (greater than 100 samples), if the p-value is less than 0.05 then the data does not show a normal distribution and were then further tested using nonparametric tests, if the tests showed a p-value greater than 0.05 then the data had a non-parametric distribution and could be compared singularly across groups or by using a 2 tailed analysis.

The tables in Appendix F – Normality Test Tables show the results for the normality tests with the non-normally distributed data highlighted; the tests were carried out for each group and across all flights as well as testing the average values for each workload which may show better normality due to the averaging smoothing effect.

As can be seen from the tables (Appendix F) the average of each ‘workload level’, the averaged data yields the most consistent and normally distributed results so instead of comparing individual flights (which are different due to the nature of the experiment) the rest of this analysis is based upon the averaged data. This reduces the outliers and make any anomalies that much more significant; this does however reduce the number of potential samples for each statistical test but it must be remembered that each average is an average of three repetitions so would equate to 45 df (Degrees of Freedom) rather than the 15 df shown. There is some potential for individual flight information, that may have been relevant to the statistical analysis, to be lost in this averaging process but due to the lack of normality any possible finding would likely be invalid for any possible statistical analysis and significance, with a much larger candidate pool individual flight analysis could well have been possible but, as discussed earlier, a large candidate pool was unobtainable.

Flight Data

The flight data (PitchV to VS_SD) is mostly non-normalised, this is to be expected as the optimum is always 0 either for variance (i.e. PitchV = Pitch Variance) or standard deviation (i.e. VS_SD = Vertical Speed Standard Deviation). With 0 being the optimum and only positive values being present, the normality curve is always skewed towards the optimum in a positive fashion; the flight data cannot, therefore, be used with a MANOVA test in this form but must be further tested for non-paramecy to see if the data has a skewed distribution and then compared within and between groups. The flight data was, in-fact, found to have a very skewed distribution which lead to the data undergoing custom paired-group variable transformations (table 4.5) to create a normal distribution from a skewed distribution with no negative values. Other transformations were tried but were either not consistent between paired data or not aggressive enough to create a normal distribution.

Table 4.5 - Flight Variable skewed data transformation table (displaying working transformations)

Variable (x)	Transformation
Pitch SD High (Standard Deviation)	$\frac{1}{\exp(x)}$
Bank SD	$\frac{1}{\exp\left(x^{\frac{1}{2}}\right)}$
Heading (HDG) SD	$\frac{1}{\exp\left(x^{\frac{1}{2}}\right)}$
Altitude (Alt) SD	$\frac{1}{\exp\left(x^{\frac{1}{4}}\right)}$
Vertical Speed (VS_) SD	$\frac{1}{\exp\left(x^{\frac{1}{4}}\right)}$

Further normality tests were then carried out to determine whether this has made the flight data viable; these tests identified that the variables that use variance (i.e. PitchV) as a measure could not be reliably transformed pair-wise into normal distributions and must, therefore, be ignored for all MANOVA tests. The variance data however is still used as part of the descriptive statistics sections relating to each research question as well as used during the further analysis section. The standard deviation (i.e. PitchSD) data could, in most cases be transformed using pair-wise transformations and does show normality as can be seen in the addition transformation columns in the tables within Appendix F; these transformations were not attempted for flight data not relating to participant performance, for example, at the Low workload level there is no human input on flight control so any output data relating to flight control only relates to the models performance rather than the participants. At the medium workload level only bank and heading data are of interest as the pitch/alt/spd data sets were all controlled by the model; at the highest level the participant had control over all parameters except for throttle and speed.

As these variables are model controlled it would not be likely that analysing them would produce any reliable indication of operator performance.

Screen Usage Data

The screen usage data also has the potential to present a skewed distribution; this is due to the measurement value being in percentage use of the screen during the flight and the optimum being unknown and not necessarily at 50%. The lack of normality with screen use can be observed most clearly within the G groups Low level workload table; only Screen 2's percentage (Screen2P) shows a normal distribution on average and even that is close to the non-normal 0.05 p-value threshold which indicates it is likely heavily skewed. This is understandable as screen 2's mean percentage value is 70.65% which means it is less likely, keeping in mind the context of the experiment, that higher percentages could be attained; the range (60% to 80%) shows that the mean value does occur pretty much centrally to the distribution. The normality tests do indicate that as the workload levels increase the screen usage data's normality does also increase to the point, with high level workloads, that a MANOVA test would be applicable for all screen usage data whereas the low level workload has a limited number of normalised variables, especially for group G; although the MANOVA test can be carried out for these variables it would not provide a statistical significance for the available group sizes specifically relating to the screens. It must be noted that at the low level workload screen 4's usage data is expected to be non-normal due to the information set available on the screen being redundant for this workload due to the simulation not requiring any flight control from the participant; this screen did, occasionally receive some percentage use at the lowest workload level but this use can be equated to an occasional glance and would range between 0% and 15% with a mean of 3.89% usage, in terms of time period this give a mean of 4.9 seconds with a range between 0 and 21.7 seconds. The longer periods spent on screen 4, displayed by the outlier candidates, could be due to any number of factors and it is possible that the viewing of this screen had a negative impact on their performance; the other screens percentage use will not be altered with respect to this screens lack of available information but the use of this screen is not evaluated statistically.

Additionally the way in which a participant uses the available displays was examined, this relates to a paired screen transition, for example the participant moves from screen 2 to screen 3; this count is logged over an entire flight and then evaluated as a percentage of the entire screen change count during a singular flight. Although a change to or from screen 1 has been observed to occur these results will not be included as part of the analysis as screen 1 does not display a useable information set during flight and would have no impact on flight performance; the other paired transition percentages have not been corrected due to the use of screen 1 and their respective percentages will remain the same. The normality test tables for the transitions percentages can be found in appendix F.

Spotting and NASA TLX Scores

When examined (see normality tables in Appendix F) for normality the spotting scores displayed interesting results in terms of each groups normality; the NG group displayed two out of three normal distributions over the flight averages (LowAv, MedAv and HighAv) with only the MedAv not showing a normal distribution. The NG's individual, non-averaged flights (i.e. Med1, Med2, Med3), do not show individual normal distribution and it is only once the Medium and High workload flights have been averaged (i.e. MedAv) that a normal distribution appears, a normal data distribution is required for further analysis so individual flights are not analysed and only the workload levels average data is examined. The G group, however, do not display a normal distribution for spotting scores at any point, even once averaged, this is due to heavy skewing towards the ideal of 100% accuracy during each flight, this indicated that the Gamers were extremely effective at the spotting task when compared to the Non-Gamers; it is not possible to adequately transform the score variable for use within MANOVA or for use with non-parametric tests, the score variable is then only to be used with descriptive statistics to evaluate group performance.

The NASA TLX generic variable does display a normal distribution throughout the averaged flights and, can therefore, be used within MANOVA analysis; the subsidiary variables (such as frustration, effort, temporal etc.) within the NASA TLX output are part of the descriptive statistics sections.

4.4 Workload Analysis

This section investigates the experimental results by comparing each group by all workload types and with each workload having its own independent section; it is then possible to compare some aspects of each workload as analysis continues.

In this context the key dependent variables are listed (tables 4.6 & 4.7), firstly, with description of the variable and then followed by reference to their availability for MANOVA comparison and respective workload level.

Table 4.6 - Viable variable list and description

Variable Group	Variable	Description
Screen Data	**Screen1P**	**Percentage overall use of questionnaire screen** *i.e as a percentage of total flight time, how long a participant is identified observing at screen 1*
	Screen2P	**Percentage overall use of virtual world screen** *i.e as a percentage of total flight time, how long a participant is identified as observing screen 2*
	Screen3P	**Percentage overall use of map display screen** *i.e as a percentage of total flight time, how long a participant is identified as observing screen 3*
	Screen4P	**Percentage overall use of instrument display screen** *i.e as a percentage of total flight time, how long a participant is identified as observing screen 4*
	Screen1Dwell	**Average dwell time for questionnaire screen** *i.e. the total time of observation for screen 1 divided by the amount of times observed*
	Screen2Dwell	**Average dwell time for virtual world screen** *i.e. the total time of observation for screen 2 divided by the amount of times observed*
	Screen3Dwell	**Average dwell time for map display screen** *i.e. the total time of observation for screen 3 divided by the amount of times observed*
	Screen4Dwell	**Average dwell time for instrument display screen** *i.e. the total time of observation for screen 4 divided by the amount of times observed*
	Av_2to3	**Percentage of total screen change of screen 2 to screen 3** *i.e. the percentage of the total number of screen change counts that are from screen 2 to screen 3*

	Av_2to4	**Percentage of total screen change of screen 2 to screen 4** *i.e. the percentage of the total number of screen change counts that are from screen 2 to screen 4*
	Av_3to2	**Percentage of total screen change of screen 3 to screen 2** *i.e. the percentage of the total number of screen change counts that are from screen 3 to screen 2*
	Av_3to4	**Percentage of total screen change of screen 3 to screen 4** *i.e. the percentage of the total number of screen change counts that are from screen 3 to screen 4*
	Av_4to2	**Percentage of total screen change of screen 4 to screen 2** *i.e. the percentage of the total number of screen change counts that are from screen 4 to screen 2*
	Av_4to3	**Percentage of total screen change of screen 4 to screen 3** *i.e. the percentage of the total number of screen change counts that are from screen 4 to screen 3*
Flight Data	**PitchSDNorm**	**Normalised standard deviation of total pitch**
	BankSDNorm	**Normalised standard deviation of total bank**
	AltSDNormMA	**Normalised standard deviation of total altitude for medium workload**
	AltSDNormHA	**Normalised standard deviation of total altitude for high workload**
	HDGNorm	**Normalised standard deviation of total heading**
	VS_SDNormHA	**Normalised standard deviation of total vertical speed**
Performance Scores	**Score**	**Percentage correctness for spotting objects**
	TLXscore	**Standardised NASA TLX overall score**

Table 4.7 - Availability of each examined variable between groups for MANOVA analysis

Available for MANOVA for both groups?			
Variable	**Low**	**Medium**	**High**
Screen2P	✓	✓	✓
Screen3P	×	×	✓
Screen4P	×	✓	✓
Screen2Dwell	✓	✓	✓
Screen3Dwell	✓	×	✓
Screen4Dwell	×	✓	✓
Av_2to3	✓	×	✓
Av_2to4	×	✓	✓
Av_3to2	✓	×	✓
Av_3to4	×	×	✓
Av_4to2	×	✓	×
Av_4to3	×	✓	✓
PitchSDNorm	×	×	✓
BankSDNorm	×	✓	✓
AltSDNormMA	×	×	✓
AltSDNormHA	×	×	✓
HDGNorm	×	✓	✓
VS_SDNormHA	×	×	✓
TLXscore	✓	✓	✓

As can be seen from Table 4.7 as workload level increases the likelihood of each variable to display a normal distribution increases; at the low workload level comparatively few variables display a normal distribution and can be considered for MANOVA analysis. This is due to two factors; the first factor is that flight data distributions, at a low workload level, cannot be considered for analysis due to the flight control being entirely automated; with the flight data being automated, firstly, the distributions tend to be heavily skewed to an optimum and, secondly, are not representative of the participants performance (i.e. Evaluation would only test the system and not the participant). The second factor is due to screen 4 not being considered a viable data set at the low workload level as it is not required for any sort of task at this level.

A possible interpretation as to why the medium workload level shows less available normalised variables than the high workload could relate to participant learning strategies coming into effect but could also be due to the increased workload forcing much more exclusive use of the available data sets to allow for successful task completion thus creating a much more standardised usage of the screens. As each flight is different therefore each workload is slightly different, even after averaging; this is further investigated within the descriptive statistics of this section and uses other, non-MANOVA associated, variables for analysis.

4.4.1 Low Workload Analysis

Overall MANOVA Analysis

Table 4.8 which shows the Levene test (see section 3.5.4) on the variables considered for MANOVA analysis; Levene's test shows that all values have a significance greater than 0.05 and this validates homogeneity of variance. The overall MANOVA results are displayed in table 4.9; this shows that Gamer/Non-Gamer grouping had no significant effect, F(6,23) = 0.374 & p = 0.888, on screen usage and TLX scores at this workload level.

Table 4.8 - Overall MANOVA results for G vs NG (lw)

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Screen2P	.028	1	28	.869
Screen2Dwell	.028	1	28	.868
Screen3Dwell	1.456	1	28	.238
Av_2to3	.085	1	28	.773
Av_3to2	1.311	1	28	.262
TLXscore	2.860	1	28	.102

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Group

Table 4.9 - Overall MANOVA results for G vs NG (lw)

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[c]
Group	Pillai's Trace	.089	.374[b]	6.000	23.000	.888	.089	2.243	.132
	Wilks' Lambda	.911	**.374[b]**	6.000	23.000	**.888**	.089	2.243	.132
	Hotelling's Trace	.098	.374[b]	6.000	23.000	.888	.089	2.243	.132
	Roy's Largest Root	.098	.374[b]	6.000	23.000	.888	.089	2.243	.132

c. Computed using alpha = .05

Table 4.10 shows the relevant results as separate ANOVA's, it can also be seen from this table that none of the computed variables display significant differences (F > 0.6 & sig(p) < 0.05) between groups.

Table 4.10 – Independent ANOVA results for computed variables (lw)

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[g]
Group	Screen2P	7.338E-005	1	7.338E-005	.014	.906	.001	.014	.052
	Screen2Dwell	1.744	1	1.744	.355*	.556	.013	.355	.089
	Screen3Dwell	.007	1	.007	.021	.886	.001	.021	.052
	Av_2to3	1.312E-005	1	1.312E-005	.002	.969	.000	.002	.050
	Av_3to2	.001	1	.001	.100	.754	.004	.100	.061
	TLXscore	381.681	1	381.681	.863*	.361	.030	.863	.146

From these results we can conclude that there is no statistically significant difference between Gamers and Non-Gamers at the low workload level, with either or both the significance (p) value being greater than 0.05 and the effect size (F) being less than 0.6, in terms of both screen usage and TLX score

Screen Usage Analysis

The descriptive comparison of the screen percentage usage, dwell and change demonstrates the similarity of both G and NG groups, this can be seen figures 4.11 to 4.12 and demonstrate the similarity of the relevant results. The associated means and standard deviations demonstrate the relative similarity of the groups with very low variations in these values being shown between groups. Figure 4.13 shows a comparison of screen change frequency with the NG group displaying a larger variance of screen change.

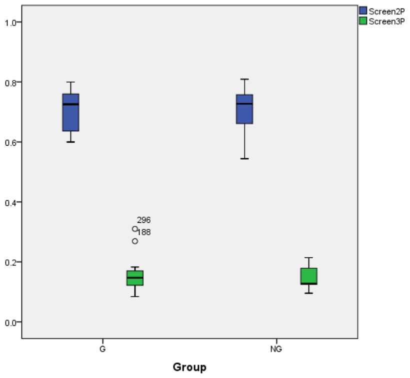


Figure 4.11 – Descriptive comparison of group percentage screen usage, screen 2 and screen 3 (1.00 = 100%)(lw).
Screen 2 - Gamer (m = 70.65% & sd = 7.12%), Non-Gamer (m = 70.96% & sd = 7.28%).
Screen 3 - Gamer (m = 15.83% & sd = 5.99%), Non-Gamer (m = 14.62% & sd = 3.611%).

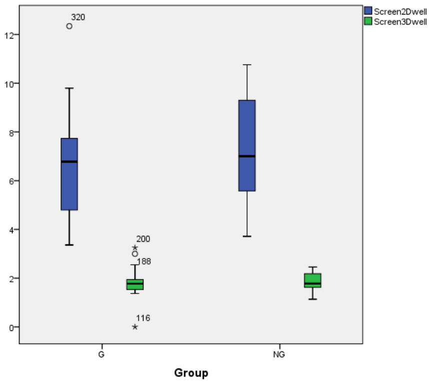


Figure 4.12 – Descriptive comparison of group average dwell time screen 2 and screen 3(lw).
Screen 2 - Gamer (m = 6.75s & sd = 2.34s), Non-Gamer (m = 7.23s & sd = 2.09s)
Screen 3 - Gamer (m = 1.82s & sd = 0.75s), Non-Gamer (m = 1.85s & sd = 0.37s).

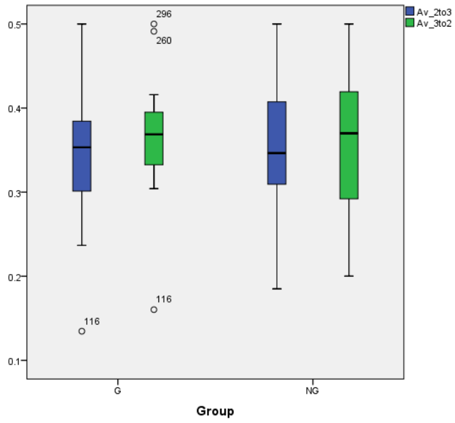


Figure 4.13 – Descriptive comparison of group screen change percentage screen 2 to 3 and 3 to 2 (lw). Av2to3 - Gamer (m = 34.58% & sd = 9.25%), Non-Gamer (m = 34.71% & sd = 9.42%) Av3to2 - Gamer (m = 36.64% & sd = 8.05%), Non-Gamer (m = 35.63% & sd = 9.40%).

Further analysis using Spearman's rank correlation (Rho, ρ) does produce some interesting findings regarding the relation of screen-to-screen data as well as screen-to-performance data. Both groups display a significant positive correlation when comparing screen 2 percentage usage with the screen 2 average dwell times (**G** – ρ = 0.864, sig = 0.000 & **NG** – ρ = 0.836, sig = 0.000) demonstrating a higher overall percentage use of the screen will lead to higher dwell times, the two groups differ however when comparing screen 3 percentage usage to its relative dwell time; only the Gamer group displays a significant correlation at this point (ρ = 0.657, sig = 0.008). This could indicate a potential formalised data acquisition and processing strategy that is stronger in the Gamer group; a formalised strategy means that a particular process for gathering information has been created, this is based upon information requirements of the candidate and their own identification of those requirements which leads to either conscious or sub-conscious data acquisition patterns. This can then be compared with the following higher workloads.

Some strong and weak correlations exist between screen change averages, percentage use and dwell times; all of the existing strong correlations between data sets (for example, percentage use and dwell time) are the same for both groups whereas there is some difference in weak correlations between the two groups which may indicate an overall difference between the two groups, but due to the strength of these no conclusions could be drawn.

No correlations exist within dwell time variables or within percentage use variables.

TLX Score Analysis

The descriptive statistics also support the conclusion, of the independent ANOVA (F = 0.863, p = 0.361), that there is no significant difference between groups with only the TLX score showing some difference in standard deviation (**G –** SD = 17.82, **NG –** SD = 23.8). Figure 4.14 demonstrates the increased standard deviation from the mean for the NG's than the G's with a predominately higher distribution at higher TLX scores. The Non-Gamers mean TLX score (m = 36.55) is also higher than that of the Gamers (m = 29.42) indicating that they found the task slightly more taxing, this could be due to the Non-Gamer's lack of familiarity with using a games based platform and, therefore, finding the situation more taxing.

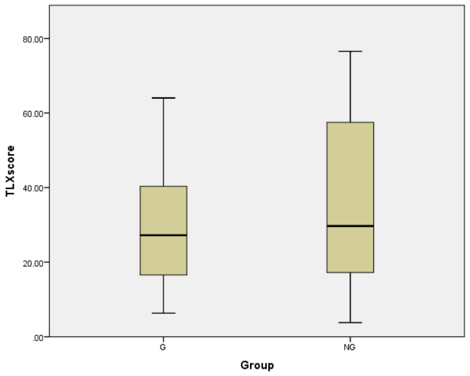


Figure 4.14 – G vs. NG for TLX score (lw)

Object Spotting Analysis

Spotting scores display a graphical performance differential (figure 4.15); although not testable through ANOVA due to non-normality; the parametric Mann-Whitney tests show that an averaged score (LowAv) of the three low level flights, of p = 0.325, which accepts the hypothesis that the distribution for both groups are similar and means that both groups had similar distributive characteristics in terms of spotting performance.

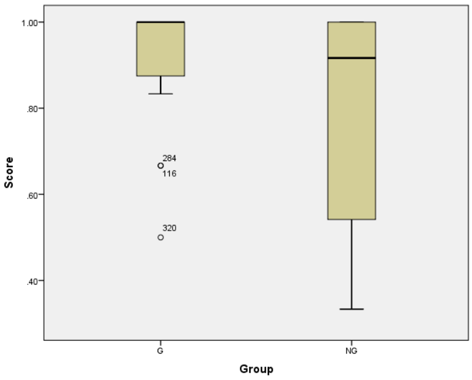


Figure 4.15 – Averaged spotting score (LowAv) by groups (lw)

Figure 4.16 shows that the G group receive predominantly higher spotting scores with a mean of 0.9 (or 90%) accuracy as opposed to the 0.79 (or 79%) accuracy of the NG group; the G group also displayed a standard deviation 0.16 (or 16%), lower than that of the NG group which was 10% higher at 0.26 (26%).

The histogram below (figure 4.16) demonstrates the G groups' predominance for correctness as opposed to the NG group.

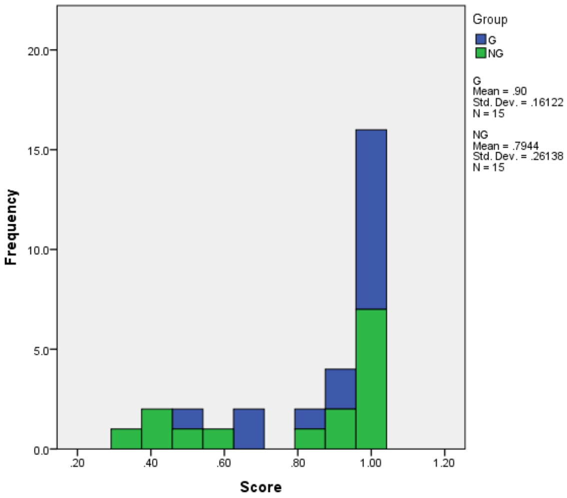


Figure 4.16 – Spotting accuracy graphical comparison (1.00 = 100%)(lw)

This strong indication of a performance differential at the same workload, due to the large difference in means and standard deviation, is an indication that gamers were more successful at spotting objects but did not necessarily differ from the non-gamers in terms of screen usage or workload indication.

Cross-Correlations

When comparing screens, TLX and spotting score only a single weak correlation was found within the Gaming group relating to the Score-TLX correlation (ρ = -0.563, sig = 0.029) suggesting that the Gamers who found the flights easier/less taxing tended to perform to a better level. The Non-Gamer group however displayed multiple correlations with screen usage, TLX score and spotting score, these can be seen in the table below (table 4.11).

This suggests that the way in which the Non-Gamers utilised the available screens had a large impact on both subjective workload levels and spotting performance; this suggests that a higher overall percentage use and dwell time of a participant on screen 2 lowers level of subjective workload and increases spotting performance.

No correlations were found for the gaming group with respect to their recent gaming activity (play estimate) indicating that recent games experience had no effect at this workload; this correlation was not attempted for the Non-Gaming group due to the limited data set.

Table 4.11 – Non-Gamer Screen, TLX, Score Correlations (lw)

Correlation	ρ	significance
Screen2P-Score	0.781	0.001
Screen2P-TLX	-0.736	0.002
Sc2Dwell-Score	0.649	0.009
Sc2Dwell-TLX	-0.521	0.046
Score-TLX	-0.651	0.009

4.4.2 Medium Workload Analysis

Overall MANOVA Analysis

Table 4.12 shows the Levene test on the variables considered for MANOVA analysis; Levene’s test show that all but two values have a significance greater than 0.05; the HDGDnorm and Av_4to2 variables show a lack of homogeneity of variance indicating that the two groups variance over these variables is markedly different. These two data sets were subsequently removed from the MANOVA but are further descriptively analysed. The overall MANOVA result is displayed in table 4.13; this shows that grouping had, again, no significant effect, F(8,21) = 1.075 & p = 0.417, on screen usage and TLX scores at this workload level.

Table 4.12 - Levene’s test for homogeneity of variance (mw)

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Screen2P	.871	1	28	.359
Screen4P	.182	1	28	.673
Screen2Dwell	.226	1	28	.638
Screen4Dwell	2.858	1	28	.102
Av_2to4	.827	1	28	.371
Av_4to2	5.509	1	28	.026**
Av_4to3	1.579	1	28	.219
BankSDNorm	1.315	1	28	.261
HDGDNorm	8.489	1	28	.007**
TLXscore	.384	1	28	.540

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Group

**. Non-homogeneity of variance

Table 4.13 - Overall MANOVA results for Gamers vs Non-Gamers (mw)

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[c]
Group	Pillai's Trace	.291	1.075[b]	8.000	21.000	.417	.291	8.600	.368
	Wilks' Lambda	.709	**1.075[b]**	8.000	21.000	**.417**	.291	8.600	.368
	Hotelling's Trace	.410	1.075[b]	8.000	21.000	.417	.291	8.600	.368
	Roy's Largest Root	.410	1.075[b]	8.000	21.000	.417	.291	8.600	.368

c. Computed using alpha = .05

Table 4.14 - Independent ANOVA results for computed variables (mw)

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[i]
Group	Screen2P	.002	1	.002	.226	.638	.008	.226	.075
	Screen4P	.003	1	.003	.699*	.410	.024	.699	.127
	Screen2Dwell	.043	1	.043	.046	.832	.002	.046	.055
	Screen4Dwell	.007	1	.007	.007	.934	.000	.007	.051
	Av_2to4	.007	1	.007	1.884*	.181	.063	1.884	.263
	Av_4to3	.001	1	.001	.570	.457	.020	.570	.113
	BankSDNorm	.030	1	.030	**6.187***	**.019****	.181	6.187	.671
	TLXscore	41.395	1	41.395	.115	.737	.004	.115	.062

i. Computed using alpha = .05

*. Significantly different (F>0.6)

**. Significantly different (p>0.05)

Table 4.14 shows the relevant results as separate ANOVA's. BankSDNorm is the only variable that displays significant difference between groups with F(1,28) = 6.187 and p = 0.19; as this represents the only available flight performance data source within the MANOVA it can be assumed that there is a strong likelihood that the Gamers were significantly better at flight control than Non-Gamers.

Screen Usage Analysis

Screen data, once again, shows no significant difference between groups; a MANOVA test using only screen related variables available at this workload gives a result of F(6,23) = 1.051 & p = 0.420, this can be compared to an only screen data MANOVA of the low workload which results in F(5,24) = 0.325 & p = 0.893; this result could indicate a possible divergence in behaviour between groups related solely to screen data. However the result is purely speculative and no conclusions can be directly drawn from it. It is worth noting however that this may indicate the Gamer and Non-Gamers may be starting to differ slightly in the way the use each screen respective to the workload level.

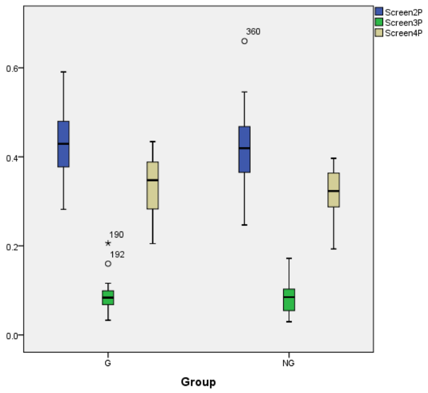


Figure 4.17 - Descriptive comparison of group percentage screen usage (1.00 = 100%)(mw)

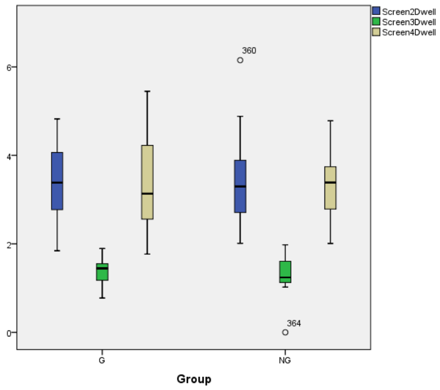


Figure 4.18 - Descriptive comparison of group average dwell time (mw)

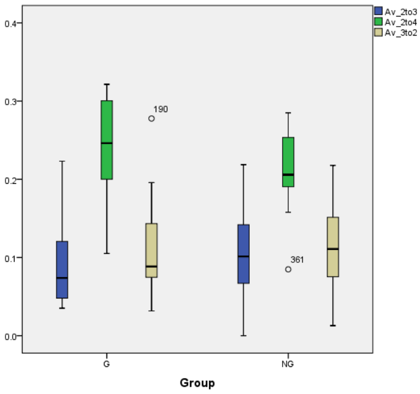


Figure 4.19 - Descriptive comparison of group screen change percentage – part 1 (mw)

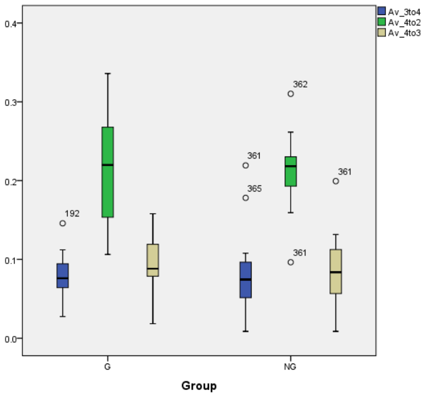


Figure 4.20 - Descriptive comparison of group screen change percentage – part 2 (mw)

Figures 4.17 to 4.20 display a graphical interpretation of the screen usage and include variables not available for MANOVA/ANOVA comparison, it is worth comparing these variables as they may give further insight into screen use.

The above comparisons again display relative similarity in terms of means but it can be seen with Av_4to2 (which was previously excluded for the MANOVA test due to non-homogeneity of variance) in figure 4.17 that there appears to be a much greater standard deviation differential between groups; the differential between groups standard deviation for this variable is 0.025 or 2.481% (see table 4.15, Av_4to2). This differential is higher than any other variables standard deviation differential (approximately 0.01 or 1% higher) but with the size of the respective means it is not likely that this is a significant difference.

Table 4.15 – Descriptive means and standard deviation for percentage screen change (mw)

Descriptives

	Group		Statistic	Std. Error
Av_4to2	G	Mean	.2152	.01867
		Std. Deviation	.07230	
	NG	Mean	.2114	.01227
		Std. Deviation	.04751	
Av_3to4	G	Mean	.0795	.00761
		Std. Deviation	.02947	
	NG	Mean	.0844	.01389
		Std. Deviation	.05378	
Av_2to3	G	Mean	.0914	.01550
		Std. Deviation	.06003	
	NG	Mean	.1034	.01605
		Std. Deviation	.06218	
Av_2to4	G	Mean	.2450	.01740
		Std. Deviation	.06738	
	NG	Mean	.2144	.01400
		Std. Deviation	.05422	
Av_3to2	G	Mean	.1107	.01663
		Std. Deviation	.06442	
	NG	Mean	.1128	.01462
		Std. Deviation	.05664	
Av_4to3	G	Mean	.0967	.00907
		Std. Deviation	.03513	
	NG	Mean	.0850	.01245
		Std. Deviation	.04823	

Further analysis with the use of Spearman’s correlations again does show some differences between Gamers and Non-Gamers in terms of inter screen usage correlation. Significant correlations were found within the Gamer group when correlating only screen data (table 4.16); the Gamer group, once again (see low workload analysis), showed significant positive correlation between a screen and its respective dwell time.

Table 4.16 – Screen data correlation for Gamers and Non-Gamers

	Gamer		**Non-Gamer**	
Correlation	ρ	significance	ρ	significance
Sc2P-Sc2Dwell	**0.796**	**0.000**	**0.761**	**0.001**
Sc3P-Sc3Dwell	**0.721**	**0.002**	0.589	0.21
Sc4P-Sc4Dwell	**0.689**	**0.004**	0.239	0.390
Sc2P-Sc4Dwell	**-0.661**	**0.007**	-0.086	0.761
Sc4P-Sc2Dwell	-0.332	0.226	-0.521	0.046

Correlations: Bold text – **Strong**, Underlined – Weak, Normal - None

This indicates that the Gamer group, with higher percentage use of a respective screen during a flight also had a corresponding increase in the average time that they spend on that screen during each glance. The Non-Gamer group only show a single strong and single weak correlation relating to this pattern suggesting that the Non-Gamers have a differing, or non-existent, strategy for information gathering and processing while the Gamers display a strong and inter-group similarity for data gathering and processing.

The non-screen respective correlations (Sc2P-Sc4Dwell & Sc4P-Sc2Dwell) show an unexpected relationship between the use of one screen and the dwell time of another; the Gamer group display a tendency that for higher screen 2 usage there is a reduction in the average dwell time of screen 4 whereas the Non-Gamers display a tendency that for higher screen 4 usage there is a reduction in screen 2 dwell time. This could indicate a difference in group prioritisation of information or data sets with the Gamers holding object spotting (use of screen 2) at a higher level than flight control (use of screen 4); this could also be a reflection of their ability and confidence to control the platform. The Non-Gamers display, in effect, the opposite of this with a weak correlation indicating they may have prioritised flight control over object spotting.

No correlation exists between screen dwell times, but a strong and weak negative correlation does exist respectively for Non-Gamers (ρ = -0.720, sig = 0.002) and Gamers (ρ = -0.593, sig = 0.02) between screen 2 and screen 4 percentage usage. This suggests that both groups have similar behaviour with screen usage as no correlation exists with screen 3 for either of the two previous variables.

Correlating screen percentage use and average dwell time to the average screen change again yielded strong and weak correlations which were similar for both groups and which also do not show a conclusive result.

Overall it seems that a Gamers potential increased prioritisation of an information set (screen) will lead to increased dwell times indicating an inbuilt strategy relating to information acquisition and prioritisation; the Non-Gamers display this potential to a lesser, if not inconclusive, extent. This pattern is further explored in the next section (high workload analysis).

Flight Data Analysis

Due to the normalised heading standard deviation failing the Levene’s test (i.e. the data displayed a non-normal distribution) it is not possible to carry out a MANOVA upon the two flight variables available, also a MANOVA with only two variables would not be significant for the relative group sizing. Instead they need to be explored descriptively, Figure 4.21 and table 4.17 demonstrate descriptively the difference in flight performance; figure 4.21, for BankSD and HDGSD, represent the non-normalised data outputs and show the difference between the two groups.

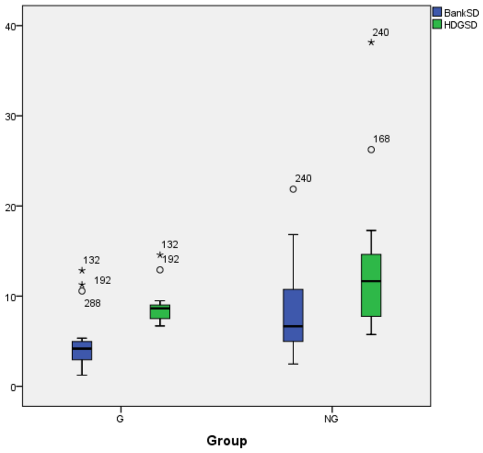


Figure 4.21 – Flight performance comparison (mw)

Table 4.17 – Flight performance comparison, mean and standard deviation (mw)

		N	Mean	Std. Deviation
BankSD	G	15	5.0322	3.56337
	NG	15	8.6446	5.66644
	Total	30	6.8384	5.00054
HDGSD	G	15	8.9215	2.15324
	NG	15	13.4513	8.60755
	Total	30	11.1864	6.58122

The variance in mean is lower for Gamers in both variables as is the standard deviation, significantly so in heading; fluctuations within bank control have an effect with heading but it is also possible to have a high bank mean and standard deviation while still retaining a low mean and standard deviation within heading. For example, it is possible to have large degrees of bank angle but still stay roughly on the same heading creating a low heading mean while having a high bank mean; this can also work conversely where a participant may only bank slightly but maintain that same angle of bank and then varying largely on the heading. This is why these two variables, although linked, are independent from one another and can be examined separately and both have independent impact on significance tests. Exploring this further with a correlation between heading and bank standard deviation yielded a weak correlation for the gaming group (ρ = 0.554, sig = 0.032) and a very strong correlation for the Non-Gaming group (ρ = 0.882 and sig = 0.00); this indicates that the Non-Gamers, when applying larger bank angles would then change heading more significantly and potentially remain more off heading whereas the gamers where less likely to change from the optimum heading even with increasing bank angles.

Irrespective of this correlation it can be seen that the Non-Gaming group performed worse in terms of flight control than the Gamers.

TLX Score Analysis

The NASA TLX ANOVA, performed in the previous MANOVA test, displayed no significant difference (F = 0.115 ,p = 0.737) between groups. Descriptively there is a small difference in mean (**G** – 47.47, **NG** – 49.82) and standard deviation (**G** – 17.79, **NG** – 20.11), this can be seen in figure 4.22; although the Non-Gamer figures are slightly higher there is no significant indication of difference of subjective workload interpretation between groups.

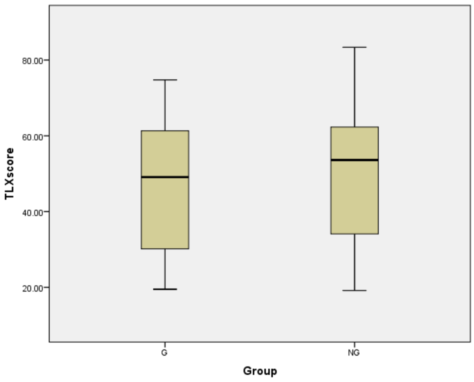


Figure 4.22 - Group TLX score comparison (mw)

Figure 4.22 indicates that although the means are similar a few of the Non-Gamer group found the medium workload more challenging than the Gamer group but when comparing to the means and standard deviation this difference is slight and is only due to a few Non-Gamer outliers exceeding the outliers of the Gaming group.

Compared to the low workload mean and standard deviation the Gamers mean has increased much more to the point of almost matching that of the Non-Gamers showing that the increase in workload has dramatically affected their interpretation of that workload, whereas the Non-Gamers only increased slightly indicating that they initially reached a ‘saturation’ point within the low workload scenario. It appears that it took this workload increase to take the Gamers to a saturation point where both groups are finding the task nearly equally taxing.

Object Spotting Analysis

Object spotting displayed a large performance differential between Gamers (mean = 0.833, SD = 0.244) and Non-Gamers (mean = 0.578, SD = 0.293); although there is a large difference (25.5%) between means showing a much greater degree of correctness the standard deviations of each group are now approximately similar suggesting that both groups have similar internal performance distribution around their respective means. The spotting results can be seen in figures 4.23 and 4.24; from these graphs the Gamer tendency for near 100% correctness can be observed with 8 of the 15 Gamers receiving perfect (100%) scores. In comparison to the Gamer group only 3 out of the 15 Non-Gamers received perfect scores. Compared to the low workload there has been a slight reduction in Gamer spotting performance (0.9 to 0.83), whereas the Non-Gamers have a major reduction in spotting score (0.79 to 0.578), a 21.2% drop in spotting score accuracy. This shows that the Non-Gamers effectiveness at task completion has been heavily reduced due to the increase in workload and added flight variables; this could also be partly due to lack of stability of the platform (caused by the model tending to bank and pitch independently if no input is applied) reducing the ability to correctly locate and identify objects. A higher ability to control the platform and keep it stable are results in higher camera and on-screen object stability. Either way, this indicates a reduction in overall effectiveness at task completion when compared to the Gamer group, the lack of stability of the platform just compounds and increases the difficulty of both tasks.

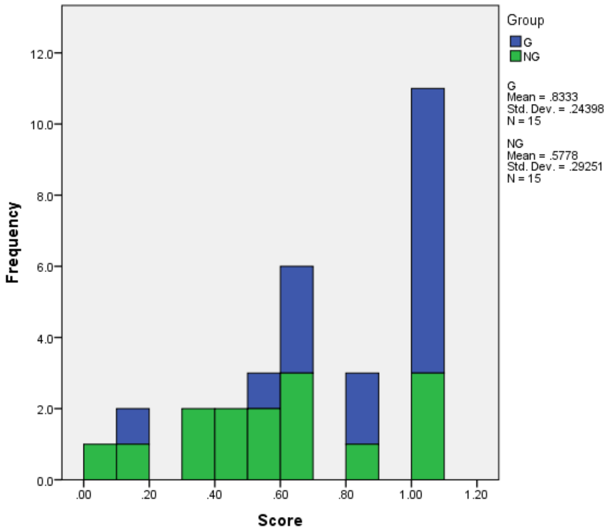


Figure 4.23 - Spotting accuracy graphical comparison (1.00 = 100%)(mw)

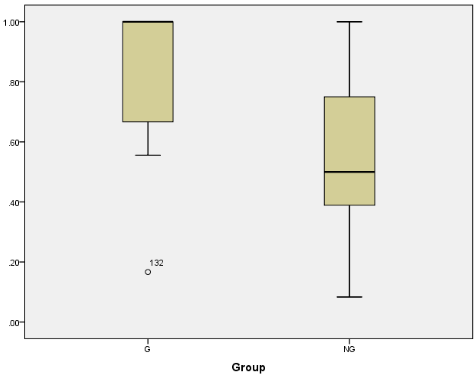


Figure 4.24 - Averaged spotting score (MedAv) by groups (mw)

Despite a large performance differential observed within the flight and object spotting data, screen data displays no significant difference within MANOVA or individual ANOVA comparison even with some results showing a trend towards a significant difference, AV_2to4 especially; this means that both groups still utilise the screens, overall, in a similar fashion in terms of most variables (percentage use, dwell times, etc.). The overall MANOVA's significance is also decreasing, from Low workload (F(6,23) = 0.374 & p = 0.888) to Medium workload (F(8,21) = 1.05 & p = 0.417) which suggests a potential in divergent overall performance of the groups.

Cross-Correlation

No correlations, either strong or weak, could be found when correlating the Gamers Screen and Flight data sets and only one weak negative correlation could be found between heading standard deviation (hdgsd) and screen 3 dwell (sc3dwell) times for the Non-Gamers (ρ = -0.539, sig = 0.038). This may indicate a tenuous link between these variables for the Non-Gamers, but is more likely a result of coincidence as no other correlations were found.

A correlation of flight data to spotting score produced a weak negative correlation between bank standard deviation and spotting score (ρ = -0.549, sig = 0.034) which may suggest that larger bank deviation causes loss of effectiveness at spotting within the Gamer group; when testing the Non-Gamer group for this same correlation a strong negative correlation was observed (ρ = -0.641, sig = 0.01) with a further weak correlation between score and heading standard deviation (ρ = -0.524, sig = 0.045). This finding supports the previous suggestion that large bank angles (poor flight control) adversely affect the Non-Gamers ability to locate and identify objects. The inability to correctly identify objects while being unable to keep the platform stable is to be expected.

No significant or even weak correlations were observed between Screen-TLX-Score data, with either group, suggesting any previous correlations or dependencies no longer exist; this could be due to the extra workload (and extra data set of screen 4) over observed within the Non-Gamer group.

A correlation between play estimation (estimated amount of hours played per week) revealed some interesting correlations within the gaming group; it is unwise to try the same correlation with the Non-Gamer group due to the limited range of the play estimate data set (only a 3 hour range) created by the group partitioning criteria (0 to max 3 hours for Non-Gamers, 4 onwards for Gamers).

Regarding play estimation and screen data a weak positive correlation was found with screen 4 percentage use ($\rho = 0.542$, sig = 0.037); this, on its own, may not indicate significance but a strong correlation was further found with screen 4 dwell times ($\rho = 0.662$, sig = 0.007). This indicates that a more recently experienced gamer (a gamer who has played more hours per week than his peer) was likely to dwell on screen 4 for longer; this again may reflect a stronger information prioritisation pattern suggested by the screen to dwell time findings.

This is not reflected in the play estimate-flight data correlation suggesting that more recently experienced Gamers had no more likelihood to perform well than less experienced Gamers.

No other correlations were found relating to play estimate showing that, overall, the amount of recent gaming experience was not directly linked to performance output at this workload.

4.4.3 High Workload Analysis

Overall MANOVA Analysis

Levene's test (table 4.18) shows a lack of homogeneity of variance with only one variable, the NASA TLX workload scores and this has been excluded from the following overall MANOVA test. Again the MANOVA test, which can be seen below in table 4.19, shows no significant difference between Gamer and Non-Gamer groups; the effect size has now increased to a significant level (F(16,13) = 0.956) while the significance level has increased (p = 0.541) but not by a large amount compared to the medium workload MANOVA and still does not display significant difference.

Table 4.18 - Levene's test for homogeneity of variance (hw)

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Screen2P	.298	1	28	.590
Screen2Dwell	.305	1	28	.585
Screen3P	.854	1	28	.363
Screen3Dwell	.001	1	28	.979
Screen4P	.098	1	28	.756
Screen4Dwell	1.318	1	28	.261
Av_2to3	.073	1	28	.789
Av_2to4	.092	1	28	.764
Av_3to2	.477	1	28	.496
Av_3to4	3.814	1	28	.061
Av_4to3	.020	1	28	.889
TLXscore	12.055	1	28	.002**
PitchSDNorm	.008	1	28	.930
BankSDNorm	2.395	1	28	.133
AltSDNormHA	2.594	1	28	.118
HDGDNorm	.006	1	28	.939
VS_SDNormHA	.064	1	28	.802

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Group

**. Non-homogeneity of variance

Table 4.19 - Overall MANOVA results for G vs NG (hw)

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[c]
Group	Pillai's Trace	.541	.956	16.000	13.000	.541	.541	15.297	.327
	Wilks' Lambda	.459	**.956***	16.000	13.000	**.541**	.541	15.297	.327
	Hotelling's Trace	1.177	.956	16.000	13.000	.541	.541	15.297	.327
	Roy's Largest Root	1.177	.956	16.000	13.000	.541	.541	15.297	.327

a. Design: Intercept + Group

b. Exact statistic

c. Computed using alpha = .05

*. Significantly different (F>0.6)

The individual ANOVA’s (table 4.20) however now display interesting significant differences within the flight data; four out of the five flight performance measurements display significant difference with only the AltSDNorm being non-significant. This result suggests that, although both groups again had no significant difference in terms of screen usage and patterns, they may have significantly different flight performance.

Table 4.20 - Independent ANOVA results for computed variables (hw)

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[p]
Group	Screen2P	.005	1	.005	.383	.541	.013	.383	.092
	Screen2Dwell	.023	1	.023	.017	.896	.001	.017	.052
	Screen3P	.001	1	.001	2.102*	.158	.070	2.102	.288
	Screen3Dwell	.036	1	.036	.931*	.343	.032	.931	.154
	Screen4P	.002	1	.002	.160	.692	.006	.160	.067
	Screen4Dwell	.020	1	.020	.003	.955	.000	.003	.050
	Av_2to3	.002	1	.002	.450	.508	.016	.450	.099
	Av_2to4	.002	1	.002	.680*	.417	.024	.680	.125
	Av_3to2	.001	1	.001	.268	.609	.009	.268	.079
	Av_3to4	.002	1	.002	1.392*	.248	.047	1.392	.207
	Av_4to3	.003	1	.003	1.298*	.264	.044	1.298	.196
	PitchSDNorm	.163	1	.163	**9.480***	**.005****	.253	9.480	.844
	BankSDNorm	.040	1	.040	**12.749***	**.001****	.313	12.749	.931
	AltSDNormHA	9.809E-006	1	9.809E-006	.948*	.339	.033	.948	.156
	HDGDNorm	.006	1	.006	**9.787***	**.004****	.259	9.787	.855
	VS_SDNormHA	.000	1	.000	**7.447***	**.011****	.210	7.447	.750

p. Computed using alpha = .05

*. Significantly different (F>0.6)

**. Significantly different (p>0.05)

Screen Usage Analysis

The independent ANOVA's have revealed no significant differences between groups; a MANOVA associated solely with the screen usage data returned non-significant values of F(11,18) = 0.883 and p = 0.572 showing that, as a group, these variables also display no significant difference. Compared to the previous workload the significance level has, in fact, increased suggesting that the hypothesis of divergent screen behaviour is incorrect. Figures 4.25 and 4.26 demonstrate this relative similarity of usage in terms of dwell and percentage use respectively. The average screen change data also displays this similarity between groups. Descriptive statistics can be found in appendix F.

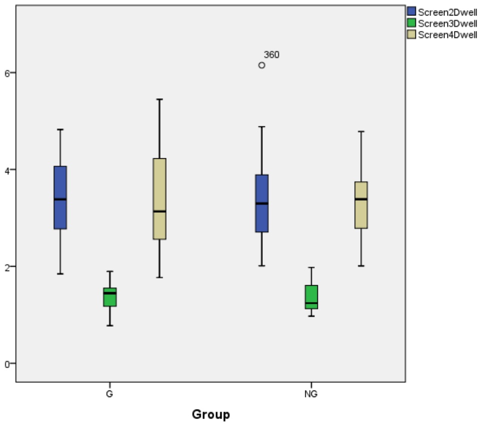


Figure 4.25 - Screen Dwell time Group Comparison

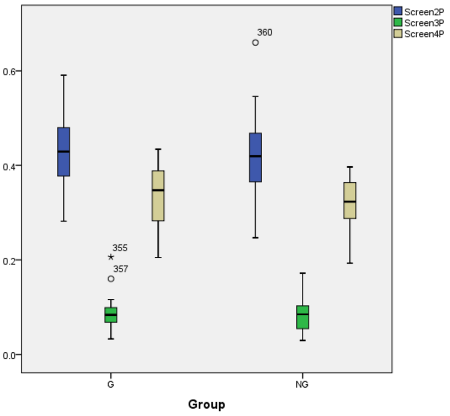


Figure 4.26 - Percentage Screen Usage Group Comparison

Correlation of screen percentage use showed some strong correlations between variables for both groups. The correlation between screen 2 and screen 4 percentage use yielded very strong negative correlations (**G** – ρ = -0.925, sig = 0.000; **NG** - ρ = -0.871, sig = 0.000), this correlation did not exist for screen 3 with the others so it would be reasonable to conclude that, for both groups, as one screen was used more the other would be used less. This is an expected result.

Correlation of only screen dwell times also produced some strong correlations but only for the Non-Gaming group (see table 4.21); this displays a potential learning behaviour by which Non-Gamers, with the increase in useful information on screen 4 increase their dwell time which then, in turn increases their average dwell time for screen 3. This indicates that they are creating a rudimentary data acquisition process that is not seen in the Gamers; this could be due to the Gamers already having existing strategies associated with data acquisition that is not based on a global increase in dwell time but more finely tuned to data prioritisation.

Table 4.21 - Screen Dwell Time Comparison (hw)

Correlations

			Screen2Dwell	Screen3Dwell	Screen4Dwell
Spearman's rho (Non-Gamers)	Screen2Dwell	Correlation Coefficient	1.000	.407	.679**
		Sig. (2-tailed)	.	.132	.005
	Screen3Dwell	Correlation Coefficient	.407	1.000	**.700****
		Sig. (2-tailed)	.132	.	**.004**
	Screen4Dwell	Correlation Coefficient	**.679****	**.700****	1.000
		Sig. (2-tailed)	**.005**	**.004**	.
Spearman's rho (Gamers)	Screen2Dwell	Correlation Coefficient	1.000	.232	.061
		Sig. (2-tailed)	.	.405	.830
	Screen3Dwell	Correlation Coefficient	.232	1.000	-.064
		Sig. (2-tailed)	.405	.	.820
	Screen4Dwell	Correlation Coefficient	.061	-.064	1.000
		Sig. (2-tailed)	.830	.820	.

**. Correlation is significant at the 0.01 level (2-tailed).

Correlation of screen percentage use to dwell times, however, shows a striking difference; the Non-Gamers had no significant correlations, either weak or strong, within the high workload level flights while the Gamers still showed both strong and weak correlations between a specific screen usage and its respective dwell times (table 4.22). This indicates that the Gamer group display a definitive, if weakening, process for gathering information; the weakening is likely due to increased workload which provides evidence that as the amount of information required to be processed increases this adversely effects the processing pattern. This effect can be definitively seen with the Non-Gamers as they now show no correlations whatsoever so any data acquisition patterns that may have previously existed have now completely degraded; the Gamers' patterns appear to be much more robust. Relating this back to table 4.21 it could be postulated that whilst the Non-Gamers' developing information acquisition strategy is a global case not associated with information type the Gamers' developing, or existing, strategy is more finely linked to information type.

Table 4.22 - Screen data correlation for Gamers and Non-Gamers

	Gamer		**Non-Gamer**	
Correlation	ρ	significance	ρ	significance
<u>Sc2P-Sc2Dwell</u>	<u>0.579</u>	<u>0.024</u>	0.429	0.111
Sc3P-Sc3Dwell	**0.671**	**0.006**	0.146	0.603
Sc4P-Sc4Dwell	**0.682**	**0.005**	0.464	0.081
<u>Sc2P-Sc4Dwell</u>	<u>-0.529</u>	<u>0.043</u>	-0.268	0.334
Sc4P-Sc2Dwell	-0.321	0.243	-0.129	0.648

Correlations: Bold text – **Strong**, Underlined – <u>Weak</u>, Normal – None

The screen percentage-dwell to average change correlation now shows some interesting results with the table available in appendix F; the Non-Gamers are showing screen change correlations (Av_...) relating to screen 3 and 4 percentage use. This is to be expected as a change in percentage use is likely to effect the amount of times that screen is looked at, and vice versa; the Gamer's, however, only show a screen change correlation to the screen 4 dwell times. This shows a potential information gathering strategy augmentation; while increased percentages of screen changes were observed both ways between screens 2 and

3, a decrease in percentage occurred both ways between screens 2 and 4. This indicates that the information gathering strategy leads to a potential separation of task related behaviour with flight data being disassociated with the virtual world (i.e. flight data available on screen 4 is no longer being equated to what happens, on-screen, in the virtual world) while relating the virtual world to the map display became much more important; this only occurred, though, with increasing flight data dwell times which could be further hypothesised as the Gamers becoming more efficient at understanding the flight data and not having to relate it to the virtual world.

Flight Data Analysis

An independent flight data MANOVA showed a significant difference between groups (F(5,24) = 2.910 & p = 0.034); as can be seen by the independent ANOVA's in table 4.20 only one of the five available flight performance measures showed no significant difference between groups, this being altitude.

With the Altitude standard deviation displaying lack of significant difference between groups, it can be suggested that even although the Non-Gamers were not as able as the Gamers to keep the platform stable, they were able to keep the platform in a similar area of operation to the Gamers; this is further backed by both vertical speed and pitch standard

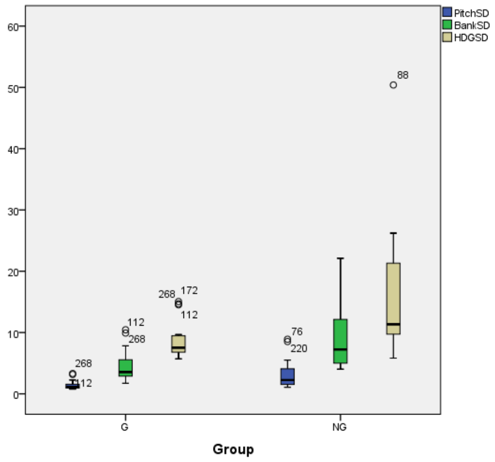


Figure 4.27 - Pitch, Bank, Heading Standard Deviation Comparison

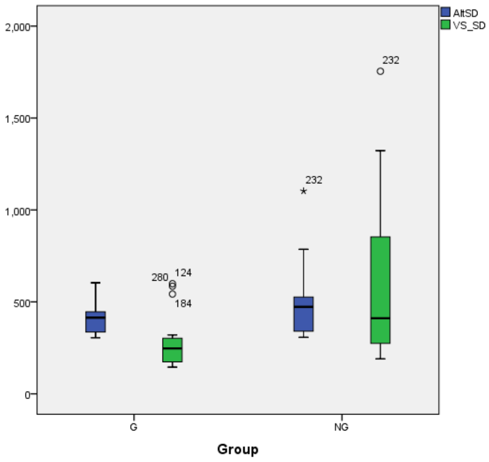


Figure 4.27 - Altitude and Vertical Speed Standard Deviation Comparison

deviation both having higher means and standard deviation (see appendix F). Descriptively this can be seen in figures 4.27 and 4.28.

Correlation identified that, in the Gaming group, multiple strong and weak correlations are present, as can be seen in table 4.23.

Table 4.23 - Flight Data Correlation (Gamers)

Correlations

			PitchSD	BankSD	AltSD	HDGSD	VS_SD
Spearman's rho	PitchSD	Correlation Coefficient	1.000	.846**	.414	.554*	.886**
		Sig. (2-tailed)	.	.000	.125	.032	.000
	BankSD	Correlation Coefficient	**.846****	1.000	.554*	.704**	.768**
		Sig. (2-tailed)	**.000**	.	.032	.003	.001
	AltSD	Correlation Coefficient	.414	.554*	1.000	.796**	.446
		Sig. (2-tailed)	.125	.032	.	.000	.095
	HDGSD	Correlation Coefficient	.554*	**.704****	**.796****	1.000	.421
		Sig. (2-tailed)	.032	**.003**	**.000**	.	.118
	VS_SD	Correlation Coefficient	**.886****	**.768****	.446	.421	1.000
		Sig. (2-tailed)	**.000**	**.001**	.095	.118	.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Unsurprisingly strong correlations exist between pitch (PitchSD) and vertical speed (VS_SD) but neither produce a correlation with altitude so even with greater deviation from the optimum in both variables the overall effect across the group on the altitude is random; this indicates different styles of flight control (high gain and low gain) both being, potentially, just as effective for this simulation.

The correlation of PitchSD and BankSD further indicates a difference of participant styles of flight control but this time a strong positive correlation exists between bank and heading deviation indicating that a high gain style of flight increases deviation from the heading. This may also be an indication of overall platform stability performance. The previous statement is validated by the very strong positive correlation between heading and altitude standard deviation indicating that if a participant has high deviation for one variable he is likely to have high deviation for the other.

Table 4.24 - Flight Data Correlation (Non-Gamers)

Correlations

			PitchSD	BankSD	AltSD	HDGSD	VS_SD
Spearman's rho	PitchSD	Correlation Coefficient	1.000	.864**	.768**	.386	.993**
		Sig. (2-tailed)	.	.000	.001	.156	.000
	BankSD	Correlation Coefficient	**.864****	1.000	.721**	.739**	.850**
		Sig. (2-tailed)	**.000**	.	.002	.002	.000
	AltSD	Correlation Coefficient	**.768****	**.721****	1.000	.311	.786**
		Sig. (2-tailed)	**.001**	**.002**	.	.260	.001
	HDGSD	Correlation Coefficient	.386	**.739****	.311	1.000	.350
		Sig. (2-tailed)	.156	**.002**	.260	.	.201
	VS_SD	Correlation Coefficient	**.993****	**.850****	**.786****	.350	1.000
		Sig. (2-tailed)	**.000**	**.000**	**.001**	.201	.

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.24 shows the same correlation previously used on the gaming group; there appear to be only strong correlations at this point with the respective weak correlations identified in the Gamer group now being strong positive correlations. These two previous tables indicate that both groups operate the platform in similar ways but with a couple of notable differences. Whereas the Gamer group had a strong correlation between heading and altitude standard deviation, the Non-Gamer group did not have this relationship with no correlation present what so ever; this suggests that Non-Gamers would focus more on one axial aspect of flight than another. Referring this to figures 4.27 and 4.28 it can be seen that altitude standard deviation did not differ significantly from the Gamers (also supported by the independent ANOVA) but the heading standard deviation was much larger (again supported by the independent ANOVA); it is likely that the Non-Gamers focused more on altitude control than heading suggesting that a tasking threshold was reached allowing only limited, and selective, use of the two overall flight parameters (heading and altitude). The Gamers however seem to have much more capacity which allows for both altitude and heading to be more effectively monitored and acted upon.

TLX Score Analysis

With the MANOVA Levene test indicating that the TLX scores are not viable for ANOVA comparison it falls to descriptive statistics to interpret the group TLX differences. The mean (**G** – 48.44, **NG** – 55.65) and standard deviation (**G** – 10.18, **NG** – 20.50) for both groups are, at this point, showing quite large differentials when compared to the previous workload where they were quite similar; the Gamers have had a very small increase in mean and a reduction in standard deviation indicating that they uniformly found the high workload only slightly more demanding. The Non-Gamers however had a significant increase in mean but, globally retained a similar standard deviation; the Non-Gamer group therefore uniformly found the high workload more taxing, while the Gamer's interpretation of difficulty increase was negligible. Figure 4.29 shows these results descriptively.

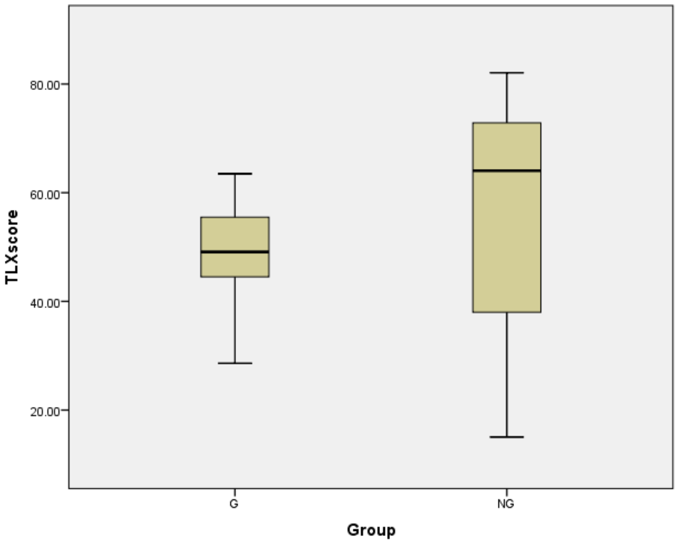


Figure 4.28 - TLX Score Comparison

Object Spotting Analysis

Object spotting score once again showed a performance differential between both groups with the Gamers (mean = 0.83, Std Dev = 0.21) proving to be more adept at this task than the Non-Gamers (mean = 0.69, Std Dev = 0.22), this can be graphically seen below (figure 4.30 & 4.31); both the Gamers and Non-Gamers still display normal and non-normal distributions respectively with the Gamer's still showing a tendency for near perfect correctness with 7 of the 15 participants with perfect scores. This continues the trend from the previous workloads of Gamers outperforming Non-Gamers.

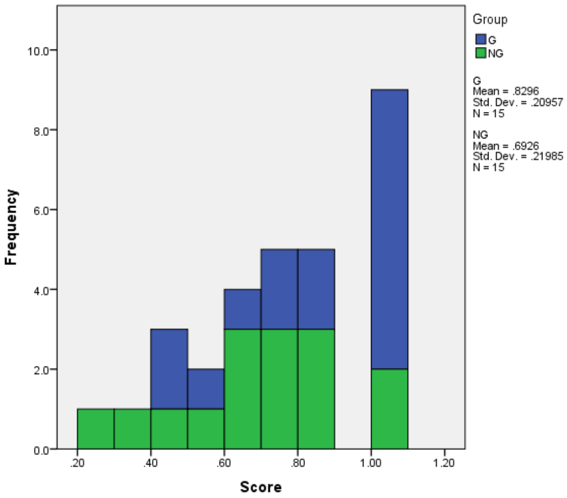


Figure 4.29 - Spotting Score G vs NG (hw)

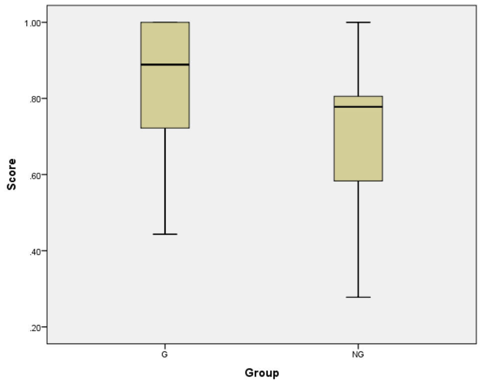


Figure 4.30 - Averaged Spotting Score Comparison (hw)

The Gamer's displayed similar characteristics in terms of mean (0.833 to 0.829) and a small improvement to standard deviation (0.24 to 0.21) suggesting the increase in workload did not adversely affect their spotting ability and accuracy. The Non-Gamers, however, showed that the workload increase had a positive effect on their ability to spot objects (mean - 0.578 to 0.693 and standard deviation 0.293 to 0.220); this suggests that the gamers, after an initial 'shock' (increase of information required for task completion initially overwhelming until information processing strategies could be formed) at the medium workload when flight control was first instigated, adapted and recovered some of their ability to correctly locate and identify objects also becoming more convergent on the group mean. This indicates learning behaviour to be present, if not in the Gamers, then certainly in the Non-Gamers.

Cross-Correlation

The correlation between flight and screen data yielded only one significant positive correlation between pitch standard deviation and screen 4 dwell times ($\rho = 0.661$, sig = 0.007), this indicates a greater amount of pitch variance for Gamers who looked at the screen longer. This could be due to a number of factors but it could be hypothesised that it is mostly likely the need to correct a large change in pitch requiring more lengthily observance of the flight instrumentation. This is supported by strong negative correlations found between screens 2 and 4 average screen changes and pitch standard deviation. With an increase in the change amount between these screens the pitch standard deviation would reduce indicating better fine control of the platform with more frequent observation. This same correlation was not observed within the Non-Gamer group, in fact the Non-Gamer group displayed no flight data to average screen change correlations what so ever showing that the process possessed by the Gamers for flight control was not used by the Non-Gamers.

Only two weak correlations appeared for the Non-Gamers with these being screen 3 percentage to heading standard deviation ($\rho = -0.546$, sig = 0.035) and screen 3 dwell to bank standard deviation ($\rho = -0.539$,sig = 0.038); although these are weak, surprisingly, it could suggest that use of screen 3 leads to improved bank and heading control likely due to increased confidence or ability to control the platform allowing more time to be allocated to the locating of objects.

Correlation of screen data to TLX score, for the Gamers, only yielded two weak correlations, once again relating to the use of screen 4. Screen 4's percentage use ($\rho = -0.604$, sig = 0.017) and dwell time ($\rho = -0.546$, sig = 0.035) could indicate that a higher use of flight instrumentation decreased the taxing effect of the high workload. It has already been shown that increased use of screen 4 (in terms of change and dwell) in the gaming group increased the amount of pitch standard deviation but did not affect altitude performance; this could indicate that increased awareness of pitch deviation decreased the participant's effective workload. It could also indicate a prioritisation of flight control which, in itself, can aid the object spotting process.

The Non-Gamers displayed two, differing, weak correlations with Screen and TLX data; these related to screen 2's percentage use ($\rho = -0.575$, sig = 0.025) and dwell times ($\rho = -0.557$, sig = 0.031); this may indicate that object spotting may have been prioritised over flight control and, with flight control able to affect spotting performance, may indicate those that were more comfortable with flight control increasingly used the virtual world to locate objects thereby lowering tasking levels.

A correlation of spotting score to screen usage only found two weak correlations which were only found in the Gamer group with screen 4 dwell times (ρ = -0.637, sig = 0.011) and average 4 to 2 screen change (ρ = 0.575, sig = 0.025); this may demonstrate that increased use of screen 4 would lead to worse spotting results suggesting that Gamers who had to make larger corrections and less frequent changes to screen 2 would not be as able to complete the spotting task as effectively.

As expected, and previously mentioned, poor flight control has an impact on ability to spot objects effectively and correctly; the two groups differed in this respect as the Gamer group's spotting score was adversely effected by pitch (ρ = -0.700, sig = 0.004), vertical speed (ρ = -0.662, sig = 0.007) and bank (ρ = -0.654, sig = 0.008) standard deviation as can be seen by their respective strong negative correlations. The Non-Gamers seemed to have their spotting ability reduced by an inability to remain on a good heading while spotting with bank (ρ = -0.659, sig = 0.007) and heading (ρ = -0.789, sig = 0.001) standard deviations showing strong negative correlations. It must be remembered at this point that, with the above correlations indicating that the Non-Gamers performed spotting better by only being affected by one axis of motion, the Gamers proved overall much better at spotting.

In trying to understand how the information increase adversely affects Gamers it was found that there was a strong positive correlation between Screen 4 dwell times and how many hours a participant played 'games' per week (ρ = 0.731, sig = 0.002). This point's to the Gamers who play more hours of games per week being more likely to dwell longer on screen 4; as this is a key data set it seems that the more experienced gamers identified this and acted accordingly. This correlation does not exist for the Non-Gamer group; they did, however, have other correlations between hours played and dwell times but, due to the limited nature of the play estimate category for the non-gamers, these correlations must be considered suspect and therefore currently disregarded.

To check if this correlation between Gamers play estimation (the period of time they estimate playing games per week) and screen 4 dwell time is a constant during lower workload flights further correlations were performed; as screen 4 was not viable during low workload flights only medium workload flights could be considered relevant. The medium workloads screen 4-to-play estimation correlation again showed a strong positive correlation (ρ = 0.682, sig = 0.007); this corroborates the author's speculation that this is part of a pattern of data acquisition behaviour and that more experienced Gamers (experienced by recent play amounts) would display a higher tendency to dwell on screen 4.

With recent gaming experience in mind, and with the previous correlation results indicating that large amounts of recent game activity being a factor, the correlative process was continued with flight data. Only one weak correlation was found (PitchSD-Play_Est, ρ = 0.526, sig = 0.044) at this point indicating that recent play amount did not affect flight performance suggesting that, although there seems to be difference in data acquisition strategies associated with recent gaming experience, recent gaming experience does not significantly influence flight performance.

Correlating again play estimate with spotting score and TLX score revealed no correlations; this indicates that screen 4 dwell times were not influenced by stress levels and had no effect on ability to correctly locate and identify objects. Overall this suggests that the screen 4 to dwell time correlation is not an indication of potential performance but is truly only a data analysis pattern influenced by amount of recent gaming.

Investigating screen 3 dwell times and percentage use yielded no correlation within the gaming group when compared to play estimate; looking at this data descriptively the percentage and dwell time changes between medium and high workloads (m = 0.09 to 0.084, std dev = 0.012 to 0.004) was extremely small with only a larger change occurring between low and medium workloads (m = 0.16 to 0.09, std dev = 0.015 to 0.012). The reducing standard deviation indicates a convergence of data analysis patterns relating to this screen further supported by an extremely small standard deviation for screen 3 at the high level flights.

No correlations were found between flight data–TLX score or TLX-Spotting score

Chapter 5 Discussion

This chapter provides a detailed interpretation of the results and assumptions derived from the previous statistical analysis chapter and from insight into information investigated during the literature review. This interpretation is carried out by investigating the results in relation to the original research questions and objectives of this thesis.

5.1 Research Summary

The aim of this research is to investigate the differences and similarities between Gamers and Non-Gamers when using an RPAS type Human-Machine interface and whether potential differences could be used to objectively measure performance of a pilot during flight. Further to this, the experiment was divided into three workload levels (low, medium and high) to identify whether these similarities and differences existed with increasing stress and activity levels and whether potential information set interaction trends (patterns associated with inter and intra screen usage) persisted or disappeared between these workload levels.

Ideally, the aim is to be able to objectively measure performance using only information set (screen) usage data while using flight, TLX and spotting performance scores as a correlation to the objective screen performance measurement. Both groups underwent identical experimental flights in exactly the same order to eliminate any discrepancy with learning profiles and workload levels; it would be unfair to expect a Non-Gamer to perform as well as a Gamer if they both started the experiment with the high workload level.

By having this structure within the experiment, data analysis patterns were expected to evolve and change within both groups and these patterns were expected to be evident and explainable in the statistical analysis.

5.2 Aim 1 - To understand task based analysis of performance and apply it to semi-automated performance measurement systems

The first aim of this research was predominantly an investigation into the use of task based analysis for performance measurement and to understand the kind of system (hardware and software) required to achieve data collection and analysis of the resulting information. Below the objectives, contained within this aim, are discussed in terms of their findings.

5.2.1 Objective 1 - To identify the literature available in the RPAS and task analysis domain applied to platform classification and the training of operators

Research associated with RPAS, although not a new field, is in some ways still in it's infancy in terms of global doctrines of operation and training; this can be seen by the lack of a global standardised licensing and training systems in place for either commercial or military platforms and operators. Military training appeared, up until recently, to be on an ad-hoc basis by utilising existing manned aircraft pilots and re-training them to use RPAS.

With this unformalised structure the literature associated with training of RPAS operators, and there associated performance measurement, was scant. This highlights the need for further research to be carried out within this field to globally define RPAS classification, use, training and performance measurement.

With the above in mind it was possible to create a potential task analysis performance measurement system, based around the Mission Essential Competency; this development is further discussed in the following objectives.

5.2.2 Objective 2 - Create a potential task analysis system based around RPAS operation and based upon current Air Force doctrine based upon the MEC system

Initial research showed the MEC system to be incompatible, in terms of structure, with a simple and usable performance measurement system; this created the need to restructure the MEC from the fundamental level upwards. With this being the case the decision was made to focus purely on the unmanned training aspect as this allowed a larger leeway in terms of measurements and structures as well as simplifying experimentation by reducing task and KSE lists and allowing for creation of new and specific UAV related MEC's, SC's and KSE.

The SC has, theoretically within this thesis, been redefined as a human factor related competency while the MEC itself remains a time and mission specific competency; both the SC and the MEC sit along the same tier of importance with the MEC relating purely to mission performance (figure 5.1) and the SC relating to operator development within a specific area of flight competence/human factor (communications or adaptability for example). As this is no longer a trademarked MEC approach it is of note that this revision in no way reflects on the original MEC system. The MEC and SC elements of this research are to be renamed Mission Competency and Human Factor Competency; it should be emphasised that this research is in no way associated with current MEC systems.

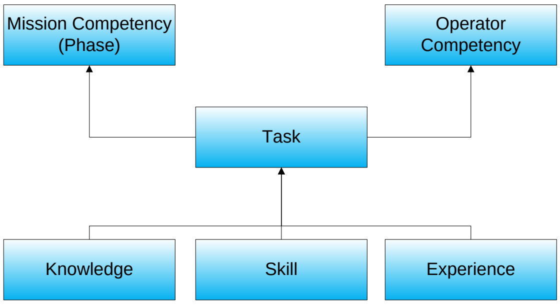


Figure 5.1 - Proposed, revised MEC Structure

This restructuring allows for a more direct measurement of the basic KSE's/tasks and translation into either mission specific or operator specific performance. Due to the potential crossover of potential KSE's they have been incorporated into a singular task.

From a training perspective the tasks identified during the case study (see Chapter 3.1) could well be manipulated into certain categories, which allow identification of a training timeline (see figure 5.2).

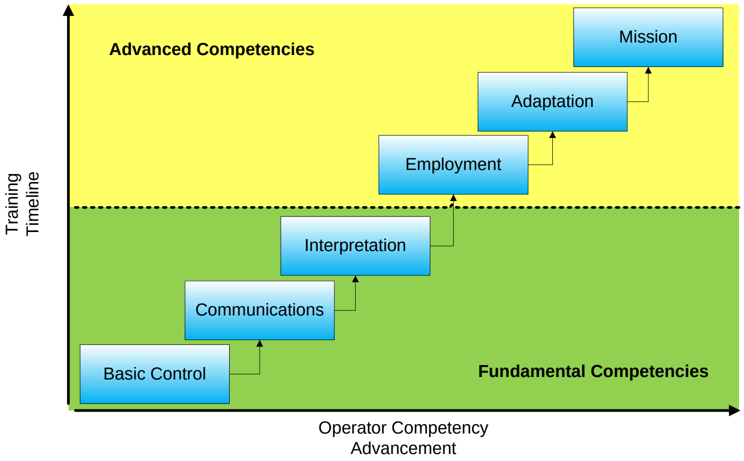


Figure 5.2 - Operator Competencies in a training timeline

Since receiving clarification of the MEC system from DSTL (Defence Science and Technology Laboratory) it appears that the experiences, in fact, inform the knowledge and skills and are not separate to it. Experiences can be used to create the knowledge and skills required for operation and can then be directly linked to a specific competency with the supporting competencies not a necessary part of the structure.

Further to the basic restructure was the investigation into the tasks used within a mission environment and the way in which they can be interpreted with regard to the human/machine interface as well as existing psychological and logical models for an operators thought/decision process. Initial analysis suggested that a basic task could be evaluated in four different, time related categories: Acquire, Analyse, Decide, and Act (AADA).

This concept was arrived at independently but, on further research, substantiation was found in the OODA approach (Observe, Orient, Decide and Act). [66]

On further inspection of the list comprising a generic MEC, and with the previous categorisation and adaptation research in mind, it was possible to create theoretical tables better categorising the tasks associated with a generic mission outlined by the example MEC [37] (see Appendix C).

Evidence based training

By observing a pilots cockpit operations in a task based manner it is possible to apply empirical data to pilot performance in terms of physical as well as, potentially, decision based metrics. Using Boyd's observe, orient, decide and act (OODA) loop it is possible to map an operators decision making process; much research has been aimed at the ‘decision’ and ‘action’ phases in terms of performance measurement. The ‘decision’ and ‘action’ analysis of RPAS flight may not be sufficient to create a good performance metric for certain tasks; using a case study (detailed in chapter 3.3) the researcher created task lists based upon observing and questioning a Subject Matter Expert (in this case a Reaper pilot) during the course of a generic surveillance mission.

There appeared to be tasks carried out by the pilot, which did not require an action and which, without a transcript of the flight, would not inform the pilots decision basis. Prior to the experimentation the Mission Essential Competency KSE and Supporting Competency lists contained within the example MEC (see appendix C were organised and in some cases expanded to allow direct interaction between KSE's and Supporting Competencies, this allowed for examination from a human factors approach.

It was realised that the KSE's may occasionally be task specific as well as being generic to multiple tasks, this allowed a postulate that each knowledge or skill could be associated to a greater or lesser degree to a specific task and would be possible with weighting to measure the performance of a particular skill or knowledge. However, research limitations have not allowed for further minute investigation of these relationships and it was decided to focus purely on the task itself rather than its continuant human factor elements.

Below is a table (table 5.1) showing basic tasks and human factors observed during the course of a basic reconnaissance flight, the subject matter expert subsequently verified these tasks.

Table 5.1 - The task and human factor lists acquired during the case study

Tasks	Human Factors
Pre-flight Planning	Plan
Configuration	Control
Take Off Control	Time Management
Autopilot Management	Interpretation
Take off Monitor	Adaptation
Peripheral Awareness	-
Object recognition	-
Re-Plan	-
Re-Task	-
Waypoint Management	-
Expedite	-
Approach Control	-
Landing Control	-

It could be interpreted that these tasks represent purely a time (phase) based approach to task analysis, this is not wholly the case as, although these tasks all have start and finish points, they may also run concurrently with each other; this is especially the case with the monitor task. The monitor task runs almost continuously as either a major or a sub task during a flight; it is understandably absent during the mission phases that do not require the platform to be monitored.

This, however, does not necessarily mean that two tasks have to be measured simultaneously they overlap each other; it could instead to be incorporated into a task where the monitor task will always be present, such as peripheral awareness and re-planning. The fact that two tasks are being measured simultaneously should not make a difference to overall performance measurement unless the relationship between these two tasks is dynamic (often to change) rather than static.

Another important upshot of reviewing tasks in this way is that they may aid with creation of a modular based cross-platform performance measurement system (see chapter 6.2.2).

As tasks have now been isolated it is important to be able to understand task performance measurement when applied to an RPAS; much research has been carried out into pilot effectiveness based upon both objective and subjective measures, mostly using basic simulator data as well as pilot interviews. This research has been supplemented by a very large amount of study regarding pilot psychology and instrument interaction while carrying out flight actions, often using head and eye- tracking solutions [67].

It may, therefore, by plausible to consider the way in which a pilot utilises available information as a performance measure in itself. Boyd's OODA loop (Observe, Orient, Decide and Act – see figure 5.3) [66] has great relevance at this point and will be discussed further within the next section.

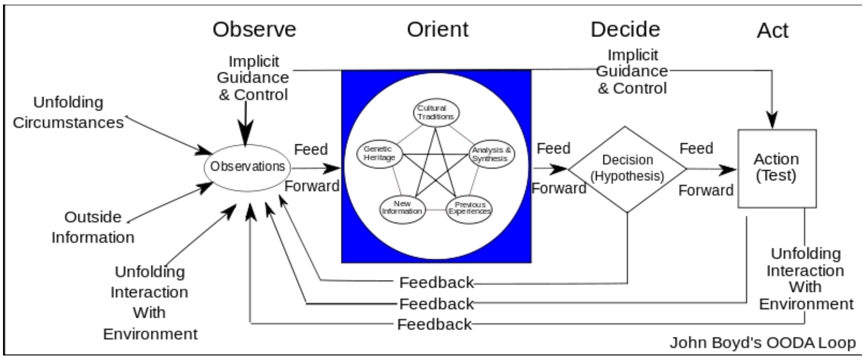


Figure 5.3 - Boyd's OODA (Observe Orient Decide Act) Loop - http://en.wikipedia.org/wiki/File:OODA.Boyd.svg

5.2.3 Objective 3 - To define the components of a semi-automated performance measurement system in order to create a demonstrator/simulator

Modular Task Based System

The concept surrounding this section pertains to the use of the previously identified tasks as an aid to, not only performance measurement, but also as a potential aid to future RPAS licensing; currently there is no supporting documentation relating to this concept but, as a modular task system can be applied to part of this research, the potential of this concept should be explored.

Looking back at previous sections within this chapter the concept of using task-based performance measurement has already been explored but how can this be applied to a real world issue?

In sections 2.1.1 & 2.1.2 the issue of RPAS licensing has already been touched on with there still being a lack of definitive global structure to RPAS pilot licensing. Once the revised task analysis structure and been created it was seen that several elements of RPAS operation could be generic to almost all RPAS systems (such as a Spotting or Monitor task); this, from a performance measurement point of view is not the case and these seemingly generic elements were, in fact, linked directly to the platform being used. But what would happen, as the STANAGs [21 & 22] indicate, if generic interfaces, data transfers and levels of autonomy became legislated and enforced?

It could be seen that the previous tasks could become, indeed, generic; it can also be extrapolated that other tasks present in RPAS control (such as take-off and landing) could also become linked to a generic type of HMI. Multiple RPASs could become operated with a standardised control interface and also have standardised operations; this would not be dissimilar from the varying types of automotive available on the market today in which they all have different objectives within their design but also offer a generic control interface in the form of steering wheel, accelerative and declarative pedals and either automatic or manual gear selection. As can be seen within the automotive market you do not require a license to operate each individual vehicle upon the market but will require differing licenses when there is a large difference in control interfaces or platform operations; often though if a specific vehicle function is required then a supplement for the basic license can be trained for (for example a trailer/towing supplement, Light Goods Vehicle or Heavy Good Vehicle) and it won't necessarily require a complete retraining of the operator.

This could indeed happen within both military and commercial RPAS communities, an example would be the development of a generic HMI for both BAE Systems Mantis and General Atomics Predator. Both have similar battlefield roles so it is not unfeasible that a generic interface could be created to operate both platforms in terms of the command and control element; each platform though will have some specific demands on training which can be seen as platform specific training. Figure 5.4 below is a representation of two differing platforms and the way in which tasks can sometimes be generic, varied or specific.

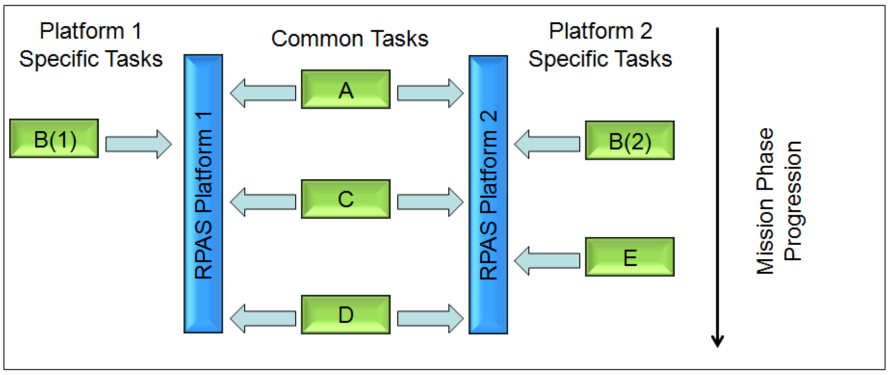


Figure 5.4 - Platform task comparison: A, C, D being common unrevised tasks, B(1) & B(2) being common but revised tasks, E is platform specific

With potential common tasks between multiple RPAS platforms identified it could be possible to create a basic license associated with a specific type of operational criteria and GCS rather than based on the type of RPAS platform in operation. Subsequent to this basic training, platform specific training (either supplied by a licensing agency or by the platform manufacturers) could then enhance the operator’s ability on a specific platform. This type of basic and specific task training can be associated with basic tasks being related to generic competencies and specific tasks being related to advanced competencies.

Semi-Automated Performance Measurement

With a remotely piloted platform the performance measurement process becomes much more simplified, as opposed to a manned aircraft, as environmental factors associated with live manned flight are no longer relevant.

Much of the basic aircraft control is now performed by the platform so the operator’s role is now less about direct control but much more focused on higher-level tasks as well as monitoring of the system.

The restructured MEC allows for information regarding the operator's higher level processes to be segregated and then analysed using optimal information/data usage; using the previously mentioned OODA loop [68] task analysis approach it should be possible to measure the performance of an operator based on the information/data acquired at the beginning of a task and then relate that to the final output of the task whether this is a physical, situational or null event.

It was noted that the analysis and decision phases of the process would be impossible to performance measure without some form of verbal communication and, for this reason the orient phase of the OODA approach has been integrated into the observation phase and renamed the analysis phase. This was only relevant to a full flight in which the tasks to be carried out were an unknown element and had to be identified post hoc.

Although there may be no optimum thought process for each task (with the thought process being partially operator specific and also very subjective) it should still be feasible to assume that an operator who observes the correct information, for roughly the desired amount of time and then completes the task successfully, in terms of a decision or action, should be given an appropriate score for good task completion. At this point it was still thought possible to measure the knowledge and skills of the pilot this however was decided to be no longer the case as multiple, concurrent knowledge and skills may be available within the task; the task itself is now to be performance measured instead of its sub-elements.

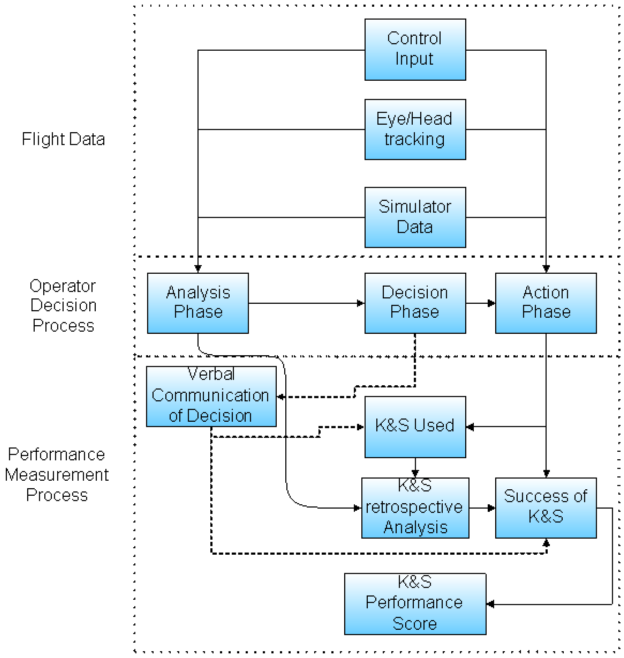


Figure 5.5 - Proposed task performance measurement process utilising Knowledge and Skills, now applied to tasks

This, of course, has its flaws; an operator could theoretically complete a task successfully without having reached the correct conclusion from the correctly observed data and this shows that an after flight debrief, in which the operator divulges his thought processes, will likely always be required to make sure of the operators competence level. What can also be observed from figure 5.5 is that there is an element of the OODA loop (see figure 5.3) contained within the above process. The operator decision process can be seen to be the *ODA aspect of the OODA loop.

An automated system, however, can still be utilised for training purposes as a tool to provide progress monitoring, flight control measurement, and objective measurement of information/data usage. This can then be used by the instructor to support his training/performance assessment.

As the research has evolved the emphasis has been taken away from creating a semi-automated performance measurement system with more of a focus on the previously mentioned research objectives; the research will still utilise these concepts to aid experimental procedure and analysis but will no longer be proposed as a current training support tool. This is not to say there is no requirement for a semi- automated performance measurement tool; the Air Force Research Laboratory (AFRL) have developed their own network based pilot performance measurement tool called PETS (Performance Effectiveness/Evaluation Tracking System) [69]

Developmental performance measurement research at the Air Force Research Laboratory in Mesa, AZ resulted in this:

> “Performance Effectiveness/Evaluation Tracking System” (PETS). PETS is a software tool that enables multi- platform, multi-level measurement ability at the individual and team level in a complex Distributed Interactive Simulation/High Level Architecture (DIS/HLA) environment. Installed at the Mesa research site, up to one million data points per minute are collected and organized into several formats differing in unit of analysis" [69: p.2]

The case study (see section 3.1) showed the need for a semi-automated tool to aid performance measurement; due to time constraints on data analysis and the length of time required to analyse the output data it was decided to create a performance evaluation tool to shorten these analysis times. A large problem associated with the analysis was the verbal communication sync to the flight and head tracking data, as well as the processing of this large amount of data; to this end a software solution was created to aid this analysis and is further detailed within section 4.2. This system is not based upon PETS but is a more basic tool that helps to amalgamate there various data sources being targeted by the experimentation.

5.2.4 Objective 4 - Creation of a simulator, with bespoke data acquisition software, based upon a semi-automated performance measurement system

As detailed within chapter 3, a simulator was created to allow for experimental investigation of operator performance based upon the use of 'information sets' with head tracking as well as a correlation to object spotting performance and flight data. This simulator was subsequently modified, in terms of its software, following the case study.

The details of the case study aims, objectives and process can be found in section 3.1; the findings of the case study, including problem identification, can be found in section 4.1.

With the changes to the simulation software and the creation of the SimPACT software the simulator proved effective at capturing relevant data and resulted in the main experimental analysis found in section 4.4.

5.2.5 Objective 5 - To create and test software to acquire and analyse data received from the simulation

Problems with conventional data gathering and analysis software (originally used within the case study) were identified and required a bespoke solution. This solution came in the form of the SimPACT software, created solely by the author, to specifically capture data from the re-designed simulator (as outlined in objective 4).

This software proved very effective at not only capturing the required data but also creating a tool to swiftly process the data into a format which was easily transferrable to statistical analysis software. It also provided further data sets which would otherwise have been very difficult to capture using the case study setup.

This software should be reusable for other similar applications and has been attached to the thesis.

5.3 Aim 2 - To demonstrate that decision-making ability and data processing capability is enhanced by experience of interactions with computer game based environments

The second aim was to demonstrate that both decision making ability and data processing capability are enhanced by sustained experience of computer game environments. The following discussion relating to the objectives contained within aim 2 demonstrate that 'gaming' experience was a factor in improved cognitive ability with distinct patterns of information collection identified within the gaming group which did not exist within the non-gaming group.

5.3.1 Objective 6: To identify whether information set usage can be considered a viable indicator of performance when applied to a task based system

The MANOVA results from all workloads show no significant differences between Gamer and Non-Gamer groups for percentage screen use and dwell times (table 5.2); this indicates that, at this basic data (non-correlated) that both groups used the screens, in terms of percentage, dwell and, sometimes, average screen change percentage in a similar fashion.

Table 5.2 – Available screen usage data MANOVA comparison results

Workload	F	p
Low	0.384	0.854
Medium	1.051	0.420
High	0.883	0.572

With this being the case a performance measurement system, based solely upon percentage use and dwell times, would not be an effective indicator of objective performance. When considering the flight, TLX and object spotting data with this similarity in screen usage it can be seen that the Gamers did, in fact, perform better than the Non-Gamers further proving that the variables, when considered individually, would not provide an explanation for the other data types performance differential. Referring this to Boyd's OODA loop it can be interpreted that, from this base screen data, the Observe phase of the loop is similar for both Gamers and Non-Gamers; this suggests that the factor effecting performance originates more in the Orient part of the loop. This hypothesis is further explored in aim 2 in which any patterns are identified and analysed.

In conclusion objective performance measurement of participants is not viable solely using percentage screen use and dwell times.

5.3.2 Objective 7: To identify key factors in stakeholder selection for experimentation via creation and application of a questionnaire

Before experimentation could begin the participants had to be quantised based upon their age and previous gaming experience. To this end the designed questionnaire was able to satisfactorily identify the key requirements of both gaming and non-gaming groups in terms of age and gaming experience (current and past) with additional, although non-vital, information regarding education and employment.

The questionnaire gave a result of a slightly higher mean age (calculated with accurate ages rather than age band mean) for the non-gaming group (25.3 years – G & 27.17 – NG) with the Non-Gaming group having a distribution leaning more towards the 31-35 year old band (26.7% of participants) compared to the Gamer group (6.7% of participants).

This shows that there was a slight predominance towards older participants for the main experiment but not large enough to be able to make a significant impact on the experimental results.

The level of gaming experience, expounded upon in section 4.4, is not in itself a result but was used to determine the criteria for participant selection; given the results identified within section 4 this seems to have been successful as key differences between the groups have been identified.

5.3.3 Objective 8: To identify if a more experienced gamer manage stepped increases in workload more effectively and accurately than a non-gamer with respect to flight stability and object spotting and identification

Throughout each of the workload levels the Gamers performed to a higher level c (see section 4.4) in terms of flight stability and object spotting, as can be seen in each workloads analysis in the previous chapter.

In terms of object spotting the Gamers, from low to medium workload had a reduction in spotting mean of only 7% while the Non-Gamers reduced by 21.2%; this indicates that the increase in workload and function heavily affected the Non-Gamers ability to correctly spot and identify objects. The medium to high transition showed a different trend with the Gamers mean spotting score remaining approximately the same and the Non-Gamers score increasing by 12%.

This indicates that the Non-Gamers suffered an initial 'shock' when transferring from low to medium workload; this is due to flight control being absent within the first workload and the introduction of flight control in the medium workload causing a type of 'information overload'. Once this initial shock of information increase has been experienced the Non-Gamers found the addition of an extra flight variable, rather than the addition of a completely new task, to be much more manageable. The object spotting data suggests that the Gamers were not as adversely affected by the addition of a new task and the further addition of a new flight variable in fact made very little difference to their spotting performance.

The NASA TLX data does not indicate that there is a large difference between either group at the low and medium workloads so it can be hypothesised that both groups experienced the initial shock of workload increase but the Gamers were much more able to handle the increase than the Non-Gamers. It is only at the medium to high workload change that a significant difference between the groups can descriptively be observed (see sections 4.4.2 & 4.4.3); the Gamers had a much lower mean and standard deviation at the higher workload than the Non-Gamers at the high workload level with only a slight increase in mean for the

Gamers and a significant reduction in standard deviation. The Non-Gamers however showed a larger increase in mean and no reduction in standard deviation.

From this it can be seen that the addition of a flight variable, rather than a task, affected the Gamers very little in terms of their respective workloads and, in fact, showed them to have much more uniform workload levels; this is supported by the lack of change to their spotting ability. The Non-Gamers, however, were significantly affected by all workload increases with increasing TLX scores observed throughout all of the workload levels; the Non-Gamers, after the initial low to medium 'shock' did improve their spotting performance but the addition of an extra flight variable did also increase their subjective TLX workload levels.

In this context 'shock' can be defined as an overload of information and/or need for extra user input that only initially affects the participant's ability to perform the required tasks.

Only two flight variables were available for comparison, and only between the medium and high workloads. The Gamers were able to perform better across the workloads than the Non-Gamers in terms of both bank and heading (see table 5.3).

Table 5.3 - Group flight data comparison (all values are in degrees)

Workload	**Variable**	**Gamers**		**Non-Gamers**	
		Mean	sd	mean	sd
Medium	BankSD	5.03	3.56	8.64	5.67
	HeadingSD	8.92	2.15	13.45	8.61
High	BankSD	4.65	2.739	9.14	5.36
	HeadingSD	8.90	3.23	16.76	11.21

It can be seen that the gamers either remained similar across the workloads or improved in their ability to keep the platform stable while the Non-Gamers decreased in competence further indicating that the Gamers were not affected by the addition flight variable while the Non-Gamers did display reduced performance with the additional variable. Inter-flight data correlations show that while both groups show positive correlations (both weak and strong) between almost all variables at the high workload indicating that poor control in one axis of motion would negatively affect control in another. The Gamers however show no correlation between PitchSD and AltSD ($\rho = 0.414$, sig = 0.125) indicating that even at higher pitch levels the whole group was able to keep the altitude variation much more stable which suggests better overall performance as well as the emergence of types of operation (high gain and low gain).

With the cross-correlations showing that flight stability is linked to ability to spot objects it is slightly surprising to find that the Non-Gamers spotting performance increased with the additional flight variable at the high workload. This suggests that the Non-Gamers were slowly adapting to using an unstable platform while spotting objects effectively.

In summary, an experienced gamer would manage increasing workloads better than a non-gamer and would only show large increases in subjective workloads in relation to the addition of new tasks rather than the addition of a variable; similarly to the subjective workload increase, the performance for a gamer would not adversely be affected by the addition of a variable but would be mildly affected by the addition of a task. A non-gamer will experience workload increase with the addition of both tasks and variables but the addition of a variable, rather than a task, would not adversely affect their performance. The addition of a task would majorly affect a non-gamers performance.

5.3.4 Objective 9: To identify if an experienced gamer's information set usage and data acquisition strategy differ from that of a non-gamer

As has already been discussed in 6.2.1 both groups have similar usage of information in terms of percentage use, dwell times and average screen change percentage; this is not the end of the story though.

At the low workload level both groups were found to have strong positive correlation between screen 2 use and respective dwell time but only the Gamer group then displayed a strong positive correlation between screen 3 use and respective dwell time; initially this suggested that the Gamers displayed a definable information gathering strategy with an increase in screen use overall meaning an increased dwell time. Hypothesising this could mean that the Gamers used a data prioritisation strategy in which an information set that required more overall observation would also need a greater amount of dwell time to take in the required information.

This trend was still evident at the medium workload (see table 4.16) with all three available screens and their respective dwell times showing strong positive correlation, a further strong negative correlation was found between screen 2 and 4’s use and dwell times showing that the longer a Gamer spent looking at screen 2, the shorter his dwell time would be on screen 4 again suggesting a data prioritisation pattern. The Non-Gamers only displayed one strong correlation at this level between screen 2 and its dwell time; two weak correlations were also found which may suggests that the Non-Gamers were also trying to prioritise information but did not use the same subconscious pattern that the Gamers do.

This pattern was still evident, although degrading, for the Gamers at the high workload level (see table 4.22) with the screen 2 use and dwell time now becoming a weak correlation; the two other inter-screen correlations though have remained strong and positive. With all three workloads showing this pattern it can be concluded that the Gamers do, in fact, have a definable information set use pattern whereas the Non-Gamers, at the high workload level, have no correlations what so ever showing that they have no definable information gathering pattern.

This is likely due to Gamers having predefined processes for partitioning data sets and gathering relevant information by being able to prioritise information sets. This pattern can be potentially linked to flight performance, it was observed at the high workload level that the Gamers had a strong correlation between screen 4 dwell times and increase in flight instability; it is likely that a Gamer, knowing he lacks proficiency at flight control, would spend longer looking at screen 4 to help compensate for the lower degree of control thereby also displaying the same information prioritisation strategy hypothesised during this section.

5.3.5 Objective 10: To identify inter-group similarities with information set usage as well as similarities with degradation of performance with increasing workload

As already investigated in 6.2.3 the Gamer group display strong information prioritisation strategies, whereas the Non-Gamer group display non-existent to mild prioritisation tendencies that disappear completely under high workload. This suggests that the Gamers have similar and much more robust strategies for data acquisition while at a macro level do not appear dissimilar from the Non-Gamers.

Flight data shows that, while Gamers overall had a slight reduction in flight performance due to increasing workload, the Non-Gamers had a much increased reduction in flight performance; this was not, however, reflected in object spotting performance as the Gamers, after an initial reduction due to the addition of the flight task, remained relatively stable in terms of performance.

The Non-Gamers also had a large initial reduction in flight performance due to the addition of the flight task, likely due to increased information 'shock' but then recovered some of this performance in the high level workload where only an axis of motion of flight was added. This demonstrates both groups learning and adaptation behaviour; the Gamers existing data acquisition strategies could be identified to be similar throughout while the Non-Gamers showed no definitive data acquisition strategy.

5.3.6 Objective 11: To identify if increased workload affects information set usage in both groups

At the macro level (just dwell and percentage use) both groups show (table 5.4) that information set usage is affected by workload increase. The largest initial change is understandably the low to medium workload level as the available information sets change from two to three. The reduction in screen 2 use is due to the need to observe flight information on screen 4, this leads to a large reduction in percentage use of screen 2 as well as a small reduction in use of screen 3; following this initial reduction the percentage use of screen 2 remains similar and constant for both groups between the medium and high transition indicating that this percentage use is acceptable when only a single variable is added. Any reduction for screen 2 at this point is only small (max of 1.4% change with the Non-Gamers); it can be seen though that the use of screen 4 does increase between medium and high workloads (7.4% & 9.2%), this ties in with the small reduction in screen 2 use and a larger reduction in screen 3 use. This shows that, at the high workload, the flight data did indeed need to be observed for longer amounts of time due to the variable increase with a lower priority being assigned to the map display. It is likely that, at this point, both groups, through the learning process, have become more adept at the use of screen 3 and feel more able to allocate time from that screen rather than screen 2.

Table 5.4 - Workload/Group percentage use comparison

Variable	**Workload Level**	**Gamer**	**Non-Gamer**
Screen 2	Low	0.707	0.710
	Medium	0.433	0.418
	High	0.430	0.404
Screen 3	Low	0.158	0.146
	Medium	0.090	0.087
	High	0.084	0.074
Screen 4	Medium	0.337	0.327
	High	0.411	0.419

As the increase in screen 4 usage does not correlate to the differential of screen 2 and screen 3 use it’s likely that the extra screen 4 usage is derived from more accurate transitions between each information set resulting in, overall, lower times spent in a void region (a region of head position not associated with any screen. Both groups have again displayed this tendency with increasing workload (table 5.5).

Table 5.5 - Void Percentage Comparison

Workload Level	Gamer	Non-Gamer
Low	0.096	0.103
Medium	0.112	0.147
High	0.074	0.090

As can be seen in table 5.5 above there is an initial increase in percentage void time between low and medium workloads, likely due to the added need to transition between three screens rather than just two and the associated need to learn best information acquisition pathways. As has been mentioned previously void time is reduced by the higher workload when information prioritisation is required to a much greater extent, this explains where the increase in percentage use of screen 4 is obtained from. The Gamers perform marginally better in that they have less percentage time spent in areas with no information but this difference is only marginal and so was not considered during statistical analysis.

When considering screen usage data correlation it can be seen that both groups are affected in different ways; as previously mentioned the Gamers have a more robust data acquisition strategy which is only affected to a small degree, with the degradation of some information set usage correlation values from strong to weak, by the increased workload with only small losses in screen-to-dwell correlations. The Non-Gamers do not have a robust strategy though and any previous correlations between percentage usage and dwell times no longer exist at the high workload level. This indicates any existing strategy apparent at the previous two workloads is not viable at the high workload levels. This supports the conclusion that large amounts of gaming activity results in robust data acquisition and management strategies that cannot easily be obtained through normal day to day life or learned via mild (less than three hours per week) experience of computer gaming.

Chapter 6 Conclusions

This chapter provides a summary and conclusion to the main research findings. It also outlines the key contributions to the RPAS field and suggests future research that would further increase RPAS knowledge and understanding.

6.1 Conclusion

There are two overriding aims to this research; firstly to identify if it is possible to adequately performance measure participant's task performance while undergoing a task based RPAS flight simulation. The second aim is to identify whether a Gamer would be a much more viable candidate for RPAS training than a Non-Gamer, through supplementary use of graduated increasing workloads, and applying to the way in which a potential candidate can gather, analyse information and then apply to a relevant task.

6.1.1 Findings 1

At the macro, or individual variable analysis, level it was concluded that it was not possible to identify any differences between the groups (see sections 4.4.1 to 4.4.3 as well as chapter 5); this indicates that, while using a head-tracker and partitioned information sets, that it is not possible to be able to performance measure both groups using information sets as an objective performance indicator. Both groups showed remarkable similarity in their percentage use and dwell times for each information set (screen) and this was confirmed with multiple MANOVA tests at each workload level. With Boyd's OODA loop in mind it can be inferred that the Observe phase of the loop is similar for both groups.

6.1.2 Findings 2

When analysing the screen data with correlation techniques (chapters 4.4.1 to 4.4.3), robust patterns began to emerge within the Gamer group that did not exist within the Non-Gamer group; these patterns show that a Gamers Observe phase is, in fact, influenced by a data prioritisation strategy in which a Gamer can be observed to increase dwell time on an information set that they comparatively use more than their compatriots. With these strong correlations being absent within the Non-Gamer group it can be concluded that Gamers, through experience of managing high quantities of disparate data, have globally developed similar strategies for effectively managing and utilising that available information. The objective performance data (both flight and object spotting) support this conclusion; the Gamers proved much more effective at completing both flight and spotting tasks to a much higher standard.

Although the objective performance differential between both groups was expected it is exciting to find a correlation between this performance differential and the way in which they utilise information sets.

6.2 Contributions

This thesis investigates the potential for objective performance measurement with the use of information sets and the potential for Gamers being more viable candidates for RPAS training than Non-Gamers. The contribution of this research was to understand if the above two aims were viable.

6.2.1 Patterns of Information Set Usage

Although it was not possible to distinguish between the Gaming and Non-Gaming groups at a basic information usage level, it was possible to identify very strong/robust data acquisition patterns (see sections in 5.2 – Cross-Correlations) within the Gamer group that were indistinct to non-existent in the Non-Gamer group. Research conducted for this thesis has clearly shown that Gamers have developed significant information gathering, monitoring and application processes that an average person does not possess (i.e. A Non-Gamer or someone who plays less than three hours of games per week).

The Gamer participant pool contained few participants who had any form of simulator experience and even fewer with flight simulator experience; it can be concluded that, although many of the Gamers had no experience of this type of 'Game', they still performed well and displayed the same information observation, orientation and application traits no matter what type or style of computer game they had experience of.

Research has shown that a 'Gamer' would make a much more viable candidate for RPAS training than a Non-Gamer and would almost certainly lead to reduced training times, cost and increased efficiency due to the pre-learned data acquisition strategies aiding them in high workload and taxing training situations as well as faster adaptation to new scenarios, protocols and interfaces. As RPAS do not require the same level of fitness or physicality (physical ability) as a manned aircraft this would allow a much larger candidate pool for RPAS training to exist; some previously rejected manned flight candidates would now become viable for training, depending on the reasons for rejection.

The identified patterns could be used as potential indicators for RPAS candidate eligibility within an initial screening program.

6.2.2 Information Set Usage Objective Performance Measurement

With regards to objective performance measurement with the use of information sets this research found, that with only the use of a head tracker, that this type of observational performance measurement is not possible. Both groups showed remarkable similarity in terms of basic information set usage, which was not expected at the beginning of the research, but this similarity does not extend to the groups flight and spotting performance with the Gaming group proving to be much more effective. With the use of this equipment it is concluded that objective information set performance usage is not viable.

6.2.3 SimPACT Software Creation

The creation of the SimPACT software as a data processing and correlation tool significantly aided the speed and accuracy of the research contained within this thesis; it can be stated that without this software that analysis of the experimentation data would not have been possible within the timeframe of research.

To the authors knowledge no other software exists currently that captures TrackIR head positioning data, through XPlane, and processes it in relation to real world co-ordinates and specific ‘areas of interest’ with the addition of processing flight data.

Although this software has not been designed to work ‘out of the box’ with any other experimental setup the software can be easily recoded to work with other experiments involving head position co-ordinate systems. This means it is not just limited to simulation but could be used for any similar application.

The processing algorithms contained within the software are also original and would prove beneficial to any experimentation which contains conditional and sequential data analysis. This software is now in use with other doctoral research projects.

6.3 Contribution Summary

In summary the contributions can be listed as below:

- New insight into how information processing differs between persons of differing experiences and how this type of pattern may be identified and quantised
- Identification of distinct and previously unobserved patterns of information processing used by experienced computer gamers
- The identification of these patterns at the recruitment level would indicate a more viable candidate reducing training costs and times
- Creation of software and associated algorithms for advanced statistical analysis of head positioning; also likely cross-transferable to other types of positional analysis. Now in use with other doctoral research projects.
- Dismissal of objective task performance measurement (for the types of tested groups) by using information set usage at a macro level although deeper analysis may allow for a different form of objective performance measurement

6.4 Further Research

Further research into the potential for objective performance measurement and RPAS candidate viability is still required.

Although it was not possible to identify objective performance measurements relating to information set usage with the available equipment, the potential for this kind of measurement cannot be completely dismissed. Ideally this experimentation would have taken place with the use of an eye-tracker, rather than a head tracker; this would have allowed for a much more detailed analysis of information set usage as well as information subsets. These subsets could include altimeter, heading indicator and areas of interest; these subsets could be considered on a time dependant scale in which they are only useful at certain points during the flight; the participant could then be measured in relation to observing these subsets at an appropriate time whereas use of them while at an inappropriate time would lead to ineffective use of available data and, consequently, a lower performance score. It is suggested that any further research investigating this topic utilises a time based ‘area of interest’ method so that detailed information can be obtained regarding the participants use of these information subsets and its equivalence to overall task performance.

The hypothesised task based performance analysis follows on from the more detailed analysis of screen usage and would require the participation of multiple subject matter experts. This pool of subject matter experts was not available for this research otherwise comparisons would have been made between existing operators and Gamers, rather than with Gamers and Non-Gamers.

During the course of this research (not detailed within the thesis) several subject matter experts identified a major factor relating to RPAS operation within a combat zone that should be immediately investigated if possible. This potential research relates to the prolonged effects of drone operation in relation to the type of operations carried out by current drone pilots and their surrounding environment.

Drone operators, unlike manned aircraft pilots, are required to continuously operate their systems for long periods of time (approximately 8 hour shifts); this though is not the main factor to be explored but only an additional factor. Whereas a manned aircraft pilot, when carrying out a ground strike mission, is only required to proceed to target, identify target, deploy ordinance, observe impact and then return to base; a drone pilot, however, when carrying out the same task, is often required to loiter after observed impact to observe the 'aftermath' of the impact. This is so that any potential missed candidates can be identified as well as potential hostile combatant affiliations and patterns; this is often viewed at the highest resolution possible so that identification of individuals can be made.

What is not thought of is the psychological effect on the drone operator when there are human (either hostile or civilian) casualties are caused and the mandate that they must observe the emotional turmoil that has been caused by these casualties to social groups. This often painful observation is further compounded by the fact that they are remotely operating the drone from thousands of miles away and are often very close to their home life. It has been identified by one subject matter expert that he required at least 1 hour of transit between the ground control station and his home environment to 'decompress' before returning to his family; in effect the drone operator goes from a war zone, in which he may have caused death and suffering to multiple individuals, to a peaceful and normal home environment in which he/she is expected to act in a normal way and participate in normal life (such as a family meal). This type of very sudden transition could lead to unknown future psychological effects as well as disruption to a normal family life.

Urgent research is suggested into the above potential psychological issue, ideally before any long term psychological issues present in current drone operators.

Appendix A – References

1. Schlag C. The New Privacy Battle: How the Expanding Use of Drones Continues to Erode Our Concept of Privacy and Privacy Rights. Journal of Technology Law & Policy Volume XIII. Spring 2013

2. Baldor LC. Next Top Guns? Air Force faces shortage of drone pilots. MSNBC Webiste. August 2012

3. Spangler S. UAV Pilot Shortage & Military Intelligence. Jetwhine. March 2009

4. Arbib S. Israel – Gaza Conflict is Being Pwned by Video Game Players. Shalomolife Website. November 2012

5. Dalamagkidis, K. Aviation History and Unmanned Flight. Springer. 2014

6. Warrel K. The Evolution of the Cruise Missile. Defence Technology Information Centre. 1985

7. Zaloga S. V-1 Flying Bomb 1942- 52: Hitler's infamous' doodlebug'. Osprey. Jan 2005

8. Kamrani F. Using on-line simulation in UAV path planning. KTH. 2007

9. Taylor JW. Jane's All The World's Aircraft 1966–67. London: Sampson Low, Marston & Company, 1966 & Popular Mechanics Magazine. June 1956: 144

10. Anon. McDonald ADM-20 Quail Missile. Strategic Air Command Website http://www.strategic-air-command.com/missiles/Aircraft-Launched_Missiles/adm-20_quail_missile.htm

11. Anon. Pilotless Photo Drone Takes Aerial Pictures. Popular Mechanics Magazine. June 1956: 144

12. Parsch A. Northrop (Radioplane) Q-4/AQM-35. Designation Systems Website. 2002 http://www.designation-systems.net/dusrm/m- 35.html

13. Sanders R. UAVs: An Israeli Military Innovation. Joint Force Quarterly. Winter 2003

14. Haghshenass F. Iran’s Asymmetric Naval Warfare. Washington Institute Policy Focus #87. September 2008

15. Hull L. Drone makes first UK 'arrest' as police catch car thief hiding under bushes. Daily Mail Newspaper. Feb 2010
http://www.dailymail.co.uk/news/article-1250177/Police-make-arrest-using-unmanned-drone.html

16. Anon. Registration requirements for Unmanned Aircraft. Civil Aviation Authority Website
http://www.caa.co.uk/default.aspx?pageid=11188

17. Anon. Airworthiness requirements for Unmanned Aircraft. Civil Aviation Authority Website
http://www.caa.co.uk/default.aspx?pageid=11187

18. Anon. Unmanned Aircraft System Operations in UK Airspace - Guidance CAP 722. Civial Aviation Authority, Safety and Airspace Regulation Group, August 2012

19. Anon. Information and guidance associated with the operation of Unmanned Aircraft Systems (UASs) and Unmanned Aerial Vehicles (UAVs). Civil Aviation Authority Webiste
http://www.caa.co.uk/default.aspx?catid=1995&pagetype=90

20. Anon. Pilot Qualifications required to operate Unmanned Aircraft. Civil Aviation Authority Website.
http://www.caa.co.uk/default.aspx?pageid=11190

21. Anon. STANAG 4586 Edition 3. NATO. Nov 2012: vi

22. Anon. STANAG 4586 AIM. NATO. Nov 209: vi

23. Cole C. The 2012 Drone Wars Briefing. Drone Wars UK Website. Jan 2012
http://dronewars.net/2012/01/01/the-2012-drone-wars-briefing/

24. Parasuraman R, Sheridan T, Wickens C. A Model for Types and Levels of Human Interaction with Automation. IEEE TRANSACTIONS ON SYSTEMS Vol 30 No.3. May 2000

25. Sheridan T, Verplank W. Human and Computer Control of Teleoperators. Man Machine Systems Laboratory Report Cambridge: MIT. 1978

26. Boyd J. The Essence of Winning and Losing. Unpublished lecture. 1996

27. Clough B. Metrics, Schmetrics! How The Heck Do You Determine A UAV's Autonomy Anyway? Defence Technology Information Centre. 2002

28. Hasslacher B, Tilden MW. Living Machines. Robotics and Autonomous Systems, Elsevier. 1995

29. Cleary M, Abramson M, Adams MB, Kolitz S. Metrics For Embedded Collaborative Intelligent Systems. Charles Stark Draper Laboratory, Inc. 2000

30. Hill AF, Cayzer F, Wilkinson PR. Effective Operator Engagement with Variable Autonomy. 2nd SEAS DTC Technical Conference. 2007

31. Bonner M, Taylor R, Fletcher K. Adaptive Automation and Decision Aiding in the Military Fast Jet Domain. Defence Evaluation and Research Agency. 2000

32. Mcarley JS, Wickens CD. Human Factors Implications of UAVs in the National Airspace. University of Illinois at Urbana-Champaign Aviation Human Factors Division. 2005

33. Lee p. Remoteness, Risk and Ethos. Airpower Review Volume 15 Number 1. Spring 2006

34. Lyons T, Ercoline W, Freeman J, Gillingham K. Epidemiology of United States Air Force spatial disorientation accidents: 1990-1991. In Aircraft accidents: Trends in aerospace medicine investigation techniques, Defence Technology Information Centre. April 1992

35. McKinley R, McIntire K, Funke M. Operator Selection for Unmanned Aerial Systems: Comparing Video Game Players and Pilots. Aviation, Space, and Environmental Medicine, Volume 82 Issue 6. 2011: 635-642

36. Dixon S, Wickens C. Automation Reliability in Unmanned Aerial Vehicle Control: A Reliance-Compliance Model of Automation Dependence in High Workload. Human Factors: The Journal of the Human Factors and Ergonomics Society 48.3. 2006: 474-486

37. Symons S, France M, Bell J, Bennett W. Linking Knowledge And Skills To Mission Essential Competency-Based Syllabus Development For Distributed Mission Operations. No. AFRL-HE-AZ-TR-2006-0041, Human Effectiveness Directorate. July 2006

38. Denning T. France M, Bell J, Bennett W. Performance Assessment in Distributed Mission Operations: Mission Essential Competency Decomposition. Interservice/Industry Training, Simulation and Education Conference, Paper No. 1616. 2004

39. Anon. Evidence-Based Training™: Turning Research Into Results For Pharmecutical Sales Training. Axiom Health Proffesional Health White Paper. 2009

40. Richford N. Definition of Evidence based practises. Ehow, Science and Education Website
http://www.ehow.co.uk/about_5048440_definition-evidence-based-practice.html

41. Anon. Evidence Based Training (EBT). Internation Air Transport Association.
http://www.iata.org/whatwedo/ops-infra/Documents/ebt.pdf

42. Wilson G. An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. The International Journal of Aviation Psychology Volume 12, Issue 1. 2002

43. Lee Y, Liu B. Inflight Workload Assessment: Comparison of Subjective and Physiological Measurements. Aviation, space, and environmental medicine 74.10. 2003: 1078-1084

44. Hart S, Staveland L. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. 1988d Environmental Medicine, Volume 74, Number 10. October 2003: 1078-1084(7)

45. Anon. NASATLX Company Statement. NASATLX Website.
http://nasatlx.com

46. Reid G, Potter S, Bressler J. Subjective Workload Assessment Technique (SWAT): A User’s Guide. Wright Patterson Air Force Base, Armstrong Aerospace Medical Research Laboratory. July 1989

47. Rubio S, Díaz E, Martín J, Puente J. Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. Universidad Complutense de Madrid. 2004

48. Duchowski A. A Breadth-First Survey of Eye Tracking Applications. Behavior Research Methods, Instruments & Computers 34.4 2002: 455-470

49. Hasse C, Grasshoff D, Bruder C. Eye-tracking parameters as a predictor of human performance in the detection of automation failures. Proceedings HFES Europe Chapter Conference Toulouse. 2012: 133-144

50. Reed H, Drance S. Essentials of Perimetry: Static and Kinetic. Oxford University Press, USA. 1972

51. Webb P. Bioastronautics Data Book. NASA SP-3006 Scientific and Technical Information Division NASA: Washington D. C. 1964

52. Anon. Landscape & Visual Impact Assessment - Landscape Value Mapping of Hong Kong. Executive Summary, Planning Department, Government of Hong Kong SAR. July 2005 http://www.epd.gov.hk/eia/register/report/eiareport/eia_1252006/html/eiareport/Part2/Section11/Sec2_11.htm

53. Anderson J, Bower G. Human Associative Memory. Psychology Press. 2014

54. Everitt J, Fletcher S, Caird-Daley A. Task Analysis of discrete and continuos skills: a dual methodology approach to human skills capture for automation.Theoretical Issues in Ergonomics Science vol. 16 no 5 513-532. 2015

55. Gonzalez C, Thomas R, Vanyukov P. The Relationships between Cognitive Ability and Dynamic Decision Making. Intelligence 33.2 2005: 169– 186

56. Nicholson B, O'Hare D. The effects of individual differences, prior experience and cognitive load on the transfer of dynamic decision-making performance. Ergonomics, 57:9. 2014: 1353-1365

57. Di Nocera F, Camilli M, Terenzi M. A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload. Journal of Cognitive Engineering and Decision Making 1(3). Dec 2007

58. Wickens C, Hutchins S, Cumming J, Carolan T. Effectiveness of Part-Task Training and Increasing-Difficulty Training Strategies A Meta-Analysis Approach. Human Factors: The Journal of the Human Factors and Ergonomics Society vol. 55 no. 2. April 2013: 461-470

59. Anon. Reaper Pilot. British Royal Air Force. 2012

60. Atkinson R, Shiffrin R. Human memory: A proposed system and its control processes. The psychology of learning and motivation 2. 1968: 89-195

61. Kallio K. Kaipainen K. Mäyrä F. Gaming Nation? Piloting the International Study of Games Cultures in Finland. 2007

62. Anon. X-Plane Plugin Website.
http://www.xpluginsdk.org

63. Anon. XHSI Sourceforge Website
http://xhsi.sourceforge.net/

64. Bussemaker J. Software Website.
http://www.utr-online.com

65. Anon. Mathworks Website.
http://www.mathworks.co.uk/products/matlab

66. Boyd J. Conceptual Spiral. A Discourse on Winning and Losing. Defense and the National Interest, Atlanta. 1192

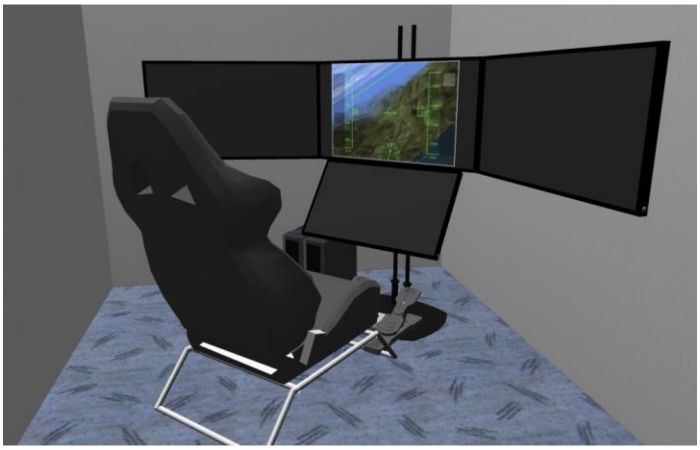
67. Sarter N, Mumaw R, Wickens C. Pilots' Monitoring Strategies and Performance on Automated Flight Decks: An Empirical Study Combining Behavioral and Eye-Tracking Data. , Human Factors: The Journal of the Human Factors and Ergonomics Society vol. 49 no. 3. June 2007: 347-357

68. Schrieber B. Transforming Training: A Perspective on the Need and Payoffs from Common Standards. Educational Publishing Foundation Vol. 25 No. 3. 2013: 294

69. Schreiber B. Stock W, Bennett, W. Distributed Mission Operations Within-Simulator Training Effectiveness Baseline Study: Metric Development and Objectively Quantifying the Degree of Learning. Human Effectiveness Directorate, Volume 2. 2006

Appendix B - Hardware and Software

Prototype GCS CAD



GCS Specifications

Computer:

Case: COOLERMASTER CM690 MKII ADVANCED CASE

CPU: Overclocked Intel® Core™i7-2700k Quad Core(3.50GHz @ max 5.00GHz)

Motherboard: ASUS® MAXIMUS IV EXTREME-Z: INTEL Z68 ROG MOTHERBOARD

Memory (RAM): 8GB KINGSTON HYPER-X GENESIS DUAL-DDR3 1600MHz, X.M.P (2 x 4GB KIT)

Graphics Card: 3GB NVIDIA GEFORCE GTX 580 - 2 DVI, HDMI, DP - 3D Vision Ready

2nd Graphics Card: 3GB NVIDIA GEFORCE GTX 580 - 2 DVI, HDMI, DP - 3D Vision Ready

Memory: - 1st Hard Disk 500GB SEAGATE Barracuda SATA-III 6Gb/s HDD, 32MB Cache (7200rpm)

DVD/BLU-RAY Drive: 24x DUAL LAYER DVD WRITER ±R/±RW/RAM

Power Supply: CORSAIR 1050W PRO SERIES™ HX1050-80 PLUS® SILVER MODULAR

Processor Cooling: COOLIT ECO II FAT BOY PUSH/PULL CONFIG LIQUID CPU COOLER

Thermal Paste: ARCTIC MX-4 EXTREME THERMAL CONDUCTIVITY COMPOUND

Sound Card: ONBOARD 6 CHANNEL (5.1) HIGH DEF AUDIO (AS STANDARD)

Network Facilities: 10/100/1000 GIGABIT LAN PORT - AS STANDARD ON ALL PCs

USB Options: 6 x USB 2.0 PORTS @ BACK PANEL (MIN 2 FRONT PORTS) AS STANDARD

Display:

3 x BenQ 27IN LED EW2730 VA 16:9

1 x Ilyama T2250MTS 22 inch Wide LCD

Peripherals:

1 x Crucial M4 128Gb Hard Drive

1 x Track IR Pro

1 x Thrustmaster HOTAS Warthog stick and throttle

1 x Saitek Pro Rudder

1 x Saitek Eclipse Keyboard

1 x Dell Mouse

Software:

Windows 7 Enterprise 64bit

Microsoft Office 2007

Microsoft Flight Sim X Gold Edition (case study)

Laminar Research X-Plane 9 (main experiment)

Predator Model

VivendoByte Data Logger (case study)

PlanG Flight Planner (case study)

FSUIPC plug-in (case study)

XUIPC plug-in (main experiment)

Google Earth

Blackbox data logger (case study)

Panel Restore (main experiment)

Rex Real-time Weather Engine (main experiment)

VFR Photogrphic scenery pack (main experiment)

Optitrak Software Development Kit (case study)

[Intentionally left blank]

Appendix C - Competencies

Revised MEC example task lists

Example MEC KSE, SC and Task lists

Mission Essential Competencies	Supporting Competencies	Knowledge	Skills	Expereinces
Organize Forces to Enable Combat Employment	Adaptability	Comm Standards	Adapts to changes in environment	Restricted Weapons Load
Detects Factor Groups in Area of Responsibility	Communication	Commit Criteria	Adapts to friendly changes	Limited Fuel Remaining
Intercept and Target Factor Groups	Decision Making	Engage Criteria	Adapts to threat changes	Operating Area Restrictions
Engage-Employ Ordnance & Deny Enemy Ordnance	Flight Battle Management	Follow-on Options	Anticipates problems	Restrictions to Visibility
Assessment/Reconstitute-Initiate Follow on Actions	Identification	Formation	Builds picture	Visual Illusions
Remain Oriented to Force Requirements Recognize	Information Management	Friendly Capabilities	Controls Intercept Geometry	Marginal/Minimal Cloud Clearance
Trigger Events that Require Shift in Phase	Situational Awareness	Mission Objectives	Develops new options	Daytime Employment
	Timeline	Package Composition	Executes merge game plan	Dusk Employment
	Weapons Engagement Zone Management	Phase of Mission	Executes short range game plan	Night Employment

		ROE	Interprets sensor output	Mountainous Terrain
		Threat Capabilities	Listens	G-Induced Physical Limitations
		Time Restrictions	Maintains formation	Degraded Comm
			Makes assessment	Degraded Nav
			Manages mission timing	Degraded Weapons Employment
			Manages stress	Battle Damage
			Multi-tasks	Supersonic Employment
			Prioritizes communications	Full Range of Adversary Air Threat/Mix
			Radar mechanization	Full Range of Adversary Ground Threat/Mix
			Rebuilds picture	Operations with Friendly IADs
			Reforms	Operations with Ownship and Friendly ECM
			Selects tactic	Operations Against Threat with Chaff/Flare
			Sorts information	Operations with Friendly Use of Chaff/Flare
			Sorts targets	Operations Against Comm Jam/Spoofing

			Speaks clearly	Operations Against Adversary ECM
			Switchology	Roe Limitations and Restrictions
				Fatigue/Time on Task
				Task Saturation
				Limited Time to Act/React to Situation
				Radar Search Responsibilities
				Targeting and Sorting Responsibilities
				Air Refuelling
				Live Weapons Employment
				Simulated Weapons Employment
				Various Initial Conditions
				Emergency Procedures
				Formation Responsibilities
				Lost Mutual Support
				Dynamic Retasking/Scramble Operations

				Various Employment Altitudes 1:1 Force Ratio 1:2 Force Ratio 1:3+ Force Ratio OCA Escort Missions OCA Sweep Missions Employment with Various Packages

Revised Task list from example Knowledge and Skills

Adaptability	Comms	Interpretation	Execution	Control	Time	Physiological	Mission
Follow-on Options	Comms Standards	Engage Criteria			Time Restrictions		Mission Objectives
	Comms Criteria	Friendly Capabilities					Package Composition
		Phase of Mission					ROE
		Threat Capabilities					
		Formation					
Adapts to changes in environment	Listens	Sorts information	Executes merge game plan	Controls Intercept Geometry	Manages mission timing	Manages stress	
Adapts to friendly changes	Prioritizes Communications	Sorts targets	Executes short range game plan	Maintains formation		Multi-tasks	
Adapts to threat changes	Speaks Clearly	Rebuilds picture		Radar mechanization			
Develops new options		Makes assessment		Reforms			
		Interprets sensor output		Switchology			
		Builds picture					
		Selects tactic					
		Anticipates problems					

Revised Task list from example Expereinces

Employment	Adaptation	Restrictions	Physiological	Responsibility	Mission
Supersonic Employment	Limited Fuel Remaining	Restricted Weapons Load	Fatigue/Time on Task	Formation Responsibilities	OCA Escort Missions
Operations Against Adversary ECM	Degraded Comms	Operating Area Restrictions	Task Saturation	Targeting and Sorting Responsibilities	OCA Sweep Missions
Operations Against Comm Jam/Spoofing	Emergency Procedures	Roe Limitations and Restrictions	G-Induced Physical Limitations	Radar Search Responsibilities	Full Range of Adversary Air Threat/Mix
Operations with Friendly Use of Chaff/Flare	Battle Damage		Visual Illusions		Full Range of Adversary Ground Threat/Mix
Daytime Employment	Lost Mutual Support				
Dusk Employment	Restrictions to Visibility				
Night Employment	Dynamic Retasking/Scramble Operations				
Employment with Various Packages	Degraded Nav				
Live Weapons Employment	Degraded Weapons Employment				
Simulated Weapons Employment	Various Employment Altitudes				
Air Refuelling	Various Initial Conditions				
Operations Against Threat with Chaff/Flare	Mountainous Terrain				
Operations with Ownship and Friendly ECM	Marginal/Minimal Cloud Clearance				

Operations with Friendly IADs	Limited Time to Act/React to Situation 1:1 Force Ratio 1:2 Force Ratio 1:3+ Force Ratio				
Many tasks become autonomous with rise in autonomy	Many tasks become autonomous with rise in autonomy	Change with increasing autonomy	Unlikely to change much	Many tasks become autonomous with rise in autonomy	Unlikely to change much

[Intentionally left blank]

Appendix D – Case Study

Questionnaires and Configuration

Questionnaire A

Questionnaire A

The information gathered by this questionnaire is anonymous and strictly confidential, if you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions.

General:

- What is your age?
- What is your gender?
- Please list the types (make and model) of unmanned commercial platforms flown (if applicable) and the number of hours flown on each platform
- Please list types (make and model) of unmanned military platform flown (if applicable) and the number of hours flown on each platform
- Have you flown in an unmanned platform in an active situation? (war/operation/theatre) y/n
- Which platforms unmanned have you flown in combat and how many hours flown during combat on each platform?
- What types of mission have you flown on unmanned platforms? (please list)
- Do you have previous manned flight experience? y/n

  (if no, proceed to question 11)

    - If yes to the above what manned commercial platform have you flown? (if applicable)
    - If yes to the above what manned military platform have you flown? (if applicable)
    - Can you estimate the number of manned hours flown in total?
    - Can you give a brief description/overview of the RPAS training process that you encountered (if possible)
    - How long did the RPAS training last approximately?
    - Can you estimate the number of virtual hours (Simulated) logged during training?

- Can you estimate the number of live hours (Real flight) logged during training?
- Do you have any other simulator experience (Microsoft flight sim x, x- plane), please state which?
- How often do you use recreational flight simulators? (hrs per week if any)

RPAS Specific:

(Please expand if necessary using a word processor or expand onto separate sheets if using paper; a mind mapping program will be made available if preferred)

- While flying, are there any non-platform related distractions (music, television, conversations) that remove your concentration from the operational displays? Please list
- Please list the team members of a ground control station and state their role and responsibilities
- Please list the types of missions you have flown
- Can these missions be broken into phases? (i.e. take- off, fly to way point, deploy, return, land)

y/n

- It has been identified that separate operators are used for take- off/landing and mission flight, which category do you fall under?

- If yes to the above can you list all mission phases for the most basic mission in chronological order?

- With regard to the above questions can you estimate the length of time taken for either mission, phases or both for all listed missions (if possible)?

- If yes to phases can you list the tasks (if possible, if too many please list most basic) required to complete all phases? (i.e. gear up, level off, assume correct direction for first way point, systems check, flight plan change)

- For the above tasks can you list if possible:
    - the parameters by which realise a task has started (i.e. reach certain height, speed etc)
    - the information you consciously observed to become aware of the task
    - the information you consciously observe to monitor/complete the task
    - your actions to complete the task
    - the information you consciously observe to verify the task has been completed
    - Estimate the time it would normally take to complete the task

Questionnaire B

On-site

GCS and Simulator:

For GCS:

- How many sets of information (screens) do you normally fly with?
- How many of those screens contain duplicate information?
- Please list what you believe to be key information groups (i.e. aircraft data, visual/camera, map etc)
- Please list, in descending order, the most used key information groups
- Please identify devices of flight control (keyboard, stick, throttle, rudder, mouse, other)
- Please list, in descending order, the most frequently used control inputs

For Simulator:

- Would you consider the provided simulator to be an adequate GCS?
- Which information sets are not displayed by the simulator that would normally be present in the GCS?
- What operational controls are not available on the simulator with respect to the GCS?
- How does the simulator frame rate compare to a GCS?

Test Flight:

- Is the level of force required on te stick correct? y/n
- If no to the above please state whther it needs to be more or less reponsive -10 to +10
- Are the model flight dynamics accurate? 1- 10
- please define why
- Are the screens of an appropriate size? 1- 10
- Is all available information easy to view? 1- 10
- Are you physically comfortable? 1-10
- Is the simulator more comfortable than a GCS? y/n
- Please define why
- Which do you prefer in an operational context in terms of comfort, GCS or Simulator? y/n
- Please define why
- Which display do you prefer in an operational context, GCS or Simulator?
- Please define why

Equipment and Configuration reference

Hardware:

Thrustmaster HOTAS Warthog Stick and Throttle

Saitek Combat Rudder Pedals

Track IR Headtracker

GT Omega Racing Simulator

3 x BenQ 27" monitors

1x Ilyama 22" Touchscreen monitor

Software:

Microsoft Flight Simulator X Deluxe

Blackbox data logger

Google Earth

Vivendobyte data logger

Custom designed Loughborough University Predator model

Configuration:

BenQ1 (display1) = MFD map and info display

BenQ2 (display 2) = Forward facing camera

BenQ3 (display 3) = Google Earth real-time link

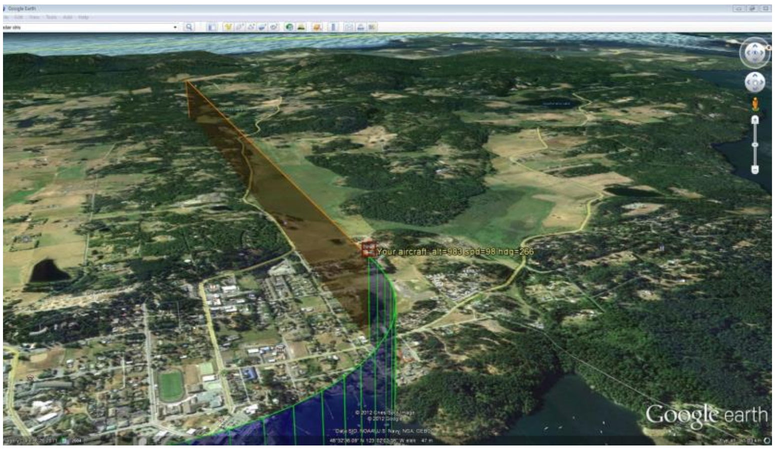
Ilyama (display 4) = Instrument Display

Display 1:

Display 2:

(ins here)

Display 3: 3D output shown but 2D also available



Display 4:


GARMIN
NAV1 117.20 113.70
NAV2 116.80 110.60
COM1 128.300 128.250
COM2 127.800 127.900
VOL
NAV
HDG
ALT
COM
CRS
BARO
RANGE
PAN
FPL
PROC
MENU
CLR
ENT
FMS
100
340°
NAV1
TAS 97KT
TRACK UP
OAT 12°C
29.92 IN
XPDR 1200 ON
LCL 11:15:55
INSET
OBS
CDI
XPDR
TMR/REF
NRST
GEAR UP
FUEL TANKS
ALT 00000
VS 00000
PR
AP
HDG
NAV
APR
REV
ALT
UP
DN

Control Sheet

Control Systems

Throttle:

- Main Stick – Forward/Back = Accelerate/Decelerate

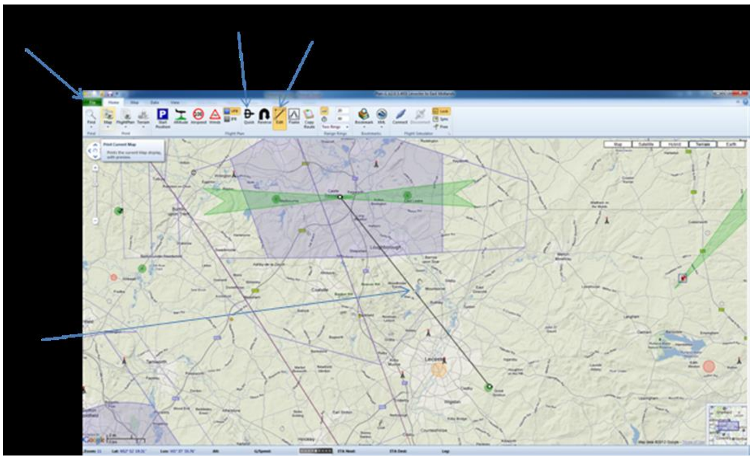
Stick:

- Main Stick
    1. Bank Left
    2. Bank Right
    3. Pitch Forward (Descend)
    4. Pitch Back (Climb)
- Hat Control - View movement
    1. Left/Pan Left
    2. Right/Pan Right
    3. Up/Pan Up
    4. Down/Pan Down
- Button A - View Reset
- Main Trigger – Brake

- H3 – Zoom View

Flight Plan Creation

In Plan G (Screen 1)**:**



- Select Quick (1) from the Flight Plan Menu
- Enter “EGBG” (Leicester) in the “from ICAO” box
- Enter “EGNX” (Castle Donnington) in the “to ICAO” box
- Click Edit (2)
- Drag the central box to either (4)
    1. An approximation of the Google Earth Highlighted area
    2. As close as possible to a set of Latitude/Longitude co-ordinates defined by the examiner
    3. An named area (village, town, city) defined by the examiner
- Select “File” (3) from the upper menu
- Select “Save As”
- Select “FSX Flight Plan.Pln”
- Name the File “*YourNamex*” where x represents how many flights plans you have previously created, this does not include edited “in mission” flight plans.

Loading Your Flight Plan

In Microsoft Flight Simulator X (FSX):

- Select “Flights”
- Select “Flight Planner”
- Select “Load”
- Select your previously created flight plan *“yournamex.pln”*
- Select “Open”
- Select “OK”
- When prompted to move Aircraft to origin select “No”

Pre-Flight Setup

Autopilot:

- Select “NAV”(1)
- Select “ALT”(2)
- Press the up and down arrows (3) until the digital Alt display reads 1000
- On the large panel select the “NAV/GPS” switch (4) and set it to GPS
- You are now ready to begin your flight

Taking Off

Please read this through first before taking off

- Apply Full Throttle (Push Throttle all the way forward)
- Using the PFD speed indicator (1) as a guide pull back **gently** on the flight stick until you are airborne, only attempt to pull back when the speed reaches 85 KTs
- Select the “AP” button on the radio panel (2) and release the flight stick
- Select the “Gear Up” option (3)
- Monitor the speed and keep it within the 90- 120Kts boundary otherwise you may damage the aircraft

Flight Plan Update

In PlanG

- Confirm that your current flight plan is being displayed in PlanG
- Make sure “Edit” and “Free” are selected
- Locate the new waypoint on the map, this waypoint will be given either in terms of
    1. Place Name
    2. Longitude and Latitude
    3. An existing marker placed in Google Earth or PlanG
- Drag the nearest white box ahead of the aircraft on the flight path to the stipulated location
- Take note of the new waypoints number, this can be seen by hovering the cursor over the white box
- Select “File” from the upper menu
- Select “Save As”
- Select “FSX Flight Plan.Pln”
- Name the File “*YourNamexedit*” where x represents your original files x value
- Click OK

Cont......

In FSX

- On the Autopilot, de-select “NAV”

- Select “Flights”
- Select “Flight Planner”
- Select “Load”
- Select your previously created flight plan *“yournamex.pln”*
- Select “Open”
- Select “OK”
- When prompted to move Aircraft to origin select “No”
- Wait for new plan to load
- Using the MFD (pictured) set your new destination to that waypoint
    1. Select FPL (1)
    2. Press the centre of the PUSH/CRSR dial(2)
    3. Use the <-/-> parts of the dial (3) to scroll through waypoints
    4. Select the desired waypoint (the one which you crated in PlanG
    5. Select MENU (4)
    6. Select ENT (5) to activate
    7. Click and hold CLR (6) to return to original MFD display
- On The Autopilot, re-select “NAV”

Landing

- Make sure you are roughly lined up on the runway
- At 1nm distance from the waypoint (can be observed in the top right of the MFD disengage “AP”
- Your are now flying manually
- Select the “Gear Down” option
- Land as you best you can using stick and throttle
- Apply brakes (Main trigger on throttle)

Data Logger Metrics

Data logger configured to store relevant data for the following sets:

- aircraft clock
- stall condition
- on-ground aircraft condition
- latitude
- longitude
- altitude
- ground altitude
- vertical speed
- heading
- true airspeed
- bank
- rpm engine
- fuel weight
- fuel flow engine 1
- indicated airspeed
- simulation rate
- flaps status
- aircraft name
- throttle engine 1
- system clock
- gear status
- autopilot master status
- overspeed status
- realism
- collision status
- collision with others aircrafts status
- parking brakes
- ambient in cloud
- time of day

Relevant output

Questionnaire A

Initial Investigation into Task Analysis of Basic Control of Predator based Remotely Piloted Aerial System

Questionnaire A

The information gathered by this questionnaire is anonymous and strictly confidential, if you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions.

General:

- What is your age? **38**
- What is your gender? **Male**

    - Please list the types (make and model) of unmanned commercial platforms flown (if applicable) and the number of hours flown on each platform

none

- Please list types (make and model) of unmanned military platform flown (if applicable) and the number of hours flown on each platform

MQ-9 Reaper 1300hours

    - Have you flown in an unmanned platform in an active situation? (war/operation/theatre)

yes

- Which platforms unmanned have you flown in combat and how many hours flown during combat on each platform?

MQ-9 Reaper 1250 operational hours

Sortie tasking was ISR throughout. The platform has an armed capability.

- What types of mission have you flown on unmanned platforms? (please list)

Intelligence Surveillance and Reconnaissance

Dynamic engagements

- Do you have previous manned flight experience?

yes

(if no, proceed to question 11)

  - If yes to the above what manned commercial platform have you flown? (if applicable)

none

  - If yes to the above what manned military platform have you flown? (if applicable)

C-130 Hercules (1100 hours)

Tucano (600)

PC-9 (1100)

- Can you estimate the number of manned hours flown in total?

3100

- Can you give a brief description/overview of the RPAS training process that you encountered (if possible)

RAF Reaper Operational Conversion Course

Effects of Controls (Basic aircraft control and operation)

ISR (intelligence gathering)

Weapons employment

Close Air Support

- How long did the RPAS training last approximately?

2 months

- Can you estimate the number of virtual hours (Simulated) logged during training?

2

- Can you estimate the number of live hours (Real flight) logged during training?

  30

- Do you have any other simulator experience (Microsoft flight sim x, x- plane), please state which?

100 hours C-130

50 hours Tucano

- 10 PC-9 How often do you use recreational flight simulators? (hrs per week if any)

  0

RPAS Specific:

(Please expand if necessary using a word processor or expand onto separate sheets if using paper; a mind mapping program will be made available if preferred)

- While flying, are there any non-platform related distractions (music, television, conversations) that remove your concentration from the operational displays? Please list

  Conversation is required to enable effective function of the platform and ensure that there is a base level of alertness throughout, particularly during low arousal periods of the clock. Training enables conversation to cease at the social level to operational communication seamlessly. Conversation is verbal only, limited to no eye contact to enable constant monitoring of aircraft and operational functions.

- Please list the team members of a ground control station and state their role and responsibilities

  Platforms vary.

MQ-9 Pilot – operate aircraft systems and conduct ISR. Mission Commander

Sensor Operator – operate sensor (Full Motion Video) primary, SAR secondary

Mission Intelligence Co-ordinator – Often a Intelligence specialist or Image Analyst. Focal point for incoming Comms and distribution forward to Pilot/SO. Manages Situational Awareness tools for display to crew.

- Please list the types of missions you have flown

Intelligence Surveillance and Reconnaissance

Dynamic engagements

- Can these missions be broken into phases? (i.e. take-off, fly to way point, deploy, return, land)

Yes

  - It has been identified that separate operators are used for take- off/landing and mission flight, which category do you fall under?

Both

- If yes to the above can you list all mission phases for the most basic mission in chronological order?

Take Off

Transit

On Task – potential re-task and further transit as required

Transit

Landing

- With regard to the above questions can you estimate the length of time taken for either mission, phases or both for all listed missions (if possible)?

T/O - 10 mins

Transit - Distance to target via ATC approved route divided by speed.

On Task- Endurance less, Take Off and transit time and the planned recovery and landing time. Or time from arrival on task until planned recovery time to facilitate a specific handback time.

Transit - As above

Landing - 10-30 minutes depending upon ATC.

- If yes to phases can you list the tasks (if possible, if too many please list most basic) required to complete all phases? (i.e. gear up, level off, assume correct direction for first way point, systems check, flight plan change)

Take Off - Gear up – check laser – basic systems checks - obtain onward clearance HANDOVER – basic systems checks – operational checks – check in with tasking agency – provide support – check out with tasking agency – obtain recovery clearance – recovery checks – HANDBACK –landing checks – land.

26. For the above tasks can you list if possible:

- the parameters by which realise a task has started (i.e. reach certain height, speed etc)
- the information you consciously observed to become aware of the task
- the information you consciously observe to monitor/complete the task
- your actions to complete the task
- the information you consciously observe to verify the task has been completed
- Estimate the time it would normally take to complete the task

Clearly a product of training.

All actions are conducted in a proscribed manner as a result of either;

Conditioned actions e.g. gear retraction at a given speed and height then into the operation and indication to ensure correct function

Range, actions to be completed by… e.g. operational checks

Priority, At the very basic level – Aviate, Navigate, Communicate. There are a large number of actions and events that have to be completed for successful operation of the aircraft and conduct of the sortie tasking. Operators are aware through training what these are, the objective, is to start in sufficient time and sequence to ensure that all are complete in the most expeditious manner. There is always the option for variation and no ‘perfect solution’ merely different opinions and techniques to achieve the overall task.

Questionnaire B

Initial Investigation into Task Analysis of Basic Control of Predator based Remotely Piloted Aerial System

Questionnaire B

The information gathered by this questionnaire is anonymous and strictly confidential, if you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions.

For Scales of -10 to 10; -10 indicates very low, 10 indicates extremely high or perfect. Please select inclusively within this range.

Predator GCS (Ground Control Station) and Loughborough University Simulator:

For The Predator GCS:

- How many sets of information (screens) do you normally fly with?

8 (2 HDD) screens, 3 routinely with usable info, 1 with aircraft management info (limited use - generally for handover so multiple crews can operate without constant reference to previous operators). Other screens run systems health info/MFW information. Can operate with 5 screens displaying critical information and systems when the tasking requires.

- How many of those screens contain duplicate information?

Ideally none. The aircraft management screen displays historic information about the flight. MFW info will be repeated behind at the MFW workstation.

- Please list what you believe to be key information groups (i.e. aircraft data, visual/camera, map etc)

FMV on the main screen plus HUD info. 2 HDD for aircraft control and navigation. FV for SA. Systems health for the contractors when there is an issue. MFW info when required but generally not essential primary vs secondary payload.

- Please list, in descending order, the most used key information groups

Aviate - HUD - FMV

Navigate - Tracker and FV for SA and positioning/operational tasking

- Please identify devices of flight control (keyboard, stick, throttle, rudder, mouse, other)

AP, Control Column, Throttle, Rudder, Mouse/Keyboard (update current mission info)

- Please list, in descending order, the most frequently used control inputs

LR - Primary Flying Controls

MC - AP Control column, keyboard/mouse, throttle.

<u>For Loughborough University Simulator:</u>

- Would you consider the provided simulator to be an adequate GCS and why (if possible)?

Yes, all primary functions available, some functions interesting and nice but not necessary. Other elements/software improvements would increase the efficiency of operation but the fundamentals are available.

- Which information sets are not displayed by the simulator that would normally be present in the GCS?

Throttle/Tq gauge.

Temperature and Pressure gauges to monitor systems.

Fuel gauges

CWP or alerting system.

- What operational controls are not available on the simulator with respect to the GCS (including buttons, panels, switches and control inputs)?

px subscale

- How does the simulator frame rate compare to a GCS?

Good not perfect. GCS frame rate will adjust dependant on bandwith available.

Test Flight:

- Is the level of force required on the stick correct? y/n

A little sensitive but close enough, familiarity would overcome this. Not used to static column. software prevents excurions outside aircraft operating limits.

- If no to the above please state whether it needs to be more or less responsive. *-10 to +10*

  A little less.

  - Are the model flight dynamics accurate? *-10 to +10*

  Close

  please define why

  Climb rate to fast 2

  roll rate 4

  turn rate -2

  - Are the screens of an appropriate size? *-10 to +10*

  yes 6

- Is all available information easy to view? *-10 to +10*

Aircraft performance instruments could be HUD or directly beneath front screen -5. All others good 5.

- Are you physically comfortable? *-10 to +10*

OK, only assessed for short period -3. ISR requires long duration operation. Control column too close, throttle variable position.

- Is the simulator more comfortable than a GCS? y/n

stby

  Please define why

- Which do you prefer in an operational context in terms of comfort, GCS or Simulator?

GCS

Please define why

big comfy padded seat. Controls in better location. (GCS is an example of poor ergonomics. Screen position vs seat position etc.

- Which display do you prefer in an operational context, GCS or Simulator?

SIm

Please define why

Far bigger screens, far better resolution, mapping software appears better, whether other programmes can be fitted would need to be checked. Correct/Better location for information still required though.

[Intentionally left blank]

Appendix E - Main Experiment

Ethical Code of Practise

CODE OF PRACTICE ON INVESTIGATIONS INVOLVING HUMAN PARTICIPANTS

1. Introduction

It is generally accepted in this country that investigations on human beings should be governed by codes such as those of the World Medical Association (The Declaration of Helsinki, 1964; revised 1975) and of the Medical Research Council. A number of professional associations and learned societies have issued similar statements of ethical principles to guide their members, amongst them the British Psychological Society and the British Sociological Association, and many institutions where investigations involving human participants are carried out have formulated codes of practice to provide more detailed guidance for their staff involved in such activity. It is now commonplace for ethical committees to have been established to supervise the ethics of investigations involving human participants and to consider individual proposals. Indeed, an increasing number of Research Councils require ethical review of projects prior to making a grant.

Investigations involving human participants are undertaken in several departments in this University in the course of teaching and research. The University seeks to ensure that the conduct of all its staff and students carrying out such work, whether human biological, psychological or sociological, conforms to accepted professional standards and is known to do so. The University has an Ethics Committee which looks at all aspects of ethical conduct at the University. It delegates responsibility for investigations on Human Participants to the Ethics Approvals (Human Participants) Sub-Committee, whose remit is to guide and assist investigators and ensure that full consideration is given to the health and safety of the participants taking part and that the rights of the participants are protected.

2. Ethics Approvals (Human Participants) Sub-Committee

2.1 Terms of Reference

i. To consider ethical issues relating to the research (including enterprise activities) and teaching of the University which involves investigations on human participants.

ii. To produce guidance documents for researchers on ethical issues that relate to investigations on human participants and to publish these guidelines on the University web-pages.

iii. To be available to give advice to staff and students of the University who wish to undertake such investigations on the ethical considerations involved.

iv. To keep the University Code of Practice on investigations on human participants under review and to recommend to the Ethics Committee such modifications as from time to time are deemed necessary.

v. As a matter of routine, to consider the ethical implications of individual proposals for investigations on human participants and to advise whether or not these are acceptable.

vi. Individual proposals will be considered on a monthly basis and decisions will be full approval, conditional approval or not approved:

Full approval will be complete approval, with no alterations needed. The study may begin straight away.

Conditional approval will consist of comments returned to the investigators. Investigators will have six weeks after receiving the comments to respond. If investigators do not respond to the comments within six weeks, conditional approval will be withdrawn. Extensions and reminders regarding the six week deadline will be provided.

Not approved will mean that the proposal has been rejected, and the investigators will have to re-apply for ethical approval.

vii. To meet once in each term to discuss proposals and other ethical issues and to meet nine other times in the year to discuss proposals only. All meetings will be reported to the Ethics Committee. Investigators will wait no longer than six weeks from submission of a proposal to an initial decision on that proposal.

viii. To communicate regularly with the Research Committee and the Health, Safety and Environmental Committee by way of an annual report via the Ethics Committee.

ix. To escalate any complex proposals, repeat offenders with regards to proposals, or issues with Schools, to the Ethics Committee for their consideration.

x. To perform an annual audit of all proposals to ensure that the correct procedures and paperwork are being maintained.

2.2 Membership

Chair: appointed by Ethics Committee.

Six representatives from the Ten Schools of the University.

Ethical and Environmental Officer (Students' Union).

Health Safety and Environmental Officer

Up to 4 co-opted members (to include one external occupational health expert).

Regularly in attendance:

A member of the Research Office or Research Team.

Details of the current membership may be obtained from the Secretary to the Sub- Committee.

2.3 Terms of Office for Members

Appointed members will be eligible for re-appointment on a consecutive basis once only, so that the maximum period of continuous service for such a member will be six years, after which the member will not be eligible for re-appointment until the expiration of one year from the end of their term of office. Co-opted members can hold office, under normal circumstances, for not more than three consecutive years after which such a member cannot be eligible for co-option until the expiration of one year from the end of their term of office.

3. Scope of the Code

This Code of Practice was initially approved by Senate on 22 July 1988 and by Council on 6 July 1988, and revised by Senate on 11 March 1992. It was further revised by the Ethical Advisory Committee in May 2003 and approved by Senate on 25 June 2003 and Council on 15 July 2003. It was further revised due to changes in the University ethics structure on 31 July 2012 and approved by the Loughborough University Ethics Committee on 18 June 2012.

3.1 Context of investigation

a. All investigations involving human participants fall within the scope of the Code (including research investigations, class teaching experiments/demonstrations/ research investigations, student projects, surveys and questionnaires) and should conform with the appropriate University and/or external guidelines. Completion of the Ethical Clearance Checklist devised by the Sub-Committee will demonstrate whether or not a proposal meets with the ethical principles adopted by the University. If a proposal does not comply with all sections of the Ethical Clearance Checklist, investigators are expected to complete a full submission to the Ethical Approvals (Human Participants) Sub-Committee after reading Section 7 of this Code. All investigators are responsible for familiarising themselves with the appropriate external guidelines for their own discipline/area of research.

b. It is essential that junior researchers /students acting as investigators are under the supervision of a senior researcher/member of staff. It is the responsibility of the supervisor to see that the junior researchers /students are aware of the relevant guidelines and to ensure that they are observed. The Ethical Approvals (Human Participants) Sub-Committee will expect supervisors to take responsibility for submitting details of proposed investigations for approval where necessary.

c. The Sub-Committee is prepared to consider protocols on a 'generic' basis where it is the intention to adopt the same procedure in a number of related investigations. A generic protocol will be cleared by the Sub-Committee for use by those investigators named on the submission under the direction of the applicant. Individuals wishing to use the approved protocol who are not named on the submission document should apply to the protocol holder for permission to practise the generic procedure. It will be the responsibility of the holder and his/her head of section to ensure that such individuals are fully competent to use the protocol before permission is given. The names of individuals cleared through this procedure should be appended to the list of investigators in the copies of the protocol document held by both the Department/School concerned and the Sub-Committee Secretary. Investigators wishing to use approved generic protocols in combination, rather than as isolated techniques, should seek clearance from the Sub-Committee.

3.2 Investigations conducted off campus

Staff or students who wish to carry out investigations involving human participants on premises other than those of the University will be expected to obtain ethical approval from any collaborating organisation as well as from the Ethical Approvals (Human Participants) Sub-Committee.

3.3 Visiting investigators

Investigators from outside the University who wish to carry out investigations involving human participants in the University will be expected to conform to the relevant sections of the University's Code of Practice and, as appropriate, submit their proposals through the Head of a University School Department to the Ethics Approvals (Human Participants) Sub-Committee for approval.

3.4 Exclusions from Code

a. Experimentation and anatomical examination in human morbid anatomy is strictly controlled by the 1984 Anatomy Act, under licence from the Secretary of State for Social Services and therefore falls outside the scope of the Code. Staff and students are advised that it is an offence to carry out dissection or experimentation on cadavers outside the control of a Licensed Teacher of Anatomy or in unlicensed premises.

b. Experimentation on animals is strictly regulated by the Home Office under the provisions of the Animals (Scientific Procedures) Act 1986 and also falls outside the scope of the Code. Staff are advised that it is an offence to carry out scientific work controlled by the Act without the appropriate licence or certificate.

c. It is not intended that the Code should apply to procedures undertaken as part of patient-care which are expected to contribute to the benefit of the individual participant.

4. Considerations Relating to Specific Types of Investigation

4.1 University class teaching experiments and demonstrations

Undergraduate or postgraduate students may be invited to participate in experiments or studies as a normal part of their programme, provided:

a. that they have the right to decline to participate in a particular procedure or, having accepted, to withdraw at any time;

b. that they are assured that neither declining nor agreeing to participate in a particular procedure will affect their academic assessment in any way;

c. that no coercion, actual or implied, or any financial inducement should be used to persuade students to participate.

4.2 Drug Studies

a. Drug studies on human participants involving new chemical entities or new combinations of drugs will need to be approved via the NHS Research Ethics Committees. Drug trials are strictly regulated by the MHRA and the University must have the appropriate licencing before any study of this nature can be carried out.

b. In the case of prescription drugs (i.e. not available over the counter), investigators should consult the checklist developed by the Ethics Approvals (Human Participants) Sub-Committee. This can be accessed at www.lboro.ac.uk/admin/committees/ethical/gn/iiupd.htm

4.3 Investigations involving contact with Human Body Fluids

All proposals for investigations involving contact with human body fluids should adhere to the Health and Safety Policy on Blood Borne Viruses, as drawn up by the Health Safety and Environment Committee, Loughborough University, and should make reference to this Policy within the submission form.

4.4 Investigations involving the use of Ionising Radiation (e.g. x-rays)

All investigators seeking approval for proposals involving the use of Ionising Radiation (e.g. x-rays) should contact the University Radiological Protection Officer for advice and should follow the guidance on Exposure to Ionising Radiation (www.lboro.ac.uk/admin/committees/ethical/gn/exir.htm) as drawn up by the Ethical Approvals (Human Participants) Sub-Committee, and should make reference to this guidance within the submission form.

4.5 Investigations involving the use of Hazardous Substances

All investigators seeking approval for proposals involving the use of hazardous substances should contact the Health, Safety and Environment Section for advice and should follow the Guidance on Exposure to Hazardous Substances (www.lboro.ac.uk/admin/committees/ethical/gn/exhs.htm) as drawn up by the Ethics Approvals (Human Participants) Sub-Committee, and should make reference to this guidance within the submission form.

5. Insurance

The University maintains in force a Public Liability Policy, which indemnifies it against its legal liability for accidental injury to persons (other than its employees) and for accidental damage to the property of others. Any unavoidable injury or damage therefore falls outside the scope of the policy.

The Insurance relates to claims arising out of all normal activities of the University, but Insurers require to be notified of anything of an unusual nature (see Section 7r). In particular, where tests on new drugs or equipment are sponsored by an external body, the trials may need to be covered by the insurance policy of the sponsoring organisation rather than the University.

6. Guidelines for Investigators

The following guidelines should be adhered to when making a full submission to the Ethics Approvals (Human Participants) Sub-Committee (i.e. if the study does not conform with all the sections on the Ethical Clearance Checklist).

a. Approval of Proposals

The University operates a two-tier system of ethical approval for investigations involving human participants. Approval may be obtained through completion of the Ethical Clearance Checklist OR a full submission to the Sub-Committee. Further information is available at the following webpage: www.lboro.ac.uk/admin/committees/ethical/#ov

The Ethical Approvals (Human Participants) Sub-Committee will advise and assist investigators where necessary on the design and conduct of such studies, to enable them to conform with the ethical guidelines set out below and obtain approval from the Sub-Committee for the work proposed.

b. Data Protection Act and Confidentiality

There should be an acknowledged obligation to protect the participants from possible harm and to preserve their right to privacy. The confidentiality of the participants should be maintained where appropriate and the investigator's intentions in the matter of confidentiality should be made known to the participants. Any investigator intending to process personal data should be made aware of and comply with the provisions of the Data Protection Act 1998. The University's Data Protection Policy can be found on the University's Data Protection Policy webpages. The Ethics Approvals (Human Participants) Sub-Committee has issued specific guidance to help investigators to comply with the requirements of the Data Protection Act which can be found at: www.lboro.ac.uk/admin/committees/ethical/gn/dp-comp.htm

c. Recruitment of Participants

The recruitment of participants should wherever possible be via a notice, or, if verbally, through a group approach rather than to individuals. Recruitment notices should clearly explain the scientific purpose of the research and details of what volunteers can expect if they agree to participate.

Staff or students of the department concerned may be invited to volunteer to take part, but special consideration should be given to the motives that might prompt them to volunteer. It is not normally desirable for students in close contact with a member of staff acting as investigator to be recruited, as they may feel vulnerable to pressure from someone in a position to influence their careers. On the other hand, it is normally reasonable for students to be recruited to take part in teaching exercises where one of the primary objectives is to enable them to make their own observations.

d. Vulnerable Groups

Recruitment from certain other groups may raise ethical issues which require special consideration. Certain groups may be incapable of giving valid consent, such as persons who lack capacity under the Mental Capacity Act, people detained under the Mental Health Act, prisoners, and people under the age of 18. An approach in such cases should be made to the authority or individual with legal responsibility for the participant. Special care should be taken in considering investigations involving the elderly and women of childbearing potential should not be recruited for any study which could be harmful to pregnancy. The Ethics Approvals (Human Participants) Sub-Committee has produced guidance on Working with Children and Young People which can be found at: www.lboro.ac.uk/admin/committees/ethical/gn/wwccop.htm. Investigators are advised to read the guidance carefully before embarking upon a research project which involves participants under the age of 18. Investigators should also read the guidance, produced by Personnel, to establish whether or not they need to seek Criminal Records Bureau clearance. This guidance can be found at: www.lboro.ac.uk/admin/personnel/recordchecks.html

e. Inclusion/Exclusion criteria

It is essential that the Ethics Approvals (Human Participants) Sub-Committee should be given full details of the basis for the selection of participants including any inclusion/exclusion criteria. Particular care should be taken to exclude participants who suffer from physical, physiological or emotional conditions which could be affected/aggravated by the proposed procedures. Submissions should include any questionnaire which is to be used in the selection process.

Where appropriate, participants should be asked about their previous medical history and be given advice on the relation of this to the proposed study. They should be asked to give permission to the investigator to contact their doctor and to authorise the doctor to release any details of their past medical history considered relevant: time should also be arranged to allow participants to consult their doctor before they agree to participate in the investigation. A generic Health Screen Questionnaire is available to download at: www.lboro.ac.uk/admin/committees/ethical/ and investigators are advised to modify (i.e. add or remove questions) it to suit their individual study.

f. Minimising Risks to Participants

No investigation involving human participants should involve more than minimal risk to their physical or mental well-being. All risks should be measured/weighed against the scientific benefit of the study. All risks should be fully explained to participants, including precautions taken to minimise those risks.

In certain circumstances, to minimise risk, the Ethics Approvals (Human Participants) Sub-Committee may require that a person with suitable medical qualifications should be responsible for an investigation or in attendance when certain procedures are carried out, or that facilities for emergency medical care should be at hand. Where appropriate, safeguards regarding communicable diseases should be taken to protect the participant, the investigator and others involved in the work. In all cases of venepuncture, a new sterile needle shall be used for each participant (see University Policy on Blood Borne Viruses).

g. Participant Information Sheets

Investigators should give each participant full details of the nature, object and duration of the proposed investigation in a form that is readily understood (this may be written or verbal depending on the targeted participants). The participant should be told what procedures the investigation will involve and whether any discomfort or inconvenience is likely to be entailed during the investigation or afterwards. Investigators should also provide information and advice about any foreseeable risks to health to which participants may be exposed. It is good practice to offer participants the opportunity to visit the location of the study, have procedures demonstrated and/or inspect/test equipment before the commencement of the investigation. This ensures that participants are fully informed about what will happen to them during the investigation.

h. Deception

There should be no deception that might affect a person's willingness to participate in an investigation, nor about the possible risks involved. It is recognised that some studies involve deception of the participant and would be invalid if this were not so. If any deception is considered necessary in a study, it should not involve the participant in any risk, such as unexpected anxiety or distress, lowering of self-esteem, or any form of long-term psychological or physical harm. Where deception is necessary, revelation should normally follow participation as a matter of course and should be designed into the experimental procedure.

i. Consent

The full, informed and voluntary consent of the participant must be obtained before the investigation begins; that is to say, consent freely given with proper understanding of the nature and consequences of what is proposed. In the cases of participants under the age of 18, or with some other potentially vulnerable groups, it may be necessary to obtain consent from the parent/guardian or carer.

The Ethics Approvals (Human Participants) Sub-Committee has produced a sample consent form which may be used/adapted by all investigators. This is available for download in Section 7 of the page: www.lboro.ac.uk/admin/committees/ethical/.

Written consent may be dispensed with only with the agreement of the Ethics Approvals (Human Participants) Sub-Committee.

j. Financial Incentives

There should be no excessive financial inducement that may cause coercion, actual or implied, and that might persuade people to take part in an investigation against their better judgement. Any payment made to volunteers should be for expenses, time, inconvenience or discomfort and never for hazard to the person. All payments to participants must be approved by the Ethics Approvals (Human Participants) Sub- Committee.

k. Withdrawal from Investigations

Participants must be free to withdraw from the investigation at any stage, without having to give any reasons, and should be told they have this right. However an opportunity should be provided in this event for participants to discuss privately their wish to withdraw. It is recognised that it may not always be possible to disaggregate data from the study once it has been anonymised and this should be clearly explained to participants at an early stage.

l. Issuing Advice to Participants

Investigators have a duty of care to participants. When planning research, investigators should consider what, if any, arrangements are needed to inform participants (or those legally responsible for the participants) of any health related (or other) problems previously unrecognised in the participant. This is particularly important if it is believed that by not doing so the participants well being is endangered. Investigators should consider whether or not it is appropriate to recommend that participants (or those legally responsible for the participants) seek qualified professional advice, but should not offer this advice personally.

m. Unexpected Damaging Consequences

Any unusual or unexpected symptoms arising or any significant untoward event affecting a participant during or after an investigation should be communicated promptly with the individual's consent to the participant's own doctor, and to the Ethics Approvals (Human Participants) Sub-Committee. The study should be stopped in the individual concerned and it should be considered whether it is advisable to stop the investigation as a whole. If a participant withdraws from an investigation, for whatever reason, the investigator should take reasonable steps to find out whether any harm has come to the individual as a result of participation in the study.

n. Completion of Investigations

The Ethics Approvals (Human Participants) Sub-Committee should be informed when a study has been completed unless the committee has removed this requirement. In particular, the Sub-Committee should be informed of any changes to the approved procedures.

o. Records of Investigations and Participants

The investigator should keep full records of all procedures carried out in a form appropriate for consultation by the Ethics Approvals (Human Participants) Sub- Committee and keep a register of participants used.

p. Location of Investigation

The places where investigations involving human participants are to be undertaken should be appropriate to the type of study and the risk involved. The Ethics Approvals (Human Participants) Sub-Committee may, at its discretion, request an inspection of the premises concerned.

q. New Equipment

Investigations involving testing new equipment on human participants should be undertaken in an appropriate location and a full risk analysis conducted to ensure that appropriate medical assistance is available if required. The Ethics Approvals (Human Participant) Sub-Committee may, at its discretion, request an inspection and/or demonstration of the new equipment before the commencement of the investigation.

r. Insurance

The Ethics Approvals (Human Participants) Sub-Committee checks some proposals with the University Insurance Officer to ensure that the University's insurance policy covers the submission subject to the usual terms and conditions. For some externally sponsored investigations, insurance cover will need to be provided by the sponsoring organisation. This is usually the case where new drugs or equipment are being tested. It is the responsibility of the applicant to arrange insurance cover for the project if it falls outside of the scope of the University's Public Liability Policy. Details of such cover should be included in the submission.

Participants should be told their position with regard to insurance cover in the event of an accident, injury, or ill-health befalling them as a result of taking part in the investigation.

7. Procedures for Submitting Full Proposals to the Ethics Approvals (Human Participants) Sub-Committee

a. What to Submit?

All protocols that do not comply with the Ethical Clearance Checklist shall be referred to the Ethics Approvals (Human Participants) Sub-Committee for consideration and approval. University class teaching exercises and demonstrations and student projects as well as research investigations can be referred to the Sub-Committee.

b. Obtaining Approval

Proposals should be prepared in accordance with Section 7 of this Code of Practice using the standard forms prepared by the Sub-Committee for this purpose, and submitted via the investigator's Head of School/Department to the Secretary of the Sub-Committee. The Head of School/Department should signify his/her awareness of the proposal being made.

Each proposal will be submitted for consideration at the next scheduled meeting of the Sub-Committee (see section e below) and the Sub-Committee will formally decide whether or not the proposal is acceptable. Exceptionally, where an earlier decision is required, the Chairperson, or his/her nominee from within the Sub-Committee, having consulted members (including the external members) as necessary, may decide whether or not a proposal is acceptable, and his/her decision will be reported to the Sub-Committee for ratification at its next meeting.

The decision of the Sub-Committee will be communicated to the investigators by the Secretary (by email).

c. Seeking Expert Guidance

The Sub-Committee expects from time to time to seek expert guidance or advice from outside its membership and it will proceed in this way in the event of failure to agree.

d. Reports to Senate

The Sub-Committee will inform the Ethics Committee of any instance where it has not been possible to reach an agreement with an investigator on a satisfactory protocol.

The Sub-Committee will submit an annual report to the Ethics Committee which will include a summary of all investigations approved by the Sub-Committee.

e. Dates of Meetings

The Sub-Committee will meet twelve times a year, three times a year to discuss ethical issues and consider proposals. There shall be a quorum at a Sub-Committee main meeting when at least five members are present.

The proposal-only meeting will occur once a month in each month that the main Sub- Committee does not meet (i.e. 9 times a year).

Dates of scheduled meetings can be found at: www.lboro.ac.uk/admin/committees/ethical/eac-m.htm

Ethical Consent form (Completed)

Commented [G1]: Where? Insert

Participant Consent form

Loughborough
University

Effect of Increasing Workload on Potential Remotely Piloted System Operators

INFORMED CONSENT FORM

(to be completed after Participant Information Sheet has been read)

The purpose and details of this study have been explained to me. I understand that this study is designed to further scientific knowledge and that all procedures have been approved by the Loughborough University Ethical Approvals (Human Participants) Sub-Committee.

I have read and understood the information sheet and this consent form.

I have had an opportunity to ask questions about my participation.

I understand that I am under no obligation to take part in the study.

I understand that I have the right to withdraw from this study at any stage for any reason, and that I will not be required to explain my reasons for withdrawing.

I understand that all the information I provide will be treated in strict confidence and will be kept anonymous and confidential to the researchers unless (under the statutory obligations of the agencies which the researchers are working with), it is judged that confidentiality will have to be breached for the safety of the participant or others.

I agree to participate in this study.

Your name

Your signature

Signature of investigator

Date

Participant information sheet

Effect of Increasing Workload on Potential Remotely Piloted System Operators

Participant Information Sheet

Mr. George Bedford, g.bedford@lboro.ac.uk Tel: 01509635674

Professor Roy S. Kalawsky, r.s.kalawsky@lboro.ac.uk Tel: 01509 635678

Advanced Virtual Reality Research Centre
Loughborough University
Loughborough
Leicestershire, LE11 3TU

What is the purpose of the study?

This investigation is designed to capture data and interpretations of experimental participants in relation to information set usage, workload and cognitive ability while operating a simulated representation of a Remotely Piloted Aerial System. It is designed to support research into the field of operator selection for Remotely Pilot systems.

Who is doing this research and why?

This study is part of a doctoral research project, supported by the EPSRC and BAE systems; it is focused on understanding the differences between two groups of potential Remotely Piloted Aerial Systems operators and will investigate the two group’s performance relating to information set usage, work load management and cognitive ability associated with prior computer games experience.

Are there any exclusion criteria?

Participants must have be either experienced gamers or non-gamers and have no physical or psychological disabilities.

Once I take part, can I change my mind?

Yes! After you have read this information and asked any questions you may have we will ask you to complete an Informed Consent Form, however if at any time, before, during or after the sessions you wish to withdraw from the study please just contact the main investigator. You can withdraw at any time, for any reason and you will not be asked to explain your reasons for withdrawing.

Will I be required to attend any sessions and where will these be?

You will be required to attend a single session at Loughborough Universities AVRRC.

How long will it take?

The session will last approximately 2 hours with no further sessions required

Is there anything I need to do before the sessions?

Complete the questionnaire linked in the e-mail or use the link below

https://docs.google.com/forms/d/1riTGSwIKUZfRIpzTla8R-pFv2LxtsZVxVii6kfuSBvY/viewform

What type of clothing should I wear?

Any form of clothing will be fine, although it is recommended to wear something comfortable.

Who should I send the questionnaire back to?

What will I be asked to do?

The experimentation will begin with the participant having a two part familiarisation session with the simulator; this will be followed by a short questionnaire and a break.

The participant will then be asked to complete nine, five minute, object spotting flights with increasing degrees of workload; breaks will be provided between each set of three flights. Each flight will contain a calibration step, the flight itself and finish with two questionnaires. An hour break for lunch will also be provided if applicable.

On completion of the nine flights the participant will be allowed a break and then will be asked to complete a final questionnaire

The participant will initially be allowed to voice any concerns or opinions he has regarding the forth coming experimentation and will be asked to complete the consent forms.

What personal information will be required from me?

Only general and anonymous information regarding age, gender, education/work background, medical (if relevant to participant acceptance) and computer games experience.

Are there any risks in participating?

No

Will my taking part in this study be kept confidential?

Yes if required. All acquired data and audio recordings will be kept in accordance with the Data Protection Act 1998 on a secure computer located at Loughborough University; all data and recordings will be strictly confidential. All data and audio recordings will be destroyed after six years of the completion of the PhD.

What will happen to the results of the study?

What do I get for participating?

£50 worth of Amazon vouchers will be awarded randomly by raffle to one of the participants on completion of all experimentation.

I have some more questions who should I contact?

Mr. George Bedford, g.bedford@lboro.ac.uk Tel: 01509635674

What if I am not happy with how the research was conducted?

If you are not happy with how the research was conducted, please contact the Mrs Zoe Stockdale, the Secretary for the University's Ethics Approvals (Human Participants) Sub-Committee:

Mrs Z Stockdale, Research Office, Rutland Building, Loughborough University, Epinal Way, Loughborough, LE11 3TU. Tel: 01509 222423. Email: Z.C.Stockdale@lboro.ac.uk

The University also has a policy relating to Research Misconduct and Whistle Blowing which is available online at http://www.lboro.ac.uk/admin/committees/ethical/Whistleblowing(2).htm. Please ensure that this link is included on the Participant Information Sheet.

Questionnaire A (Pre-Study)

Effect of Increasing Workload on Potential Remotely Piloted System Operators

Questionnaire A

The information gathered by this questionnaire is anonymous and strictly confidential, if you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions. Submission of this questionnaire does not automatically constitute guaranteed acceptance as a research participant.

Please answer the following questions truthfully and to the best of your ability

Participant Information:

1. What gender are you?

2. What is your age?

3. What is your educational background?

4. What is your current employment

5. Do you or have you ever played computer games (if no then do not proceed to further questions)?

6. Please list the types of computer games you most often play (for example: First person Shooter, Role Play, Strategy, Simulation such as Gran Turismo or X-Plane, MMO)

7. Please estimate the amount of time you currently play computer games per week

8. Has this amount of time been the norm for the last year?

9. When was the last time you played approximately more than 3 hours of computer games per week?

10. Please estimate how long you have been playing computer games

11. Would you be available for experimentation during the period of….?

12. If not available during the period above please list your availability

13. Do you have any physical or psychological disabilities such as epilepsy, limited movement, visual impairment beyond correctable solutions (please list if applicable)?

Questionnaire B (Familiarisation)

Effect of Increasing Workload on Potential Remotely Piloted System Operators

Questionnaire B (Simulator Familiarisation)

The information gathered by this questionnaire is anonymous and strictly confidential, if you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions.

Please answer the following questions truthfully and to the best of your ability

Do you understand the purposes of this experimentation?

Do you find the simulator comfortable? (if no then please explain why)

Do you find using all of the available displays comfortable? (if no then please explain why)

Do you find the head tracker arrangement comfortable? (if no then please explain why)

Do you find the camera controls intuitive? (if no then please explain why)

Do you find the flight controls intuitive? (if no then please explain why)

Do you feel that the camera control interface has been adequately explained to you? (if no then please explain why)

Do you feel that the flight control interface has been adequately explained to you? (if no then please explain why)

Do you feel confident in the use of the camera controls for object location? 1-10

Do you feel confident in the use of the flight controls for maintaining platform stability? 1-10

Are there any changes to the simulator that you would like to be made before experimentation begins? (if yes then please explain what and why)

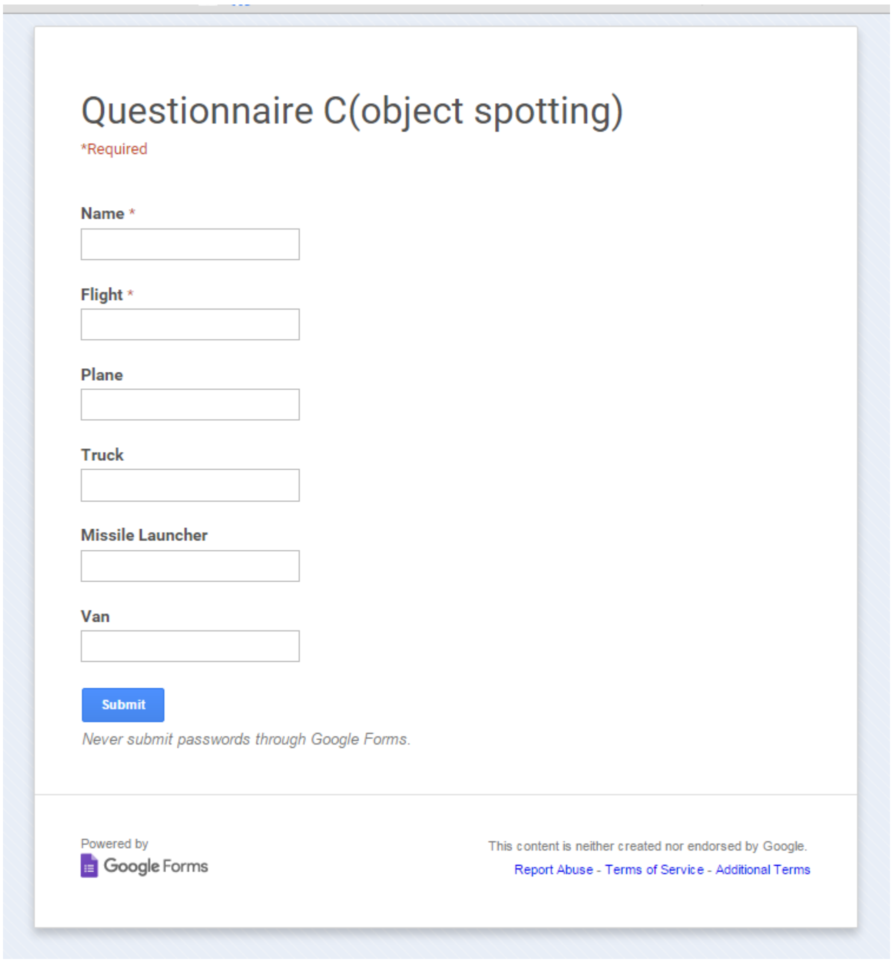
Are there any issues with the simulator interface that you feel would impact on your performance? (if yes then please explain what and why)

Are you still happy with proceeding with experimentation? (if no then please explain why)

Questionnaire C (Objects)

Effect of Increasing Workload on Potential Remotely Piloted System Operators

Questionnaire C (Object Spotting)



Questionnaire C(object spotting)

*Required

Name *

Flight *

Plane

Truck

Missile Launcher

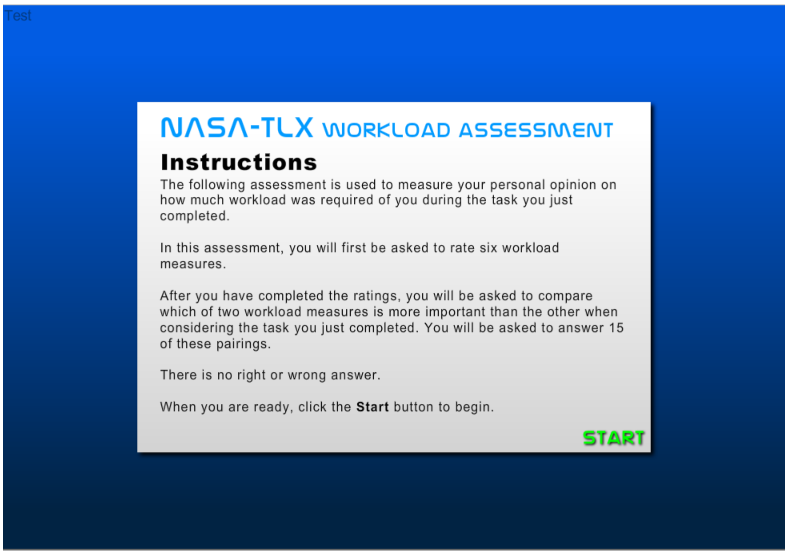
Van

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google.

NASA TLX

Step 1



NASA-TLX WORKLOAD ASSESSMENT

Instructions

The following assessment is used to measure your personal opinion on how much workload was required of you during the task you just completed.

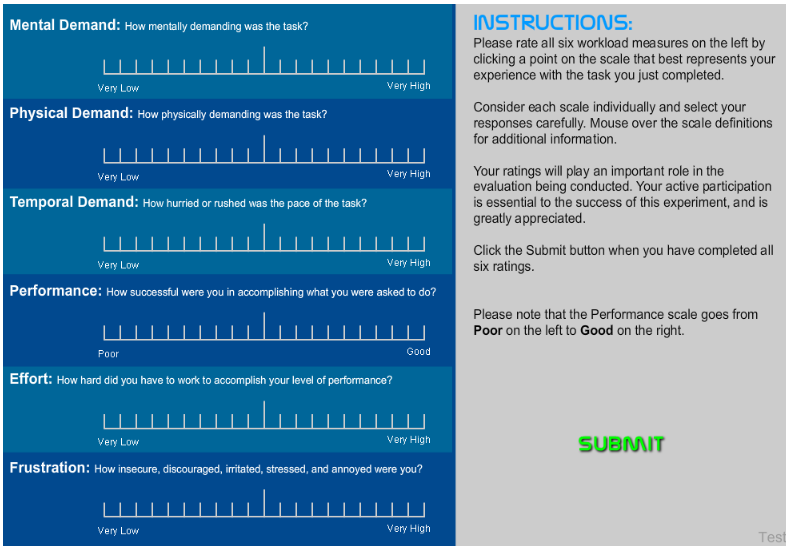
In this assessment, you will first be asked to rate six workload measures.

After you have completed the ratings, you will be asked to compare which of two workload measures is more important than the other when considering the task you just completed. You will be asked to answer 15 of these pairings.

There is no right or wrong answer.

When you are ready, click the **Start** button to begin.

Step 2



Mental Demand: How mentally demanding was the task?

Physical Demand: How physically demanding was the task?

Temporal Demand: How hurried or rushed was the pace of the task?

Performance: How successful were you in accomplishing what you were asked to do?

Effort: How hard did you have to work to accomplish your level of performance?

Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

INSTRUCTIONS:

Please rate all six workload measures on the left by clicking a point on the scale that best represents your experience with the task you just completed.

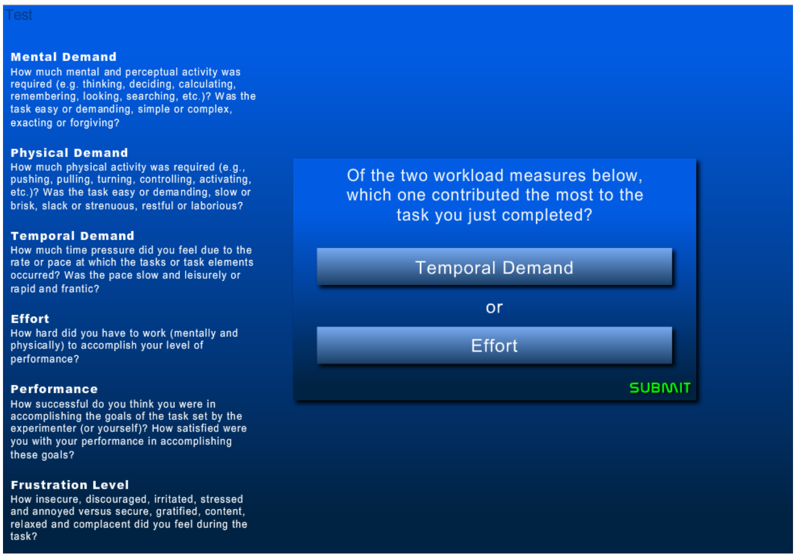
Consider each scale individually and select your responses carefully. Mouse over the scale definitions for additional information.

Your ratings will play an important role in the evaluation being conducted. Your active participation is essential to the success of this experiment, and is greatly appreciated.

Click the Submit button when you have completed all six ratings.

Please note that the Performance scale goes from **Poor** on the left to **Good** on the right.

Step 3



Mental Demand
How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Physical Demand
How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand
How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Effort
How hard did you have to work (mentally and physically) to accomplish your level of performance?

Performance
How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Frustration Level
How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Of the two workload measures below, which one contributed the most to the task you just completed?

Final Output

userID	experimental ID	participant ID	tlx_Score
g.bedford@lboro.ac.uk	Low1	Test	49.53

Questionnaire D (Debrief)

The information gathered by this questionnaire is anonymous and strictly confidential, if you have any questions or issues with any of the questions within this questionnaire please contact George Bedford (g.bedford@lboro.ac.uk). This questionnaire is voluntary and you may decide not to answer any of the following questions.

* Required

Please insert your first name and surname initial *

example: GeorgeB

Were you satisfied with your experimental performance? * Scale 1-10

Do you feel that you became better at flight performance as you progressed through experimentation? *

Please indicate how easy the camera was to operate * Scale 1-10

Please indicate how easy the flight controls were to operate * Scale 1-10

Please indicate how easy it was to identify objects * Scale 1-10

Please indicate how easy it was to identify object numbers * Scale 1-10

Do you feel that there was a workload increase across all three flight sets? *

Please indicate the degree by which increased workload impaired your ability to locate and identify objects * Scale 1-10

Which set of flights did you find required the most effort? *

Please list in descending order. Example: High, Medium, Low

Do you feel that the field of vision restriction impaired your performance? *

Please indicate the degree by which field of vision restriction impaired your ability to maintain platform stability * Scale 1-10

Please indicate the degree by which field of vision restriction impaired your ability to locate and identify objects * Scale 1-10

Do you feel the simulator scenery adequately matched that of the Google Earth display? *

If no to the previous question please identify why

Do you feel you were able to adequately orient the objects location from the camera using the search area on the Google Earth display? *

If no to the previous question please identify why

Do you feel physically tired? * Scale 1-10

Do you feel mentally stressed? * Scale 1-10

Do you have any other physical discomfort relating to the experimentation just performed? *

If yes to the previous question please identify why

Do you feel you had enough time to familiarise yourself with the simulator before experimentation? *

Do you feel that simulation was: *

too long

too short

adequate

Where the objects of a size and definition that was adequate? *

Is there anything you would change about the simulator that you feel would aid your performance?

Is there anything you would change about the simulator that would aid comfort?

Thank you! You have now completed the experiment

[Intentionally left blank]

Appendix F - Experimental Data

Participants

Group A & B:

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-25	12	40.0	40.0	40.0
	26-30	13	43.3	43.3	83.3
	31-35	5	16.7	16.7	100.0
	Total	30	100.0	100.0	

Descriptive Statistics

	N	Mean	Std. Deviation
Age	30	26.23	4.309
Valid N (listwise)	30		

Employment

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Proffesional	14	46.7	46.7	46.7
	Student	15	50.0	50.0	96.7
	Unemployed	1	3.3	3.3	100.0
	Total	30	100.0	100.0	

Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Further	4	13.3	13.3	13.3
	Higher	26	86.7	86.7	100.0
	Total	30	100.0	100.0	

Game Related Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Play Estimation (hr)	30	0	25.00	7.6383	8.66960
How Long Playing Total (yrs)	30	0.2	28.00	16.0400	6.50356
How Many Weeks Since	30	0	999	154.43	275.147
Valid N (listwise)	30				

Group A (G):

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-25	7	46.7	46.7	46.7
	26-30	7	46.7	46.7	93.3
	31-35	1	6.7	6.7	100.0
	Total	15	100.0	100.0	

Descriptive Statistics

	N	Mean	Std. Deviation
Age	15	25.30	3.886
Valid N (listwise)	15		

Employment

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Proffesional	4	26.7	26.7	26.7
	Student	10	66.7	66.7	93.3
	Unemployed	1	6.7	6.7	100.0
	Total	15	100.0	100.0	

Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Further	1	6.7	6.7	6.7
	Higher	14	93.3	93.3	100.0
	Total	15	100.0	100.0	

Game Related Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Play Estimation (hr)	15	5.00	25.00	14.6667	7.00680
How Long Playing Total (yrs)	15	6.00	25.00	16.5333	5.35679
How Many Weeks Since	15	0	1	0.33	0.488
Valid N (listwise)	15				

Group B (NG):

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-25	5	33.3	33.3	33.3
	26-30	6	40.0	40.0	73.3
	31-35	4	26.7	26.7	100.0
	Total	15	100.0	100.0	

Descriptive Statistics

	N	Mean	Std. Deviation
Age	15	27.17	4.636
Valid N (listwise)	15		

Employment

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Proffesional	10	66.7	66.7	66.7
	Student	5	33.3	33.3	100.0
	Total	15	100.0	100.0	

Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Further	3	20.0	20.0	20.0
	Higher	12	80.0	80.0	100.0
	Total	15	100.0	100.0	

Game Related Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Play Estimation (hr)	15	0	3.00	0.61	0.86359
How Long Playing Total (yrs)	15	0.2	28.00	15.5467	7.64179
How Many Weeks Since	15	0	999	308.53	325.473
Valid N (listwise)	15				

Normality Test Tables

Group G Low Workload Flights and Average

G

	Group		Significance (Shapiro-Wilk)				Trans
		df	Low1	Low2	Low3	LowAvg	LowAvg
PitchV	G	15	.000**	.000**	.000**	.000**	-
PitchSD	G	15	.000**	.264	.001**	.000**	-
BankV	G	15	.000**	.838	.002**	.307	-
BankSD	G	15	.001**	.712	.002**	.141	-
SpdV	G	15	.000**	.000**	.000**	.000**	-
SpdSD	G	15	.000**	.000**	.000**	.000**	-
AltV	G	15	.000**	.000**	.000**	.000**	-
AltSD	G	15	.000**	.000**	.000**	.001**	-
HDGV	G	15	.000**	.247	.000**	.000**	-
HDGSD	G	15	.000**	.264	.001**	.000**	-
VS_V	G	15	.000**	.055	.960	.015**	-
VS_SD	G	15	.000**	.177	.894	.035**	-
Screen2P	G	15	.124	.726	.572	.083	-
Screen2C	G	15	.058	.634	.200	.038**	-
Screen2Dwell	G	15	.390	.302	.063	.310	-
Screen3P	G	15	.264	.003**	.402	.010**	-
Screen3C	G	15	.003**	.002**	.229	.000**	-
Screen3Dwell	G	15	.006**	.110	.028**	.077	-
Screen4P	G	15	.000**	.000**	.000**	.000**	-
Screen4C	G	15	.000**	.000**	.000**	.000**	-
Screen4Dwell	G	15	.000**	.000**	.000**	.000**	-
VoidP	G	15	.284	.008**	.194	.127	-
VoidC	G	15	.125	.348	.160	.364	-
VoidDwell	G	15	.016	.115	.015	.276	-
Time	G	15	.076	.515	.034	.178	-
TLXscore	G	15	.189	.540	.160	.505	-
Score	G	15	.000**	.000**	.000**	.000**	-
TScore	G	15	.012	.012	.012	.012	-

*. This is a lower bound of the true significance.
**. No Significance (p<0.05)
a. Lilliefors Significance Correction

Group G Medium Workload Flights and Average

	Group		Significance (Shapiro-Wilk)				Trans
		df	Med1	Med2	Med3	MedAvg	MedAvg
PitchV	G	15	.000**	.000**	.000**	.000**	-
PitchSD	G	15	.011**	.055	.287	.000**	-
BankV	G	15	.000**	.000**	.013**	.000**	-
BankSD	G	15	.000**	.000**	.299	.003**	.502
SpdV	G	15	.000**	.000**	.000**	.000**	-
SpdSD	G	15	.000**	.000**	.000**	.000**	-
AltV	G	15	.000**	.000**	.839	.045**	-
AltSD	G	15	.000**	.000**	.895	.014**	.372
HDGV	G	15	.000**	.003**	.013**	.000**	-
HDGSD	G	15	.011**	.055	.287	.003**	.185
VS_V	G	15	.000**	.000**	.000**	.000**	-
VS_SD	G	15	.000**	.000**	.000**	.000**	-
Screen2P	G	15	.601	.186	.718	.761	-
Screen2C	G	15	.349	.015**	.001**	.313	-
Screen2Dwell	G	15	.571	.055	.538	.930	-
Screen3P	G	15	.002**	.336	.566	.048**	-
Screen3C	G	15	.107	.865	.576	.760	-
Screen3Dwell	G	15	.197	.309	.424	.434	-
Screen4P	G	15	.450	.974	.509	.616	-
Screen4C	G	15	.785	.163	.029**	.924	-
Screen4Dwell	G	15	.172	.037**	.589	.247	-
VoidP	G	15	.015**	.179	.211	.203	-
VoidC	G	15	.973	.038**	.011**	.325	-
VoidDwell	G	15	.006**	.268	.269	.021**	-
Time	G	15	.341	.401	.574	.574	-
TLXscore	G	15	.429	.758	.442	.214	-
Score	G	15	.000**	.000**	.000**	.001**	
TScore	G	15	.012**	.012**	.012**	.012	-

*. This is a lower bound of the true significance.
**. No Significance (p<0.05)
a. Lilliefors Significance Correction

Group G High Workload

	Group		Significance (Shapiro-Wilk)				Trans
		df	High1	High2	High3	HighAvg	HighAvg
PitchV	G	15	.000**	.000**	.001**	.000**	-
PitchSD	G	15	.020**	.001**	.020**	.001**	.094
BankV	G	15	.000**	.000**	.001**	.000**	-
BankSD	G	15	.003**	.002**	.093	.010**	.781
SpdV	G	15	.000**	.000**	.000**	.000**	-
SpdSD	G	15	.000**	.000**	.000**	.001**	-
AltV	G	15	.000**	.058	.000**	.259	-
AltSD	G	15	.000**	.073	.000**	.093	.455
HDGV	G	15	.000**	.000**	.002**	.001**	-
HDGSD	G	15	.020**	.001**	.020**	.003**	.165
VS_V	G	15	.000**	.000**	.001**	.000**	-
VS_SD	G	15	.001**	.001**	.068	.002**	.155
Screen2P	G	15	.321	.870	.434	.517	-
Screen2C	G	15	.869	.135	.612	.103	-
Screen2Dwell	G	15	.619	.443	.854	.791	-
Screen3P	G	15	.458	.033**	.301	.901	-
Screen3C	G	15	.764	.001**	.856	.356	-
Screen3Dwell	G	15	.058	.327	.167	.651	-
Screen4P	G	15	.478	.935	.348	.198	-
Screen4C	G	15	.013**	.676	.243	.254	-
Screen4Dwell	G	15	.999	.000**	.043**	.242	-
VoidP	G	15	.030**	.002**	.033**	.011**	-
VoidC	G	15	.341	.325	.826	.114	-
VoidDwell	G	15	.000**	.000**	.576	.003**	-
TIme	G	15	.446	.393	.314	.549	-
TLXscore	G	15	.261	.324	.073	.356	-
Score	G	15	.000**	.000**	.000**	.003**	
TScore	G	15	.012**	.012**	.012**	.012**	-

*. This is a lower bound of the true significance.
**. No Significance (p<0.05)
a. Lilliefors Significance Correction

Group NG Low Workload Flights and Average

	Group		Significance (Shapiro-Wilk)				Trans
		df	Low1	Low2	Low3	LowAvg	LowAvg
PitchV	NG	15	.000**	.000**	.000**	.001**	-
PitchSD	NG	15	.000**	.000**	.000**	.018**	-
BankV	NG	15	.166	.004**	.498	.024**	-
BankSD	NG	15	.456	.011**	.587	.052	-
SpdV	NG	15	.000**	.000**	.000**	.000**	-
SpdSD	NG	15	.000**	.000**	.000**	.001**	-
AltV	NG	15	.000**	.000**	.000**	.001**	-
AltSD	NG	15	.000**	.000**	.000**	.008**	-
HDGV	NG	15	.000**	.000**	.000**	.000**	-
HDGSD	NG	15	.000**	.000**	.000**	.001**	-
VS_V	NG	15	.000**	.000**	.164	.000**	-
VS_SD	NG	15	.000**	.000**	.275	.000**	-
Screen2P	NG	15	.111	.235	.632	.449	-
Screen2C	NG	15	.644	.320	.127	.513	-
Screen2Dwell	NG	15	.043**	.138	.429	.374	-
Screen3P	NG	15	.729	.007**	.775	.099	-
Screen3C	NG	15	.110	.002**	.589	.145	-
Screen3Dwell	NG	15	.232	.431	.284	.667	-
Screen4P	NG	15	.000**	.000**	.000**	.000**	-
Screen4C	NG	15	.000**	.000**	.000**	.000**	-
Screen4Dwell	NG	15	.000**	.000**	.000**	.000**	-
VoidP	NG	15	.013**	.245	.005**	.533	-
VoidC	NG	15	.356	.397	.221	.435	-
VoidDwell	NG	15	.050	.517	.006**	.557	-
TIme	NG	15	.673	.886	.350	.744	-
TLXscore	NG	15	.090	.217	.022**	.107	-
Score	NG	15	.000**	.000**	.000**	.001**	-
TScore	NG	15	.562	.562	.562	.562	-

*. This is a lower bound of the true significance.
**. No Significance (p<0.05)
a. Lilliefors Significance Correction

Group NG Medium Workload Flights and Average

	Group		Significance (Shapiro-Wilk)				Trans
		df	Med1	Med2	Med3	MedAvg	MedAvg
PitchV	NG	15	.000**	.000**	.000**	.000**	-
PitchSD	NG	15	.000**	.014**	.000**	.000**	-
BankV	NG	15	.000**	.000**	.000**	.000**	-
BankSD	NG	15	.000**	.000**	.017**	.022**	.366
SpdV	NG	15	.000**	.000**	.000**	.000**	-
SpdSD	NG	15	.000**	.000**	.000**	.000**	-
AltV	NG	15	.000**	.000**	.019**	.000**	-
AltSD	NG	15	.000**	.000**	.019**	.000**	.222
HDGV	NG	15	.000**	.000**	.000**	.000**	-
HDGSD	NG	15	.000**	.014**	.000**	.001**	.579
VS_V	NG	15	.000**	.000**	.000**	.000**	-
VS_SD	NG	15	.000**	.000**	.000**	.000**	-
Screen2P	NG	15	.934	.616	.335	.801	-
Screen2C	NG	15	.239	.506	.140	.070	-
Screen2Dwell	NG	15	.436	.052**	.036	.217	-
Screen3P	NG	15	.132	.417	.940	.409	-
Screen3C	NG	15	.785	.041**	.862	.320	-
Screen3Dwell	NG	15	.817	.534	.997	.019**	-
Screen4P	NG	15	.856	.038**	.033**	.212	-
Screen4C	NG	15	.490	.036**	.738	.792	-
Screen4Dwell	NG	15	.673	.353	.489	.961	-
VoidP	NG	15	.048**	.009**	.333	.114	-
VoidC	NG	15	.143	.699	.067	.008**	-
VoidDwell	NG	15	.002**	.002**	.198	.063	-
TIme	NG	15	.086	.046**	.614	.020**	-
TLXscore	NG	15	.057	.104	.139	.687	-
Score	NG	15	.000**	.000**	.000**	.408	-
TScore	NG	15	.562	.562	.562	.562	-

*. This is a lower bound of the true significance.
**. No Significance (p<0.05)
a. Lilliefors Significance Correction

Group NG High Workload Flights and Average

	Group		Significance (Shapiro-Wilk)				Trans
		df	High1	High2	High3	HighAvg	HighAv
PitchV	NG	15	.000**	.000**	.000**	.000**	-
PitchSD	NG	15	.001**	.000**	.000**	.003**	.056
BankV	NG	15	.000**	.000**	.000**	.000**	-
BankSD	NG	15	.003**	.000**	.001**	.024**	.265
SpdV	NG	15	.000**	.000**	.000**	.000**	-
SpdSD	NG	15	.000**	.000**	.000**	.000**	-
AltV	NG	15	.000**	.001**	.000**	.000**	-
AltSD	NG	15	.000**	.021**	.000**	.002**	.223
HDGV	NG	15	.000**	.000**	.000**	.000**	-
HDGSD	NG	15	.001**	.000**	.000**	.002**	.170
VS_V	NG	15	.000**	.000**	.000**	.000**	-
VS_SD	NG	15	.001**	.000**	.000**	.008**	.266
Screen2P	NG	15	.046**	.660	.838	.409	-
Screen2C	NG	15	.147	.075	.242	.046**	-
Screen2Dwell	NG	15	.003**	.032**	.114	.190	-
Screen3P	NG	15	.190	.285	.903	.621	-
Screen3C	NG	15	.001**	.033**	.293	.088	-
Screen3Dwell	NG	15	.598	.143	.442	.069	-
Screen4P	NG	15	.313	.645	.268	.533	-
Screen4C	NG	15	.038**	.001**	.495	.073	-
Screen4Dwell	NG	15	.527	.001**	.203	.424	-
VoidP	NG	15	.313	.152	.007**	.455	-
VoidC	NG	15	.005**	.019**	.107	.073	-
VoidDwell	NG	15	.003**	.108	.004**	.169	-
TIme	NG	15	.002**	.004**	.026**	.006**	-
TLXscore	NG	15	.806	.912	.169	.232	-
Score	NG	15	.005**	.000**	.007**	.290	-
TScore	NG	15	.562	.562	.562	.562	-

*. This is a lower bound of the true significance.
**. No Significance (p<0.05)
a. Lilliefors Significance Correction

Gamer Paired Screen Percentage Normality Test

Flight	Paired Screen Percentage (significance - p)					
	2 to 3	2 to 4	3 to 2	3 to 4	4 to 2	4 to 3
Low1	.037**	-	.017**	-	-	-
Low2	.101	-	.094	-	-	-
Low3	.004**	-	.007**	-	-	-
LowAv	**.672**	**-**	**.167**	**-**	**-**	**-**
Med1	.151	.453	.000**	.473	.806	.253
Med2	.136	.312	.638	.001**	.328	.247
Med3	.570	.246	.210	.068	.945	.341
MedAv	**.009****	**.154**	**.032****	**.898**	**.491**	**.909**
High1	.546	.505	.119	.120	.105	.104
High2	.221	.094	.317	.193	.250	.426
High3	.794	.607	.105	.383	.013**	.740
HighAv	**.429**	**.590**	**.320**	**.353**	**.002****	**.639**

** $p < 0.05$ Shapiro-Wilk non-significant/non-normal

Non-Gamer Paired Screen Percentage Normality Test

Flight	Paired Screen Percentage (significance - p)					
	2 to 3	2 to 4	3 to 2	3 to 4	4 to 2	4 to 3
Low1	.002**	-	.007**	-	-	-
Low2	.055	-	.149	-	-	-
Low3	.001**	-	.001**	-	-	-
LowAv	**.733**	**-**	**.638**	**-**	**-**	**-**
Med1	.007**	.264	.022**	.632	.148	.248
Med2	.635	.469	.116	.002**	.607	.164
Med3	.486	.745	.721	.989	.550	.621
MedAv	**.973**	**.289**	**.904**	**.040****	**.423**	**.556**
High1	.362	.723	.509	.417	.956	.051
High2	.475	.606	.451	.002**	.639	.001**
High3	.649	.683	.948	.930	.398	.351
HighAv	**.085**	**.372**	**.169**	**.208**	**.755**	**.196**

** $p < 0.05$ Shapiro-Wilk non-significant/non-normal

Descriptive Statistics Tables

Low Workload

Descriptive Statistics

	Group	Mean	Std. Deviation	N
Screen2P	G	.7065	.07119	15
	NG	.7096	.07279	15
	Total	.7080	.07076	30
Screen2Dwell	G	6.7493	2.34111	15
	NG	7.2316	2.08708	15
	Total	6.9905	2.19292	30
Screen3Dwell	G	1.8214	.75213	15
	NG	1.8527	.36651	15
	Total	1.8371	.58155	30
Av_2to3	G	.3458	.09245	15
	NG	.3471	.09415	15
	Total	.3464	.09168	30
Av_3to2	G	.3664	.08047	15
	NG	.3563	.09398	15
	Total	.3614	.08612	30
TLXscore	G	29.4153	17.82070	15
	NG	36.5491	23.80043	15
	Total	32.9822	20.97471	30

N: Number of Participants

Medium Workload

Descriptive Statistics

	Group	Mean	Std. Deviation	N
Screen2P	G	.4337	.07366	15
	NG	.4178	.10677	15
	Total	.4257	.09049	30
Screen4P	G	.3371	.06989	15
	NG	.3165	.06471	15
	Total	.3268	.06700	30
Screen2Dwell	G	3.3679	.87242	15
	NG	3.4440	1.06782	15
	Total	3.4059	.95885	30
Screen4Dwell	G	3.3926	1.20438	15
	NG	3.3613	.79099	15
	Total	3.3769	1.00128	30
Av_2to4	G	.2450	.06738	15
	NG	.2144	.05422	15
	Total	.2297	.06208	30
Av_4to2	G	.2152	.07230	15
	NG	.2114	.04751	15
	Total	.2133	.06014	30
Av_4to3	G	.0967	.03513	15
	NG	.0850	.04823	15
	Total	.0908	.04188	30
BankSDNorm	G	.1445	.08080	15
	NG	.0813	.05602	15
	Total	.1129	.07548	30
HDGDNorm	G	.0538	.01483	15
	NG	.0404	.02579	15
	Total	.0471	.02176	30
TLXscore	G	47.4709	17.78887	15
	NG	49.8202	20.11207	15
	Total	48.6456	18.69404	30

High Workload

Descriptive Statistics

	Group	Mean	Std. Deviation	N
Screen2P	G	.4295	.10450	15
	NG	.4040	.12075	15
	Total	.4168	.11171	30
Screen2Dwell	G	3.7153	1.10582	15
	NG	3.6602	1.18077	15
	Total	3.6877	1.12436	30
Screen3P	G	.0843	.01678	15
	NG	.0738	.02277	15
	Total	.0790	.02038	30
Screen3Dwell	G	1.1525	.20462	15
	NG	1.2222	.19113	15
	Total	1.1873	.19775	30
Screen4P	G	.4108	.10116	15
	NG	.4266	.11458	15
	Total	.4187	.10650	30
Screen4Dwell	G	5.3567	2.83225	15
	NG	5.4088	2.12999	15
	Total	5.3828	2.46241	30
Av_2to3	G	.1613	.06391	15
	NG	.1460	.06068	15
	Total	.1537	.06172	30
Av_2to4	G	.2289	.05360	15
	NG	.2449	.05227	15
	Total	.2369	.05265	30
Av_3to2	G	.1729	.06570	15
	NG	.1596	.07518	15
	Total	.1662	.06970	30
Av_3to4	G	.0997	.02592	15
	NG	.0816	.05371	15
	Total	.0906	.04245	30
Av_4to3	G	.1123	.04738	15
	NG	.0924	.04791	15
	Total	.1023	.04789	30
TLXscore	G	48.4353	10.18209	15
	NG	55.6553	20.50618	15
	Total	52.0453	16.32584	30
PitchSDNorm	G	.0584	.02335	15
	NG	.0310	.02461	15
	Total	.0447	.02738	30
BankSDNorm	G	.1430	.06692	15
	NG	.0699	.04258	15
	Total	.1064	.06648	30
AltSDNormHA	G	.0115	.00270	15
	NG	.0103	.00366	15
	Total	.0109	.00321	30
HDGDNorm	G	.0584	.02335	15
	NG	.0310	.02461	15
	Total	.0447	.02738	30
VS_SDNormHA	G	.0199	.00822	15
	NG	.0119	.00785	15
	Total	.0159	.00889	30

Screen Correlation – Spearman's Rho

Gamers

Correlations

			Screen2P	Screen3P	Screen4P
Low Workload	Screen2Dwell	Correlation Coefficient	**.864**[**]	-.239	-
		Sig. (2-tailed)	**.000**	.390	-
		N	15	15	-
	Screen3Dwell	Correlation Coefficient	.111	**.657**[**]	-
		Sig. (2-tailed)	.694	**.008**	-
		N	15	15	-
	Screen4Dwell	Correlation Coefficient	-	-	-
		Sig. (2-tailed)	-	-	-
		N	-	-	-
Medium Workload	Screen2Dwell	Correlation Coefficient	**.796**[**]	.225	-.332
		Sig. (2-tailed)	**.000**	.420	.226
		N	15	15	15
	Screen3Dwell	Correlation Coefficient	.314	**.721**[**]	-.200
		Sig. (2-tailed)	.254	**.002**	.475
		N	15	15	15
	Screen4Dwell	Correlation Coefficient	**-.661**[**]	.336	**.689**[**]
		Sig. (2-tailed)	**.007**	.221	**.004**
		N	15	15	15
High Workload	Screen2Dwell	Correlation Coefficient	**.579**[*]	-.329	-.321
		Sig. (2-tailed)	**.024**	.232	.243
		N	15	15	15
	Screen3Dwell	Correlation Coefficient	.225	**.671**[**]	-.186
		Sig. (2-tailed)	.420	**.006**	.508
		N	15	15	15
	Screen4Dwell	Correlation Coefficient	**-.529**[*]	-.071	**.682**[**]
		Sig. (2-tailed)	**.043**	.800	**.005**
		N	15	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Non-Gamers

Correlations

			Screen2P	Screen3P	Screen4P
Low Workload	Screen2Dwell	Correlation Coefficient	**.836****	.086	-
		Sig. (2-tailed)	**.000**	.761	-
		N	15	15	-
	Screen3Dwell	Correlation Coefficient	-.282	.371	-
		Sig. (2-tailed)	.308	.173	
		N	15	15	-
	Screen4Dwell	Correlation Coefficient	-.117	-.222	-
		Sig. (2-tailed)	-	-	-
		N	-	-	-
Medium Workload	Screen2Dwell	Correlation Coefficient	**.761****	-.439	**-.521***
		Sig. (2-tailed)	**.001**	.101	**.046**
		N	15	15	15
	Screen3Dwell	Correlation Coefficient	-.175	**.589***	.146
		Sig. (2-tailed)	.533	**.021**	.603
		N	15	15	15
	Screen4Dwell	Correlation Coefficient	-.086	.032	.239
		Sig. (2-tailed)	.761	.909	.390
		N	15	15	15
High Workload	Screen2Dwell	Correlation Coefficient	.429	-.129	-.129
		Sig. (2-tailed)	.111	.648	.648
		N	15	15	15
	Screen3Dwell	Correlation Coefficient	-.225	.146	.289
		Sig. (2-tailed)	.420	.603	.296
		N	15	15	15
	Screen4Dwell	Correlation Coefficient	-.268	-.150	.464
		Sig. (2-tailed)	.334	.594	.081
		N	15	15	15

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Screen Percentage, Dwell to Average change correlation

Gamer		Screen2P	Screen3P	Screen4P	Screen2Dwell	Screen3Dwell	Screen4Dwell
Non-Gamer		Screen2P	Screen3P	Screen4P	Screen2Dwell	Screen3Dwell	Screen4Dwell
Av_2to3	Correlation	.507	**.686****	-.486	.236	.182	.079
	Sig. (2-tailed)	.054	**.005**	.066	.398	.516	.781
Av_2to4	Correlation	.132	-.511	-.075	-.104	-.243	-.368
	Sig. (2-tailed)	.639	.052	.791	.713	.383	.177
Av_3to2	Correlation	.629*	**.639***	**-.604***	.243	-.157	-.068
	Sig. (2-tailed)	.012	**.010**	**.017**	.383	.576	.810
Av_3to4	Correlation	-.754**	-.011	**.800****	-.311	.214	.161
	Sig. (2-tailed)	.001	.970	**.000**	.260	.443	.567
Av_4to2	Correlation	-.079	**-.539***	.118	-.343	-.143	-.375
	Sig. (2-tailed)	.781	**.038**	.676	.211	.612	.168
Av_4to3	Correlation	-.614*	.046	**.739****	-.104	-.075	.143
	Sig. (2-tailed)	.015	.869	**.002**	.713	.791	.612