CRANFIELD UNIVERSITY

A J ALLSOP

DETERMINING THE EXTENT TO WHICH SIMULATION CAN BE USED TO TRAIN RAF PILOTS TO FLY AND FIGHT THE EUROFIGHTER TYPHOON

CRANFIELD DEFENCE AND SECURITY

PHD THESIS

Academic Year: 2016-17

Supervisors: K Knowles, A J Saddington

March 2017

CRANFIELD UNIVERSITY

CRANFIELD DEFENCE AND SECURITY

PHD THESIS

Academic Year 2016-2017

A J ALLSOP

Determining the Extent To Which Simulation Can Be Used to Train RAF Pilots to Fly and Fight the Eurofighter Typhoon

Supervisors: K Knowles, A J Saddington

March 2017

ABSTRACT

This research examines the extent to which simulation can be used to train pilots of the Royal Air Force to fly and fight the Eurofighter Typhoon, and is the culmination of a series of trials over a period of 4 years. The approach was threefold, firstly examining the performance of students trained entirely on the Operational Conversion Unit's full syllabus in the simulator and then tested against their peers on each of the four phases in live flight, secondly investigating the cultural acceptance levels of the present Typhoon pilots and lastly using lessons learnt to generate and test a syllabus to train Typhoon pilots to Multi Role Combat Ready in 40% of the present time. It was found that increasing the proportion of synthetics from the lowest Live Synthetic Balance (LSB) of 75:25 used on the front-line meets a cultural and resource barrier at 50:50. This did not represent the maximum LSB achievable however with the heavily synthetic Multi-Role Syllabus reaching an LSB of 21:79 with successful completion of the end of course test.

Cultural acceptance of the simulator had correlations with the squadron a pilot was assigned to, the manner in which the simulators were programmed for use and the experience level of the pilot. No evidence was found within the sample to suggest age had an effect. Recommendations on minimum proportions of live and synthetic training was mapped for each of the required tasks and comparisons of these were made across complexity levels.

Resource savings found by the generating and testing a Multi Role Combat Ready syllabus that recognised and incorporated all the strengths, weaknesses and lessons identified in the previous trials generated a saving of approximately 9 months and 100 Typhoon live flying hours per student, equivalent to approximately 1300 man maintenance hours that could be reinvested into personnel in the form of leave, adventurous training or development.

ACKNOWLEDGEMENTS

Thank you to Nicola, Ethan and Isobel Allsop for their unwavering support through the most difficult of times.

Thanks to Al Seymour who dared to take a risk and in doing so differentiated himself from his peers.

Dedicated to the fighting edge of the Royal Air Force, the men and women who sacrifice more than the public will ever know.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

A2A Air to Air Role.

ACT Air Combat Training.

AFRL Air Force Research Laboratory. A US facility studying many facets of aviation, one of which is simulation.

AFT Advanced Flying Training.

AO Areas of Operation. Possible areas of conflict around the globe.

AOA Aircraft Operating Authority.

ASTA Aircrew Synthetic Training Aid: Very high fidelity simulator built by the 4 nation consortium that is the standard simulator for all nations operating Eurofighter.

AVM Air Vice Marshall: A senior rank in the RAF.

BFM Basic Fighter Manoeuvres.

BFT Basic Flying Training.

CA Counter Air. A task that trains longer range combat using missiles.

CAA Civil Aviation Authority. The body governing civilian aviation in the UK.

CAP Combat Air Patrol. A tactically significant position from which a formation of aircraft will guard significant assets, usually established as an an orbit.

CFT Crew Flying Task. The tasks required to maintain each pilot at a suitable level of competency for particular roles. The overall task sets the number of flying hours per month a pilot will be funded for.

CR Combat Ready.

CT Cockpit Trainer.

C-T Continuation Training. Training undertaken by pilots to keep them current in particular skillsets.

DA Defensive Actions.

DACT Dissimilar Air Combat Training.

DASS Defensive Aids Sub Suite: The suite of sensors that detect electromagnetic emissions, classify threats and dispense protection eg. flares or jamming.

DP Duty Pilot. Experienced pilot positioned in the air traffic tower to assist with safety issues of airborne aircraft.

ECM Electronic Counter Measures. Methods of protecting an aircraft using the electromagnetic spectrum.

EDCT Emulated Deployable Cockpit Trainer: Mid-fidelity simulator built and operated by BAE Systems.

EoP End-of-phase Check Ride.

EWI Electronic Warfare Instructor. The RAF's cadre of specialists in the field and employment of electronic warfare

FAA Federal Aviation Authority. The governing body for US aviation.

FOB Flying Order Book.

FMS Full Mission Simulators.

FQMD Forward Quarter Missile Defence.

HVAD High Value Asset Defence. Training to protect High Value assets such as tankers.

IP Instructor Pilots. Qualified instructors that provide the instruction on the Conversion Unit.

IRT Instrument Rating Test.

JAA Joint Aviation Authority. The governing body for European aviation.

JAR Joint Aviation Requirements. The requirements for European aviation as laid down by the JAA.

JSF Joint Strike Fighter: Next generation fighter produced by the US. The RAF intends to have a combat fleet consisting of only Typhoons and JSF by 2020.

LCR Limited Combat Ready.

LOSB Live Optional and Synthetic *Blend*. Stated minima (%) of the Live and Synthetic training, leaving an amount that could be *opted* to be undertaken in either environment.

LRA Long Range Aviation. Russian long range patrol aircraft such as the Bear or Tu22.

LSB Live / Synthetic Balance: The balance of live and synthetic training stated as a percentage ie 40:60 is 40% live 60% synthetic.

LSJ Life Survival Jacket.

MEC Mission Essential Competencies.

MoD Ministry of Defence. Overarching body controlling the UK military.

MoDREC Ministry of Defence Research Ethics Committee.

MR Multi Role: A combat discipline involving air to ground and air to air skills in the same mission.

NAO National Audit Office.

NME Shorthand for Enemy.

OCU Operational Conversion Unit. The training unit that takes student pilot proficient in the Hawk trainer and instructs them in combat operations using a front-line aircraft.

PB Trial PANDORAS BUZZARD. The first of three trials within this paper.

POM Plane of Motion. The ability to recognise a manoeuvre of a hostile aircraft and position to employ guns against it.

PSC Production Software Configuration. A method of identifying the software standard resident a particular aircraft or simulator.

QPI Qualified Pilots Instructors, see IPs.

QRA Quick Reaction Alert. UK Homeland defence aircraft sitting at readiness 24 hours a day, 365 days a year.

QWI Qualified Weapons Instructor.

RAF Royal Air Force.

RCS Radar Cross Section. The measure of a target's radar reflectivity. Dependant upon factors such as target aspect and host aircraft radar frequency.

SA Situational Awareness. The ability to visualise the 3D battlespace in order to maintain safety or prosecute an attack correctly.

SAM Surface-to-Air Missile systems.

SME Subject Matter Experts. Individuals deemed to be suitably qualified, or of such experience, in a field as to be designated as expert.

SOP Standard Operating Procedures. Explicitly stated methods of conducting processes that are learned by rote, intended to give common and safe ways of working when under pressure.

SQ Sub-Question. A series of questions that collectively answer the research question.

Sqn Squadron.

SRF Sortie Report Forms.

TNA Training Needs Analysis.

USAF United States Air Force.

VID Visual Identification. Methods of achieving a position on an unknown aircraft to provide a visual identification of its type.

VMS Visual Mutual Support. The standard fighting formation for a pair of aircraft – line abreast at 1.5 miles.

WEZ Weapons Engagement Zone. The volume of space from which an aircraft can employ its armament upon the enemy.

NOTATION

α	Qualifying Significance level
H_o	Null Hypothesis
H_A	Alternate Hypothesis
L_c	Learning time of control group in the *live* environment
L_x	Learning time of trial group in the *synthetic* environment
N_{min}	Minimal total cost per subject studied
Sp_c	Support live flying required in the control group syllabus
Sp_x	Support live flying required in the trial group syllabus
S_x	Learning time of trial group in the simulator
σ_x	Standard deviation
$P_{1,2 \text{ and } 3}$	Number of pilots at Qualification Level 1,2 and 3
p	Statistical significant of the value found
$t(n)$	Degrees of freedom, where n is the number of
$\bar{x}$	Mean of a sample

Chapter 1. INTRODUCTION

After providing a background and context to the problem the research question will be introduced. The literature review will then expose the research to date as well as highlighting previous issues inherent in research of this type. The research itself seeks to triangulate a solution from three different viewpoints: performance of the student pilots, culture and resources.

As will be explained the impact of this work has been to increase the synthetic proportion of the Typhoon Conversion Unit’s Syllabus from 40% to 75%, to demonstrate a method of reducing the time to train a pilot to Combat Ready from 11 months to 3, and save a total of 137 flying hours and 1700 maintenance man hours per pilot trained. This represents £274,000 per pilot in terms marginal costs such as fuel and daily consumables, and £12,604,000 when using the Royal Air Force’s (RAF) own full capitation figures. In the process it has achieved the European first of a first solo in a fighter aircraft directly from the simulator without the use of surrogate aircraft or designated twin seat trainer.

New knowledge has been developed in understanding the limits of simulation when using it within a high proportion of a syllabus. Additionally relationships between the culture and the perceived limits of synthetics in daily training has also been explored.

The work has resulted in the following achievements:

a. Numerous presentations: the Royal Aeronautical Society’s Flight Simulation Conference (June 2013), the RAFs Central Flying School’s International Conference (2012) well as the RAF’s Chief of the Air Staff, Air Officer Commanding 1 Group and a number of US, German and French Officers of Air Rank.

b. Award of the RAF’s Central Flying School Trenchard Memorial Trophy for ‘Excellence in the Art of Instruction’ (2012).

c. Joint award from the Company of Educators and No. 22 (Training) Group of the RAF for ‘The Most Significant Contribution to Education and Training’ (2013).

d. RAF’s No. 1 Group Operational Innovation Award (2013).

e. The award of an MBE in the New Year’s Honours (2014).

1.1 BACKGROUND

The Eurofighter Typhoon is the Royal Air Force’s Air Superiority Fighter for the next 30 years and with infinite resources all flying training would be conducted in the real-world environment (known as “live” training). However in the real world the RAF faces ever increasing fiscal restrictions, impacting the funding for fuel costs, spares, manpower and exercise participation to identify a few. The net results are substantial reductions in flying hours for pilots and a corresponding loss of knowledge to a level proportional to the hours experienced, impacting safety and effectiveness.

Synthetic training (through the use of flight simulators) will be asked to restore elements of this knowledge imbalance and training deficit. True capabilities and limits of high-fidelity simulation in the instructional and knowledge-transfer role have yet to be tested and the lessons incorporated into RAF pilot training. This research intends to search for elements of those limits and collate evidence to determine the environment (live or synthetic) and the extent to which synthetics can be used within Typhoon training.

Before defining the research question, and in order to provide context, the problem will first be framed through presentation of the pertinent historical needs, issues and difficulties facing the RAF with respect to this field before laying out the research question itself.

1.2 DEMONSTRATION OF THE NEED

Three months after funding for this research was secured the Permanent Under Secretary for the Ministry of Defence (MoD) and two Air Vice Marshalls (AVMs), the Directors of Combat Air and Information Superiority, were asked to answer the questions of the Public Accounts Committee in the House of Commons (Hansard, 2011). During questions fourteen to twenty three of the session the Committee expressed surprise at the small number of combat pilots trained to conduct Multi Role (MR) operations. AVM Hillier, in his response, stated the reason for such a small number being trained as “flying hours are expensive, so what we do not want to do is apply flying hours to keep people with a particular skill set that we do not expect to deploy on operations.” The Chair questioned the value for money of having 52 Typhoon aircraft with only 8 pilots capable of using it to its full capability.

In the same month as the senior RAF officers and civil servants faced the Committee, the National Audit Office (NAO) released their report as to the Management of the Typhoon Project (Chambers *et al.*, 2011). In their observations of the risks to training the NAO states that “shortfalls in the number of hours available for training reduce the range of flying competencies of pilots and increases flying risks” (p19).

The need and reason for the funding of this research was highlighted during the preceding year, 2010. Owing to a combination of factors flying hours available to train the pilots were becoming an increasing concern. Figure 1-1 from Chambers et al. (2011), shows the proficiency of pilots as a consequence of the flying hours received.

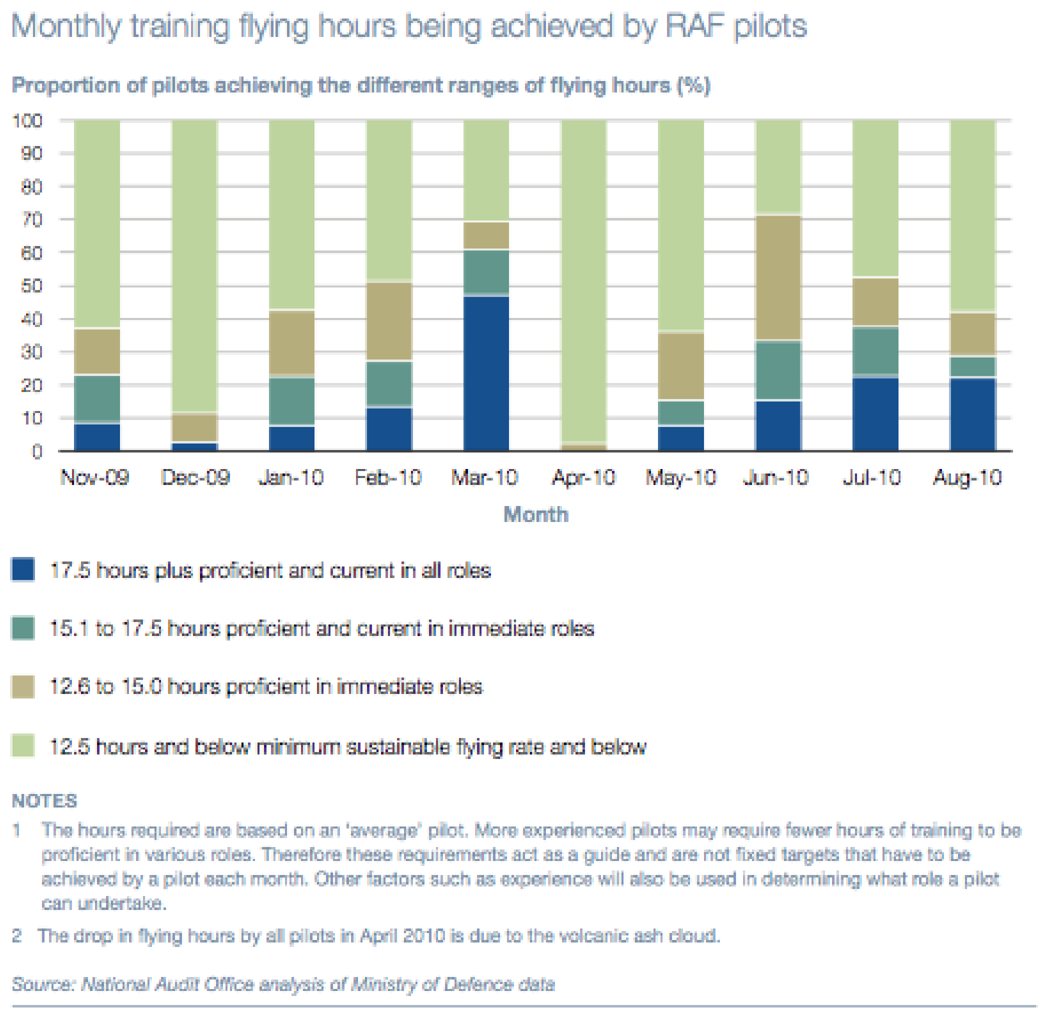

Figure 1-1. Monthly Flying Hours Achieved by RAF Typhoon Pilots

It can be seen that in the majority of months 50% or more of the pilots received training that provided only the minimum sustainable flying rate. The result was that five pilots were temporarily grounded because, as AVM Hillier stated to the Public Accounts Committee, “they were not getting enough flying to maintain their currency and skills.” (Hansard 2011, question 70).

These examples highlight, at the highest levels, the difficulties of providing training in expensive assets such as fast jets within the environment of ever increasing fiscal restrictions and the consequent drive to achieve maximum efficiency and effectiveness. When one considers that in 2010 the cost per single flight hour of a Typhoon was given as £70,000 by Peter Luff the Permanent Under Secretary for Defence (Hansard, 2010), it can be seen why comments are drawn from such high levels[1]. The effect of these difficulties can be seen both visibly, such as the examples above but also at the non-quantifiable individual level. There are only a few Typhoon pilots and each one grounded or not able to undertake a particular role forces the remaining pilots to shoulder the extra duties such as Quick Reaction Alert (QRA), Falkland Islands detachments and combat duties such as Op ELLAMY (Libya). This extra loading is felt not only by the pilots but their families as well.

The deficiencies, questioned at governmental levels, showed that in 2011 the RAF had a very public need: to train pilots more quickly and achieve greater competencies with the limited resources available and to do so within the environment of national cutbacks and global recession. Whilst these observations were made within the context of 2011 the forthcoming desire to export the aircraft and conduct export pilot training as well as the, at that time unseen, Libyan operations would do nothing but raise this need higher in the priorities.

1.3 BIRTH OF A REQUIREMENT

The literature review will demonstrate that, during 2010, the RAF had a weak understanding that simulation might provide a solution to elements of the problem. It had, however not articulated this any more firmly than a verbal aspiration to achieve a 50:50 Live Synthetic Balance (LSB) by between 2015 and 2020. Papers produced (see section 2.2.1) drove desire for increased synthetics with headlines of savings that could be achieved. Arguing against the increased use of synthetics were the pilots themselves who would forcefully attack any idea that training on the scale proposed could be conducted in the simulator. There was no practical understanding of the effect of moving large amounts of training into the simulator and the true capabilities and limits of high fidelity simulation in the instructional and knowledge transfer role had not yet been tested or any lessons incorporated into RAF pilot training.

---

[1] More recent figures place the Full Cost Capitation rate at £92,000 (Air Command HQ 2012).

The fast-jet fleet of the RAF in 2010 consisted of Harriers, Tornado GR4s, a few remaining Tornado F3s and Typhoons. With the exception of the Typhoons all these aircraft were at least twenty years old, with the original design being much older. Training for these various platforms had been built upon tried, tested and proven processes that centred around live flight. The students would be instructed on the ground then shown a technique in the simulator, if it was capable, before being re-taught in a twin seat aircraft and repeating in an operational aircraft. The process was costly in both time and money.

Simulators for these tasks included a mixed fleet of older procedural trainers that, lacking any visuals, allowed training for sorties such as instrument tests and head-down cockpit tasks only. The more developed simulators did have rudimentary visuals that allowed some out-of-the-window tasks to be performed but they were not networked, forcing the pilot to train in an isolated environment rather than the 4-ship tactical formation that was the norm on combat units. Recent procurement of Hawk aircraft, however, recognised the need for better accompanying simulation. The Hawk T2 used at RAF Valley employed the latest technology to provide synthetic training solutions (see Figure 1-2 and Figure 1-3).

Figure 1-2. Hawk Cockpit Trainer.

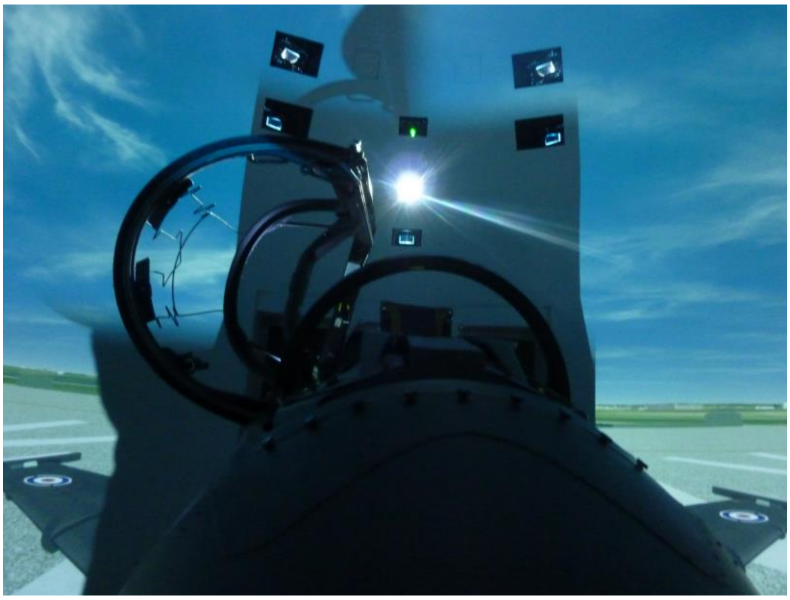

Figure 1-3. Hawk Full Mission Dome

The increasing use of visuals in the updated simulation and the provision of a networked capability allowed aircraft like the Hawk to train correct positions in combat formation and allow the pilot to practice techniques with reference to the ground, such as air-to-ground gunnery. Despite these improvements the training syllabus still repeated simulation sorties airborne rather than replacing a live flight *in toto*, i.e. these state-of-the-art devices were seen as a reinforcement to airborne learning rather than replacement. This was supported in an interview in September 2011 with Mr Lloyd-Jones of the instructional team at RAF Valley who stated that “live flight was perceived and trusted as the only method of ensuring a pilot’s knowledge was satisfactory”.

The Typhoon simulator, known as the Advanced Synthetic Training Aid (ASTA), promised a different level of fidelity both in terms of the visuals and the models of the enemy combatants. These models would be run attempting to create a synthetic world controlled using comparable real world parameters such as decibels-based radar energy, aspect dependant Radar Cross Sections and real aircraft software. The desire was aspirational, the results in 2010 were disappointing; years behind schedule and suffering from a lack of robust software. Recognition of these issues had resulted in the procurement of 2 Emulated Deployable Cockpit Trainers (EDCTs) which used emulated software as opposed to the real aircraft software. The result was a comparatively cheap solution that provided a level of fidelity that provided students with an understanding of some of the displays and controls and allowed training in the majority of tasks to a basic level. It was however unable to reproduce emergency situations in high fidelity or, owing to its poor GUI, allow control of anymore than two enemy entities. Nevertheless it represented value for money even provided a best guess of forthcoming hardware and software such as the Lightning III weapon targeting pod and Air to Surface weapon computer coding. Use of this equipment was limited to the older style of training; that of replicating a sortie prior to repeating all the aspects again airborne. This training was limited to certain sorties due to the limiting visuals of 240 degrees in azimuth meaning that the leader lost visual with his wingman after the first manoeuvre. Additionally the over-optimistic performance of the radar software, ensured an overly tenacious lock on the target aircraft.

The advantages of emulation over simulation were cost and the speed of software corrections, however it suffered from a major drawback - realism. Emulation copied the desired performance, as laid down in the contract of each software load rather than what was actually resident in the aircraft. As a result the early versions of the Emergency Flight Reference Cards, used by the pilots to correctly diagnose and remedy aircraft emergencies, were based on these emulated solutions and subsequently contained numerous errors that were only discovered as the maturity of the simulated software progressed.

Thus the RAF had seen the delivery of vastly improved simulation systems that appeared to offer a leap in capability, their papers (Harper and Hillier, 2007; JTES, 2009; Wells *et al.*, 2009; Air Staff, 2010; MoD, 2010b) suggested the financial savings that could be realised. In contrast however there was no clear direction or strategy to deliver a trusted-results based method of altering the years of culturally-entrenched training techniques, particularly in light of the varying simulator's degrees of replication and ability.

1.4 POST STRATEGIC DEFENCE AND SECURITY REVIEW (OCT 10)

In October 2010 the new coalition government returned the first Strategic Defence and Security Review of the millennium. The short-term result for the RAF was the almost immediate scrapping of the Harrier GR9 and Nimrod fleets as well as an announcement of forthcoming redundancies of 7000 personnel, approximately 14% of the serving manpower. This was accompanied by the open statements that the RAF would move to a fast-jet fleet consisting of only 2 types and that the Joint Strike Fighter (JSF) would be postponed from 2013 to 2020. This was caused by the £36bn deficit in the defence budget accumulated in the preceding years. Against this backdrop demand came for increased use of synthetics in order to save money, which in turn would set the 50:50 Live Synthetic Balance (LSB) strategic aspiration.

1.5 THE FAST-JET SYNTHETIC PICTURE 2011

By 2011 the EDCTs had been repositioned at RAF Leuchars to support the build up of the 3rd and 4th front-line Typhoon squadrons. Their use was limited to the monthly currencies for emergencies and occasional demonstrations of tactical concepts. Use of these devices in an instructional context was limited. ASTA's software by contrast was becoming more reliable than the past but suffered significant lag in terms of concurrency. As such ASTA was unusable for sorties or instruction using the Defensive Aids Sub Suite (DASS) amongst other items. Progressive attitudes on the use of the ASTA to instruct Student Pilots (SPs) had seen the content of the Operational Conversion Unit (OCU) syllabus grow to 54 % simulator 46% live flight, however this was for only a single phase of the 4 phases of the training, the rest still lagging markedly behind. These figures were in excess of any seen before in the RAF but still adhered to the old style of practicing for a sortie that would be repeated in live flight. The front-line training during this period was conducted almost entirely in live flight with only the mandatory monthly emergency simulator impacting their LSB.

1.6 RISK

The desire, in 2011, for a 1:1 LSB by 2015 was articulated in fleet plan 31S. There was, however, a lack of evidence that supported this aim and the balance had not been tested, neither had the synthetic devices shown a capability that would provide it: of the research and papers that had been produced to support the plan (Harper and Hillier, 2007; JTES, 2009; Wells *et al.*, 2009; Air Staff, 2010; MoD, 2010b) all highlighted the potential financial rewards but rarely mentioned the effect on training or operational capability of the RAF. Only one paper mentions practical methods of achieving output, such as the need for a Training Needs Analysis (Wells *et al.*, 2009, p. 8).

The risk facing the RAF, therefore, was that the desired sacrifice of the live flying hours in order to pay for the cheaper simulators may yield both a synthetic environment that was unable to achieve the required pilot standard and a challenging flying budget that was unable to provide the flying hours to retrain the sub-standard pilots. The effect would be to perpetuate the problems demonstrated in the OP ELLAMY era: sacrificing the skills, training and currency of certain pilots such that others may train to a proficient level.

1.7 RESEARCH QUESTION

The overarching research question was intended to seek the extent to which simulation could be employed within Typhoon training, in order that the RAF might become more efficient with its resources in both fiscal and manpower terms. Research, rather than a simplistic immediate employment of simulation, was required for 3 reasons. Firstly, at the strategic level this new training medium and emphasis was to be able to assure that defence of the realm could be maintained within the constraints of present and future SDSRs. Secondly, at the tactical level training within simulation should be able to output sufficient numbers of pilots that could reach the required standards. Thirdly, any limitations or areas where simulation was weak were to be understood in order that they could be avoided should heavy simulation syllabi be employed in training.   Thus the guiding research question was:

" What is the maximum extent to which pilots of the Eurofighter Typhoon can incorporate synthetic training? "

Unlike some academic research this work was to be undertaken within a practical environment. In order to be considered a success, and due to the significant costs of the assets and risks involved, the outcome of the research question was visibly to demonstrate employment and associated success and failure *within the practical context*, and most importantly have addressed the practical aspects of the problem. Analysis of the research question in context returned a number of barriers to a solution, Figure 1-4 provides a pictorial representation of the layered issues opposing the incorporation of increased simulation usage. At the heart of the problem lay pilot performance, should this work have been unable to demonstrate a standard at least equal to present methodologies then the use of a heavy synthetic syllabus would be immediately discounted as ‘sub-standard’. The middle band focuses on the more holistic issues that surround the research question, each having the ability to significantly weight any solution to the question.

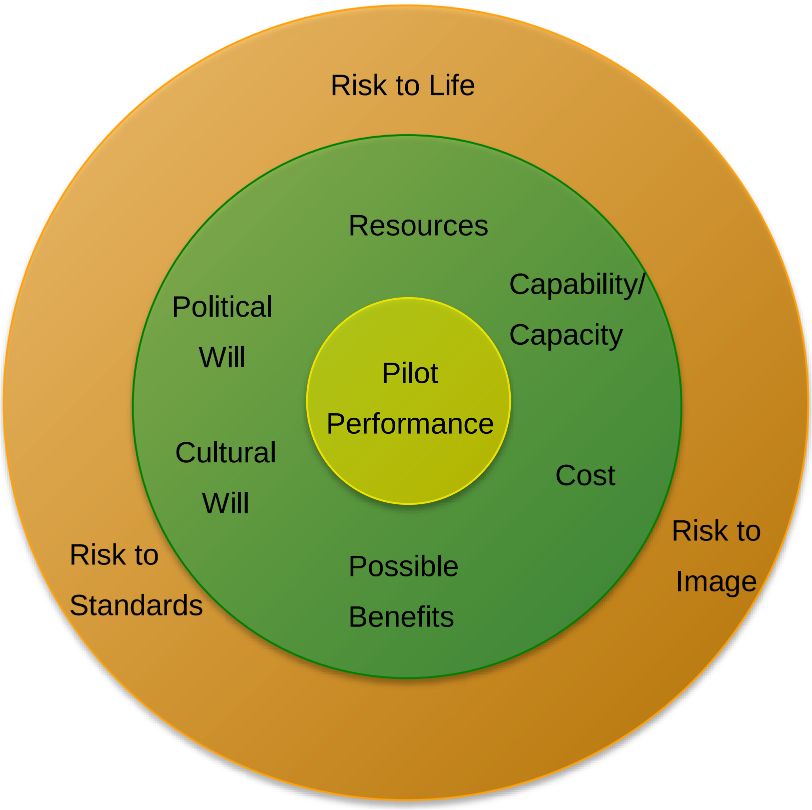

Figure 1-4. Pictorial Representation of Issues

Finally, and most importantly, the outer band considers the issue of risk. In order that the research might be incorporated into the Eurofighter's core methodology it would have to be sanctioned by at least the rank of Group Captain, although initial work was expected to require agreement from significantly higher. All officers agreeing to the research would need to satisfy themselves (and the Military Aviation Authority) that the risks had been adequately addressed. In conjunction with these risks the research's small sample sizes (discussed in section 2.5.3) would find difficulty in proving definitively the success or otherwise of elements of the methodology, when considering the entire population. Thus answers to research question required triangulation through a number of differing but complementary approaches, and in doing so provide a series of independent answers that when viewed holistically *reduced the level of risk facing senior officers when considering the incorporation of a heavily synthetic syllabus.*

Viewing the question in this manner allowed the identification of three themes that were consequently used to guide the literature review and provide a “handrail” when dissecting the research question in Section 1.8. The three themes are:

1 Performance. Incorporating the issues of pilot performance, risk to life, risk to standards, possible benefits and capability.

2 Culture. Taken from the issues of cultural will, political will and risk to image.

3 Resource Use. From the issues of capacity, resources, cost and possible benefits.

1.8 DISSECTING THE RESEARCH QUESTION

Before undertaking the literature and knowledge review the research question and the associated layered issues from Section 1.7 were broken down into a series of sub-questions and objectives, shown below. Each of the objectives are considered during the review so as to source answers and guidance to inform the methodology. In this manner the objectives were to focus the overall methodology towards a series of smaller answers that, when viewed together, would answer the research question fully whilst being informed by previous research.

Thus from the guiding research question,

“ What is the maximum extent to which pilots of the Eurofighter Typhoon can incorporate synthetic training? ”

the following sub questions (SQs), addressing the layered issues, were identified:

SQ1 “ What limits synthetic training in the instructional environment? ”

SQ2 “ What is the cultural limit of synthetic use? ”

SQ3 “ Do resources limit the employment of synthetics? ”

1.8.1 SQ1 - WHAT LIMITS SYNTHETIC TRAINING IN THE INSTRUCTIONAL ENVIRONMENT?

This SQ addresses the first theme and primarily the inner layer of issues (Figure 1-4) that is concerned with pilot performance when simulation is used for their training. Specifically will a heavily synthetic syllabus reduce the overall performance of the pilot? The hypothesis for this phase being:

“*OCU training can be conducted using simulation alone*.”

The guiding objectives are:

a. **Objective 1**. Assess the suitability of ASTA to train each phase of the OCU syllabus; first solo, 1v1 Combat, Counter Air and Low Slow QRA training when compared to the present methodology.

b. **Objective 2**. Identify any variables that could affect the transfer of training.

c. **Objective 3**. Identify any technical limits prohibiting the employment of a highly synthetic syllabus to train students.

d. **Objective 4**. Determine the level of training transfer between ASTA and the live environment.

e. **Objective 5**. Generation of a risk register to determine the risks during the first solo flight, with particular reference to risk to life.

1.8.2 SQ2 - WHAT IS THE CULTURAL LIMIT OF SYNTHETIC USE?

Sub-Question 2 examines the second theme of the research question, that of culture, in particular it looks at aspects of, and barriers to, the employment of simulation within day to day training.

a. **Objective 1**. Determine the present level of cultural acceptance of synthetic training use in front-line Typhoon training and investigate if this is common across both Typhoon bases.

b. **Objective 2.** Identify any factors, such as experience levels, that correlate to the level of acceptance.

c. **Objective 3.** Prove or disprove the two commonly held beliefs that were held by senior ranks and witnessed during briefings in 2010 (those senior ranks being Air Vice Marshal S. Atha, Air Commodore G. Waterfall and Wg Cdr A. Seymour):

i. Younger pilots are more accepting of simulation than older pilots.

ii. Acceptance of simulation is based upon a pilot's qualifications.

d. **Objective 4**. Determine the minimum levels of LSB for each of the required tasks, as subjectively assessed by the present Typhoon pilots.

e. **Objective 5**. Investigate the effect of threat complexity on the LSB.

1.8.3 SQ3 - DO RESOURCES LIMIT THE EMPLOYMENT OF SYNTHETICS?

The third theme of the primary research question was focused on the ability of the pool of resources to conduct simulation. This would require the construction and proof of concept for the final training area, that of the Combat Ready work up (the section of training conducted immediately after joining the Front-line from the OCU, usually in the order of 12 months).

a. **Objective 1.** Construct and demonstrate proof of concept of a CR work-up syllabus that minimises the resource input of the front-line squadrons.

1. **Objective 2.** Determine the levels of resources required to conduct this syllabus and any resource savings made by its methodology.

2. **Objective 3.** Determine if the resources available are sufficient to employ this syllabus across all CR work up pilots.

In summary the research sought to determine the limits of exploitation of simulation within the context of the Eurofighter Typhoon which would *reduce the level of risk* facing the RAF, should heavy synthetic syllabi be endorsed. The inability to answer the question categorically through synthetic pilot training alone was recognised, given the effect of cost and time on likely sample sizes. Thus a triangulation path to a solution was plotted to provide both evidence and an in-depth understanding of the capabilities of fast-jet simulation based upon assets presently available, in order to inform future fast-jet training strategy. It was envisaged that the research would be used to source cost savings, maintaining present capability, or direct the reinvestment of resources to increase operational capability without an overall increase in cost.

1.9 OVERVIEW OF THE THESIS

In order to address each of the sub-questions that were determined after the dissecting of the research question (section 1.8), 3 trials were established to target each of the three key themes and were to run in series, taking a total of 3.25 years to complete. A parallel employment of the trials was considered and discounted as the workload for the primary researcher (a serving officer) would have been too high to have been conducted ‘part-time’. The creation of ‘trial’ and ‘project’ status allowed the author to bid for resource outside the norm and enabled the agreement of the RAF command chain to the research as a whole. The trial / project names and the corresponding sub-questions are provided below:

a. **Trial PANDORAS BUZZARD** (Chapter 3. SQ1 – What limits synthetic training in the instructional environment? An examination of the effect on output performance of training pilots entirely in the simulator with a single test live flight in each phase. The aim being to determine if simulation alone was a valid way of generating experience and highlight any limitations.

b. **Trial PANDORAS DIAMOND** (Chapter 4. SQ2 – What is the cultural limit of synthetic use? A questionnaire based investigation of an informed audience, the Combat Ready Typhoon pilots, on the limits of day-to-day training in the simulator, with a view to proving the Commander with cultural limits of the employment of synthetics.

c. **Project JENX** (Chapter 5. SQ3 – Do resources limit the employment of synthetics? A test of concept: training a pilot through a heavily synthetic syllabus to Combat Ready on the front line, based on the learning of the previous trials. The intent to being to determine the nature and scale of savings possible.

Post a review of the available literature and knowledge, the chapters consider each of the trial / projects in turn and, after providing an overview, state the design with reasoning, methodology, results and analysis. A final, separate section of analysis viewing the research as a whole is then provided to return the focus to the overarching research question.

Chapter 2. LITERATURE AND KNOWLEDGE REVIEW

2.1 INTRODUCTION

The review is intended to provide the context in which the subsequent trials were to take place. This is done through an exploration of literature and commences with an exposé of the RAF's knowledge in the area at the time both at a strategic and lower tactical level, through papers and policy documents. Once the context is established other Air Forces and the Civilian stance are compared and contrasted to present the RAF's position from a wider perspective. Having established a more global picture the review returns to detail: examining experimentation methods within the field that have informed these positions. Finally the theories devolved and limitations encountered from employing the experimental outcomes in practice are identified and made explicit.

2.2 THE RAF

2.2.1 UNDERSTANDING THE PICTURE – STRATEGIC LEVEL

The need for more increased use of simulation had been recognised within the RAF but the first attempt at a strategy to progress towards viable high level training was produced in 2007 (Harper and Hillier, 2007) with the stated aim that:

> ***"All force elements [are] able to train in a realistic, complex and hostile joint scenario with real or representative equipment"***

The importance of this document is illustrated by considering that this was the first time at this level that the requirement to train collectively was articulated for the synthetic environment. Collective training is defined by the United Kingdom Glossary of Joint and Multinational Terms and Definitions as 'individuals, units and command formations are collectively prepared for operations' (MoD, 2006). Prior to this point simulation for the RAF's aircraft was entirely stand-alone. Tornado F3, GR4 and Harrier GR9 had devices capable of training single pilots in elements of mission requirements but lacked the ability to train as full combat formations, thus any training above the individual level had to be done in the air. Harper & Hiller recognised that as the platforms' capability, security and agility increased this live training would become unaffordable.

The motivation of the paper was therefore financial and, whilst it recognised that a well designed collective synthetic environment could reduce the gap between training output and competency (Harper and Hillier, 2007, p. 3) and thus increase defence capability, it endeavoured to fund this new ability through the savings made by reducing live flight by 15% (p5). The paper went on to state that, with respect to the optimal LSB, 'there was no agreed coherent position' and that to return a solution in 'optimal time' Subject Matter Experts (SMEs) should make 'subjective judgements' (p5). It is argued therefore that this future, unknown capability and its abilities, proposed by the paper were to be financed by the present day training system; with an unknown effect on the near-term output and competency gap and subsequent RAF contribution to UK defence capability.

The unknown nature of this desired capability is further highlighted by a paper on simulator strategy produced two years later that stated that the underlying technology for networked simulation was 'still relatively immature and not all users were convinced of its value' (JTES, 2009). Again the idea of altering the live / synthetic balance to fund this simulation was stated without an explanation of the present or near future effect.

Two years after the initial strategy paper was written Air Command sought to determine the optimal blend for the Tornado and Typhoon aircraft (Wells *et al.*, 2009). The authors highlight the concern over Crew Flying Task (CFT), referred to as the minimum sustainable flying hours in the NAO report (Chambers *et al.*, 2011), stating that this is based on 'limited availability of statistical information and significant military judgement' (Wells *et al.*, 2009, p. iii) and suggested that a review of CFT be set up as a matter of flight safety priority. Their roadmap towards the optimal LSB blend, however, reduces CFT below this minimum in 2017 assuming that synthetics will suffice from a flight safety viewpoint. The original strategy paper, however, stated that the increase in simulation capability would be bought at the expense of live flight; should this reduction below the minimum CFT be recognised as a flight safety hazard, the funds to buy back live flight would already have been sunk into simulation technology. This would leave the RAF unable to return to sufficient numbers of hours to provide safe live flight and without a synthetic environment to replace them. Aside from the flight safety issue that this would present, the affect of insufficient hours on defence capability is adroitly illustrated in a non-theoretical context by both AVM Hillier at the Parliamentary Accounts Committee and the NAO report discussed in the Section 1.2.

Unlike any of the preceding reports Wells' Optimal Blend paper does recognise the need for a Training Needs Analysis (TNA) to be carried out (Wells et al. 2009, p.8) in order to determine the scope of training activities that are required to achieve the operational output. At the heart of the decision-making process are the subjective decisions of the Subject Matter Experts (SMEs) over what level of simulation fidelity is required to conduct training of the student (Wells et al. 2009, p.6). These decisions however, would be unlikely to be tested against a student until the simulation prototype stage, by which time considerable money would have had to have been found from live flight to fund the development of the new technology.

During the run up to the SDSR the simulation and training workstrand produced a Study Report intended to inform this important debate (MoD, 2010b). In it training in simulation was quoted as costing '5 to 20% less' than Live and suggested that £300-£600m pa could be saved if 'on average, 25% of current Live Training was transferred to Simulation'. Upon examining where these figures were obtained the paper utilises a schematic from an Australian Department of Defence report - Reducing the Cost of Ownership through Simulation (2008). This report had examined the worldwide LSB of coalition partners and plotted percentage of training conducted in simulation against cost saving. The weakness of the report was that it lumped together land, sea and air synthetic savings despite the fact that the environments and thus requirements to be modelled differ greatly between the disciplines. As an example consider solely the air environment and the generic platforms of a transport aircraft and a fighter. The transport aircraft simulation might only require detailed models of the aircraft performance, handling characteristics and visual databases of the airports it would be likely to land at. The fighter on the other hand would require all these plus radar, Forward Looking Infra Red (FLIR), Night Vision Goggles (NVG's), Defensive Aids Sub Suite (DASS), Reconnaissance pods and weapons modelling. To operate in the environment it would fight in, the simulation of enemies would require correct Radar Cross Sections, enemy weapon modelling, combat doctrines and radar models. So it can be seen that providing a cost-saving figure across all realms of simulation would be relevant only to a highly strategic level such as Chief of Defence rather than the individual heads of the RAF, Navy or Army.

To question the inclusion and relevance of the Australian DoD report within the SDSR simulation report further, the cost of simulation versus live flight is given as 5-20%; exactly the same as the bracket given by another, very much earlier, report investigating cost effectiveness of flight simulators in the military (Orlansky and String, 1977). This report was written in 1977 and used limited fidelity devices, the majority without visual systems and used for procedures and thus of limited use compared with today. This would infer, therefore, that despite almost 40 years of simulation and computer development, the percentage cost savings remain constant.

In order to make the large cost savings desired the amount of training undertaken by simulation would have to be increased. The UK's percentage of defence training conducted in simulation is shown to be in the order of 23% whilst the majority of the remaining allied partners are clumped between 40-50%. The extra synthetic training requires better simulators and these would appear to increase the comparative cost of simulation versus live training by 10% according to the SDSR study (MoD, 2010b). This still represents considerable savings considering the larger amount of training that can be undertaken. In order to increase savings further the impact needs to be felt within the numbers of aircraft procured or maintained within the fleet and the manpower required to service them (Harper and Hillier, 2007; MoD, 2010b). To provide the strategic level decision makers of the RAF with the confidence to heed the need to 'embrace this step change' (MoD 2010) there is need to provide greater evidence that the 1:1 balance advised will consistently provide the pilot with the abilities that UK defence requires if the live flying hours are to be sacrificed for this aim.

This is supported by the official response of an Air Staff meeting in which concerns are expressed that should the targets of 25% simulation by 2015 and 50% by 2020 not be reached 'we may create long term (and expensive) capability recuperation issues' (Air Staff, 2010). Indeed the Air Staff explicitly state that there is still little evidence to support the significant savings claimed and, just as importantly, there is no guidance as to how to maintain training output during the transition.

A final point to draw out at the strategic level is that in none of the literature reviewed was the 1:1 LSB intended for the RAF articulated with greater fidelity than a generic figure for an entire aircraft type. This assumes therefore that simulation is of equal relevance, and indeed equally capable, to train all squadrons and units within that aircraft type. The Typhoon Force has (as of Sep 2011) three front-line squadrons, one Operational Conversion Unit (OCU) and one Trials squadron. The OCU trains students to Limited Combat Ready – QRA standard, the front-line squadron develops the student further to a Combat Ready – Multi Role pilot. The Trials unit, however, works at the very edge of the aircraft's understanding, developing and assessing new technology and recommending methods of exploiting the new technology tactically. Within this context it can be seen that each of the elements have overlapping yet differing demands on the capabilities of a simulator. To complicate the picture further there are a total of 6 different Typhoon aircraft types used across the squadrons, known as Blocks, and this does not include the twin and single seat distinctions. The 1:1 LSB aspiration is unlikely to be equally successful across each element of this vista. If this is the case then should some of the squadrons do very little simulation yet others replace live flying altogether? So if the inference within the LSB desire is live flight replacement rather than augmentation; safety implications of live flight replacement need to be considered particularly carefully. The penalty for getting this wrong is at best a serious incident, at worst – death of a serviceman and loss of an aircraft.

2.2.2 UNDERSTANDING THE PICTURE – TACTICAL LEVEL

In September 2011, at the commencement of this research, the LSB on the Typhoon OCU - 29 Sqn, was almost exactly 3:2 with 69.5 hours of live Typhoon against 45 spent in the simulator (29 Squadron, 2011b). This reflects the RAF fast jet training pipeline as a whole (Blyth 2015 p113) which utilises almost exactly the same proportions. This contrasts with the Typhoon Force as a whole, including the front-line, which had an average LSB of 6:1 (Wells *et al.*, 2009) in FY 08/09. This comparatively large value of simulation in the 29 Sqn OCU syllabus was largely due to an extensive syllabus rewrite 18 months earlier, which looked to lever the synthetic advantage as far as was considered possible. This LSB therefore represents the best contribution possible from the synthetic environment as assessed by the SMEs at the time. The savings this provided in terms of aircraft flying and man maintenance hours cannot be overlooked. To illustrate this point consider the following: firstly, of the 69.5 live flying hours in the present syllabus only 37.75 are flown by the student pilot the rest was support flying ie. as wingman or threat aircraft (29 Squadron, 2011b), and secondly according to a question to the Officer Commanding Typhoon Engineering during March 2011 each live flight hour required 22 man hours of maintenance[2]. It can be seen, therefore, that for each hour of live flight placed in simulation almost another full hour of support flying (eg Typhoon being used as a target aircraft or an instructor as wingman in a Typhoon) is saved and thus

[2] In 2014 the 22 hours had reduced to 13 engineering man hours per live flying hour.

the man hours of maintenance saved is in the region of just under 44 hours for each hour a student flew.

In increasing the synthetic proportion struggles much past this figure as a number of significant issues are encountered. ASTA is the label given to the Typhoon simulation devices as a whole and whilst they will be discussed in more detail in the experiment discussion later it is worth identifying, at an overview level, the shortfalls of the equipment and the impact of the tactical level view of simulation. The first significant issue is simulator functionality. The lack of any debriefing or usable playback facility prevents reinforcement of any lessons learnt and has thus, to date, relegated learning to the practice of motor skills or the witnessing of setup parameters of the forthcoming sortie.

The second issue is currency. The simulator has traditionally suffered from out of date radar software, displays and controls that do not match the aircraft and a Defensive Aids Sub Suite (DASS) that did not work. Whilst these problems were experienced on previous simulators the lack of a successful solution has underlined the unsatisfactory nature of these new generation simulators.

The third issue is reliability. The ASTA initially had poor serviceability rates, which contributed to the reluctance of pilots to drive to the simulator building for anything other than their mandated monthly emergency sortie.

The final issue is that of culture. Fast-jet pilots measure their experience by live flying hours (RAF Form 414, the RAF flying logbook), the coveted 1000hrs-on-type badge is measured on these live hours. Additionally the document called the ‘White Ticket’ (29 Squadron, 2011b) represents the hours provided by command for the achievement of the set task. Squadron Commanders are continuously required to demonstrate how many live hours they have flown against this measure. There is no incentive, even from a higher level, for a pilot to enter the simulator for anything other than compulsory training.

These barriers have been the reason for lack of adoption of any of the recommendations from the first ASTA training thesis (Lockwood, 2006). This thesis was the first serious effort from within the Typhoon force to improve simulation and its contribution to the fighting effort of the fleet. Whilst the paper did examine elements of reliability and identified the lack of a debriefing facility it did not contain any practical evidence of the training potential of the simulator. Instead it used SME judgment to determine the simulator's ability to train against a set of Mission Essential Competencies (MECs). Lockwood's recommendation that each pilot conducts 24 tactical sorties per year within the simulator (2006, p.36) is based upon a recognition of reduction of opposition aircraft to train against (known as red air) rather than a qualitative view of synthetic capabilities.

Section 2.2 has sought to demonstrate that the RAF's desire to investigate simulation stopped short of anything other than an 'on paper' analysis, which used only generic evidence from other militaries with no knowledge of whether this was applicable in the specific environments intended. Additionally only select individuals recognised the possible long-term impact of a decision to follow a heavy simulation path without any further evaluation. Complementing the 'theory only' strategic view were the four issues at the tactical level that provided an effective deterrent to any increased employment of simulation. The result was an effective barrier to any realistic investigation in the ability of ASTA to train pilots.

2.3 OTHER AIR FORCES – PRESENT-DAY COMPARABLE LSB

This section will examine the closest rivals to the Eurofighter Typhoon in terms of role and performance and compare their LSBs with a view to measuring the exploitation of simulation by the world's leading nations. This will allow Tyohoon's progress to be viewed within a relevant context and give some indication if the strategic aim of 1:1 balance is realistic.

Within the Typhoon partner nations the UK leads the proportion of training conducted within simulation, an outcome largely borne from procurement rather than design. The RAF presently has a single base with two full mission simulators and two cockpit trainers all of which link together via a local network to allow full mission rehearsal as a four-ship formation and provide a quantity of devices that ensures there is the volume to train large numbers of pilot synthetically. In contrast Italy and Germany have procured only two devices at each of their main bases, the intent being to link between bases via a secure network which, to date, has not been developed. The result is that both front line and their OCUs compete over fewer resources per base, restricting the employment of synthetics.

At the commencement of this research the Saudi Air Force were yet to receive their simulators, discussions as to the suitability of the four nation product had led to a decision to procure a BAE Systems proprietary solution which was delivered in 2012. All the Saudi pilots trained by the author during 2009-2012 exhibited a reticence to employ simulation outside the normal methods of practicing for emergencies or pre flight preparation similar to those that had been used by the RAF for many years. The attitude was best summed up by Colonel Mohammed Al-Shahrani (later Brigadier-General and Base Commander of Taif, the first Saudi Typhoon base), who stated during a conversation with the author on 22 Jun 2009 “we have different cultures, we wish to fly not simulate but we will watch the RAF.” The result of this reluctance has led the Saudi Typhoon training system to have a LSB ratio in in the vicinity of 9:1 (stated during a discussion with BAe’s Chief Typhoon pilot, Mr A King on 21 December 2016).

Away from the partner nations the closest fighter in terms of performance, technology and capability is the French Air Force’s Rafale. Use of this aircraft as a comparison is particularly relevant as the Typhoon and Rafale shared their initial design phase before the French separated from the other partner nations to produce the Rafale independently. Nevertheless, the intent and much of the aircraft design is common. Their OCU produces pilots to a similar standard to the RAF - LCR (Limited Combat Ready) and does so with an LSB of 63 live hours and 37 synthetic (Exchange Officer 2011), proportions virtually identical to that of 29 Sqn.

In the US Air Force a RAND investigation by Ausink quantifies the LSB for the US fighter – the F-15, a fighter of the same generation as the Typhoon, at 10:1. Questionnaire respondents, however, had stated a desire to increase simulation to an LSB of 2:1 (Ausink *et al.*, 2011, p. 17). Oddly the F-22, presently the world's only 5th generation fighter in service, which was built without a two-seat training variant and had the need to protect much of its capabilities from prying eyes in the live environment might have been expected to employ simulation for much of its training. Ausink's work however (Ausink *et al.*, 2011, pp. 13, 23–24) found that pilots were recommended to achieve 10 live hours to each simulator hour - the same as the F-15, a generation older. The figures actually achieved per month are much lower at 6 live and 2 synthetic although the subjective opinion of the pilots was that this should be much higher than the recommended figures – the average desired being 13 live to 9 synthetic, an approximate ratio of 3:2 (Ausink *et al.*, 2011, p. 13).

It can be seen from the statistics above that the 2011 Typhoon OCU overall LSB ratio of 3:2 can be seen as broadly comparable to its competition. Thus there was no external stimulus on the RAF to increase the synthetic proportion of training, neither was there an indicator from the peer level that simulation held the answer to the low hours, experience and training displayed in the NAO report, Figure 1-1.

2.4 THE CIVILIAN VIEW

The use of simulators has been widespread in the civilian transport context for over a decade and their use has been incorporated into training by both the Civil Aviation Authority (CAA) (CAA, 2012) and the Federal Aviation Authority (FAA) (FAA, 2012). Their use and inclusion in training is tightly governed and the Authority's approval for that use is controlled under strict guidelines. The approval governs elements such as visual design and acuity, cockpit equipment, motion and sound replication (JAA, 2012). Whilst efforts have been made to bring military simulation representation in line with their civilian counterparts, the efforts have been complicated by the fact that there is no unified view on what the critical items are that should be tested, differing as they do between each of the combat aircraft based on each's purpose. The relevant UK military Joint Service Publication (MoD 2010, Regulation 375.105.2) orders that each Aircraft Operating Authority are to approve the use of each simulator on an annual basis but how this is done is left to the discretion of each authority. Work by the Frazer-Nash Consultancy Ltd (Frazer-Nash, 2010, p. 185) guides the RAF on methods of doing this. The relevant area of this document, detailing how each phase is tested, is 11 pages long. The section applicable to the area of combat aviation – Tactical Operations Phase, is only 2 lines:

> "Demonstrate the tactical or operational performance of the whole simulator as a weapons system. SMEs are responsible for the content and conduct of this phase of this assessment."

This demonstrates that there is little comprehensive guidance on how combat aircraft simulation should be determined as 'fit-for-purpose'. Areas that might have a civilian read across are provisioned for but guidance for any true military application is lacking.

It can be seen therefore that Joint Aviation Requirements (JAR), CAA, FAA and JAA regulations are specific to the civilian context, a context that is only partially relevant to the military application. This is argued similarly in the section 2.2.1 Understanding the Picture: if it was to be deemed necessary to model every requirement and system of a combat aircraft to the fidelity level required by the civilian sector the cost penalties would begin to outweigh the simulators' financial advantages. Not all of these expensive representations can be authoritatively deemed necessary for training. As an example consider the JAR requirement for simulator motion: a meta analysis for the US Navy (Jacobs *et al.*, 1990) found that the "use of motion cueing added little to the training environment for jets, and may even have detracted from training for some tasks." Whilst this analysis was conducted in 1990 it remains relevant as this was the time period during which the ASTA was being conceptually designed; ASTA contains no motion capability. Had money been diverted into the design and build of motion replication in order to mirror the JAR regulations the meta analysis of the time would have indicated that this would have provided little training and financial value.

This section has argued that whilst civilian requirements for simulators provide a useful comparator on how combat simulators should be designed and tested they cannot be directly transferred into the military application. The lack of direct guidance on measuring a combat simulator's fitness for purpose highlights the lack of specialism and knowledge in this area.

2.5 EXPERIMENTATION

If combat simulation's relevance and level of contribution to pilot training is to be understood then, in the demonstrated absence of any regulations governing its use or design, an examination of experimental research should provide a clearer picture of its capabilities. This section will initially determine the extent to which simulation has been investigated and included within both civilian and military flying training, with particular focus on hardware, the issues affecting transfer of training and the historical precedence of sample sizes. Subsequently a study of the ability of simulators to impart the relevant skillsets will be sought, in order to determine better the risk of non-transference of these skillsets in the trial phase.

2.5.1 CIVILIAN

The civilian market does offer an insight into the capability of simple simulators or PC-based simulators to train pilots (Talleur et al. 2003, D'Alessandro 2007). Whilst there are trials that show weak transfer of training in the civilian context (Dennis and Harris, 1998) the majority demonstrate strong correlations for the transfer (Atkins *et al.*, 2002; Talleur *et al.*, 2003;

D’Alessandro, 2008). Whilst this research explores, largely, the more procedural nature of flying such as Instrument Flying (IF) or highly specific tasks such as landing from a radar approach, as a body of evidence it does examine all experience levels of pilots from the beginner through to pilots with over 2000 flying hours. Additionally Talleur’s work testing 106 pilots in the skills involved with Instrument Flying found Personal Computer Aviation Training Devices (PCATDs) to be effective in maintaining currency and enhancing instrument proficiency.

The largest non-airline LSB found being used was 2:3 in a mocked-up Cessna 172 aircraft (Macchiarella, Brady and Arban, 2005). This research, however, did not propose to test whether the student pilot was ‘to standard’ in the air but rather use the simulator training as a stepping stone to shorten the follow-on airborne training.

D’Alessandro found that PC-based simulators did successfully provide training for introductory skills and tasks but were less able to transfer training as the tasks became more complex. He believed this to be a function of the fidelity of the devices on which the training was being undertaken, underlined by the fact that there was a point at which the transfer of training became less effective for each task. This matches the conclusions of Salas et al who offer that simulation should be considered as a tool for training rather than training itself and that development of the tool must concentrate on the learning rather than striving for realism (Salas, Bowers and Rhodenizer, 1998). Salas et al support the D’Alessandro findings: showing that low-grade devices are capable of providing the level of training required as they are designed to teach to a specific task; in this example that is procedural flying and basic skills.

The civilian experimental experiences are relevant to the military context through their demonstration of training transfer at the foundation level of military pilot’s skills. The analysis by D’Alessandro and Salas et al. infer that the limitations of these experiments are the simulators themselves and thus increased simulation fidelity contains the possibility of increasing the complexity of tasks that can be trained with them.

2.5.2 MILITARY

Military simulators, by virtue of the funding behind them, have tended to be more complex than the PC simulators the civilian experiments have used. The following section will examine the experimental experiences of the military community and determine if the limits are the simulators, as D'Alessandro states, or a function of human ability.

Once simulation became fully established, in early 1980, and possible benefits became recognised a plethora of articles and experiments were published. By way of précising history to this point a Jacobs et al produced a meta analysis that considered a total of 247 articles, of which only 26 had sufficient information for any statistical analysis (Jacobs *et al.*, 1990). Mirroring the civilian experience of years later these experiments could be grouped into takeoff, approach to landing and landing; the most basic of fighter pilots skills. Nevertheless the major finding was that "simulators consistently produced improvements in training for jets".

The late 1980s and early 1990s saw an increased ability to replicate the visual environment and, subsequently, testing of what the US calls Basic Fighter Manoeuvres (BFM), the RAF calls 'Combat' and the layman calls 'dogfighting' was available. Bell and Waag examined the limited previous literature (Bell and Waag, 1998) and offered the major observation that whilst opinions of the SMEs are a necessary requirement their opinion is of limited validity in a scientific sense. This observation was particularly relevant to the issue of measuring performance as well as having critical relevance to the second of the research question's themes – culture. Research into 1970s and 80s trials (Seaman 1999) supports Bell stating that 'The subjective data produced by [SME] evaluations do not provide the quantitative indices of …performance improvement or training transfer' and that they may not be 'sufficiently sensitive' (1999, p.21). These are critical observations as all instruction both in the RAF and all other western air forces the author has experienced over 23 years in the RAF assess through the use of SME assessment, the implication being that this method is invalid for assessing a transfer of training. Seaman and Bell's view is true if control of both the training and testing environment can be dominated to an extent that all variables remain constant between the sample pilots. The nirvana being the ability to move away from SME opinion and make such scientifically acceptable quotations as "75.26% more enemy striker kills" and "54.77% fewer F-16 [Fighting Falcon] kills" (Portrey, Keck and Schreiber, 2006, p. v). These statistics assume that variables such as radar quality, radar attenuation over range, issues with the display of information, weather, aircraft energy state, radar cross section of target aircraft, formation disposition and target aspect are able to be fixed for each sortie. Some of these may be fixed if a lower fidelity or an emulated simulation solution is used, as the radar performance and weather can be pre-set and each sortie can start from set parameters. This is not able to be done for simulated solutions that utilise rehosted aircraft software. Rehosted software utilises actual aircraft software and as such contact on a target will be achieved as a function of the radar's scan rate, Pulse Repetition Frequency (PRF) in use and pilot's radar settings. Thus it is possible for two targets at the same range to be detected by the radar at differing ranges on two different runs, even if all other variables are fixed.

Thus the opinion that SME assessments do not provide hard figures for subsequent analysis is valid for quasi-transfer i.e. testing the new skillsets in a different simulator of the same aircraft type, or wholly emulated simulator-based trials where variables can be fixed. In the live environment the problem is complicated by the infinite combinations of aircraft positions and energy states that a firing solution can exist for, multiple limits on which a solution would become invalid and a number of rules of thumb that should be applied before and during firing. The most careful experiment will constrain a large number of these variables but an SME, when providing his opinion, will still apply a large quantity of tacit knowledge and gestalt observations for each situation, gained from experience, which is not able to be gathered by quantitative data metrics. It is therefore proposed that in the case of re-hosted simulation or measurement of transfer of training into live flight that the validity of the SME's opinion is increased over and above quantitative measurements.

Further investigation into the experiments examined show that in some instances the Instructor Pilots (IPs) did their assessment from the "enemy" aircraft in a linked device (Seaman 1999, p.20). In this instance the research is insensitive, the subjective assessment being negatively affected as understanding a pilots intentions, abilities and energy states is difficult when separated by (a simulated) 6000' of distance across the combat circle. Finally the type of students tested ranged from Combat Ready (CR) ie training just complete, to highly experienced Qualified Weapons Instructor (QWI) students (Seaman 1999, p.21). In these cases the sensitivity of the results would be reduced as adherence to the aircraft type's Standard Operating Procedures (SOPs), which govern missile shot doctrine and aircraft tactical employment, would be weaker for the CR pilots than the QWI pilots for whom they would be completely ingrained. Thus the further down the training pipeline the experiments tested the more their subjects would act as intelligent clones and similar scores become increasingly likely.

The environment in which the training is tested is also debated. Many of the trials in the meta analysis utilise quasi-transfer testing, whereas a smaller number of experiments test the subjects in live flight. In his 1991 paper 'The Value of Air Combat Simulation, Strong Opinions but Little Evidence' Waag asserts that live flight testing is the only way:

"In other words [live flight test] evidence is the only <u>sufficient</u> condition for establishing the effectiveness of simulation training" (Waag, 1991, p. 4).

7 years later in 1998 having completed a review evaluating the effectiveness of flight simulators for training from all research between 1966 and 1998, he and Bell highlighted the same point - in order for transfer of training to be proven it must be witnessed in the air:

> “Many training researchers believe that such transfer is the only sufficient condition for establishing the effectiveness of simulation training” (Bell and Waag, 1998).

Contemporary meta analysis (de Winter, Dodou and Mulder, 2012) contributes to the opinion by demonstrating that testing using quasi-transfer from simulator to simulator was more favourable (Cohen’s[3] d=0.73) than in true transfer to live flight (d=0.10). The effects being attenuated in true transfer versus that of quasi-transfer. This paper is at first glance contradictory to Bell and Waag’s statement above. However more in depth investigation of de Winter et al’s work highlights a lack of the exact type of tasks being undertaken (repetitive or highly complex) and more importantly the quasi transfer experimentation is, for the large majority undertaken between the years 1995 and 2010. In contrast the true transfer experiments all occurred between 1970 and 1985 yet there is no observation of the effect that improvements in simulation technology would have on the training transfer.

This area of the literature review pertains to the first theme of the research question – pilot performance, and specifically its testing and measurement. The above work indicates that if the aim of the simulator is to train for live flight then it is favourable for that transfer to be assessed *in* live flight. As simulation has improved and live flying costs increased quasi-transfer testing has appeared to become the natural proving ground. This is largely built on the ability to reproduce consistently the testing environment facilitating the natural allies of quantitative metrics and quasi-transfer testing. However as the testing regime intended for the research was to mirror the non-negotiable standard used presently, the tests would be not only non-repetitive but dynamic and complex, thus it was testing in live flight that was believed to be the only method of truly testing training transfer. The inability to conduct realistic scenarios in a quasi-transfer setting was also a supporting factor to this decision: in the case of fast-jet training there are no simulators available that

---

[3] Cohen’s d is an effect size, indicating a standardized difference between 2 means.

would produce an environment capable of the roll rate or the provision of a G onset rate of 9G per second as found in the Typhoon. Thus any expectation of quasi-transfer results to be the same or directly applicable to true transfer would be additionally unfair. Most importantly however the Eurofighter Typhoon fraternity has only one flight simulator type capable of training, thus quasi-transfer experimentation would be transported to a non-representative cockpit which would have an unknown effect on results. Finally, as made clear in the dissection of the research question, this research was to take place within the practical environment, thus the question's second theme of culture was considered to be particularly important. It was believed the research would have a better chance of acceptance from the community if successes and failures were exposed within the present testing regime. For these reasons the decision was made to side with Bell and Waag and test the pilots in live flight rather than use a quasi-transfer methodology. Having made this decision the use of SME opinion rather than quantitative method for measurement was the only available option. Whilst the limits of SME subjective opinions have been highlighted quantitative metrics for testing in live flight have yet to be agreed upon or validated and thus the SMEs assessment, with its reliance on experience, tacit, gestalt observations and widespread usage were to be used.

2.5.3 SAMPLE SIZES

A common element of all the literature reviewed was the (small) sample sizes and limited number of tasks assessed. Vaden and Hall's meta analysis of the effect of motion on training transfer in the preceding 24 years found only 7 experiments, of which the sample sizes ranged from 8 to 36 – giving a mean of 22 (Vaden and Hall, 2005). If Bell and Waag's criterion that witnessing true transfer to live flight is the only measure of establishing simulation effectiveness is applied to these experiments, in effect stripping out the quasi-transfer trials, then the sample sizes mean falls further, to 18. None of the remaining trials were conducted post 1979 and all the transfer to live flight tests were conducted in a T-37 training aircraft, thus falling within Bell and Waag's bounds of a less complex and resource intensive regime. As can be seen in Figure 2-1 the T-37 was a training aircraft, equivalent to the RAF Tucano, rather than the front-line aircraft of the day. As such trials using this aircraft would have cost less to run, allowing the testing of larger numbers of students, but the testing of more basic tasks than the research intended here. Nevertheless the small sample size demonstrates, even in basic military pilot training, the historic difficulty of collecting the numbers that would be academically acceptable in other fields.

Figure 2-1. Cessna T-37 Trainer

De Winter, Joost and Dodou’s work (2012) in the same field used an expanded list of previous experimentation including those from Vaden & Hall (2005) but added helicopters, transport and civilian training aircraft. Of the 24 experiments between 1962 and 2004 examined only 2 were not in Vaden and Hall (2005) and fitted the context of live flight testing and military fast jet. The first of these examined BFM instructor ratings for basic fighter tasks on the front-line fighter of the day – the F-4 Phantom (Pohlman and Reed, 1978). The experiment used 8 pilots trained in a simulator with motion, 8 without motion and a control group of 6 with no simulation. The second experiment also attempted to ascertain the effect of simulator-with-motion, but for air-to-surface delivery of weapons in a fast jet trainer - the Northrop F-5. Again the sample size was comparable, with 8 in the simulator group and 8 in the control (Gray and Fuller, 1977).

The small samples when using front-line aircraft are not limited to motion assessment. Work using the Swedish front-line fighter, the JA-37 Viggen (Figure 2-2) looked at the similarity and differences of the psycho-physiological reactions between simulated and live sorties (Magnusson, 2002). The research examined a complex sortie, equivalent to that anticipated in this research, but used only 6 pilots.

Figure 2-2. JA-37 Viggen

Research into small sample sizes found that they could be defensible under certain conditions. Bacchetti's (2010) research into the field of clinical trials proposed that a pragmatic strategy was to use the maximum sample size that was reasonably feasible. Reasonably feasible was defined as 'practical constraints', 'exhausting the pool of easily studied subjects' or when restricted by cost barriers – the value of the information gathered outweighed the cost to gain it. In this last case this approximated closely to n_{min} - their mathematically-derived value for minimised total cost per subject studied (Bacchetti, McCulloch and Segal, 2008, p. 6). This in turn produced a better projected value to cost. Within the context of the present trial this would equate to increased efficiency of costly assets and an increased throughput of student pilots.

Bock et al. (2002) offered that small representative samples could provide at least four types of valid findings:

a. All or none conclusions. If the entire sample returns the same result the conclusion is *likely* to be true.

b. 'Some' conclusions. Can be considered a negative case of the all or none conclusion.

c. Generating ideas. Small samples may offer an insight into new methods or processes to be further explored.

d. Support to the status quo. Allied to substantial *a priori* evidence a finding from a small sample that supports this evidence may justify the conclusion that the sample supports the *a priori* evidence.

The breakdown of the research question seeks to focus on the first three of these possible findings. The challenge of *substantial a priori* evidence for support of the final finding, status quo, was discounted as previous Front-line fast-jet research true transfer training is not believed sufficient in quantity or of a contemporary setting to qualify as a sufficient body of evidence. Nevertheless the value of the information that could be gathered from a, b or c above is considerable and would inform considerable future expenditure in a field where the MoD has a noticeable paucity.

The intention of this section is to show that of the previous work attempted within this field over the last 50 years there is very limited evidence of the true transfer (ie testing of the transfer of training in live flight) within a military fast-jet context. The research uses very small sample sizes due to the length of time, pilot numbers available and cost involved in front-line training for combat pilots. Like this paper's research all have been intended to guide policy decisions on simulation use and, despite their sample sizes have been included, along with experiments using even smaller sample sizes, in meta analysis to determine the effect on the validity of training transfer. As this research intends to test over similar parameters it must at least match its predecessor's sample sizes. No matter the historical support for small sample sizes in this field it was recognised that these alone would not be sufficient to answer the research question to the depth required, hence the triangulation approach of the intended method.

2.5.4 METRICS

In order to understand the value of synthetic training suitable metrics are required to provide indications of effects. Research into the experiments to date found a number of higher level common metrics used in the majority of experiments.

Percentage Transfer demonstrates the amount of live flying saved by simulation relative to the live environment (Orlansky and String, 1977, p. 26; Roscoe and Williges, 1980, p. 183; Alexander *et al.*, 2005, p. 3).

$$\textbf{Percent transfer (PT)} = ((L_c - L_x) / L_c) * 100 \qquad (1)$$

Where L_c = Learning time of control group in the *live* environment.

L_x = Learning time of trial group in the *synthetic* environment.

The transfer effectiveness ratio; demonstrating the ratio of time saved in the live environment against the time spent training in the simulator.

$$\textbf{Transfer Effectiveness Ratio (TER)} = (L_c - L_x) / S_x \qquad (2)$$

Where S_x = Learning time of trial group in the simulator.

These equations are well used within the literature researched but consider time saved from an individual's viewpoint. As this thesis' research question is situated within the practitioner domain it is possible that limitations of simulator employment is likely to face real-world constraints. To this end the metrics will be modified to measure the number of total resources used:

$$\textbf{Support Flying Transferred (SFT)} = ((Sp_c - Sp_x) / Sp_c) * 100 \qquad (3)$$

Where Sp_c = Support live flying[4] required in the control group syllabus.

Sp_x = Support live flying required in the trial group syllabus.

Support Flying Hour Efficiency Ratio (SER) = Total Live hours (inc support) / Total student live hours required (4)

Both efficiency ratios can be broken down into cumulative or incremental figures to determine when the efficiency or effectiveness curves drop below an acceptable level to the organisation or researcher (Roscoe and Williges, 1980, p. 187). These metrics are robust indicators of efficiency when comparing comparable syllabuses however they make the assumption that the student that emerges is of suitable standard. To determine if this is the case comparison can either be done using quantitative or subjective measures although, as argued in Section 2.5.2 above, the decision on which to use is largely determined by the testing location. Portrey's (Portrey, Keck and Schreiber, 2006) work testing teams within in the synthetic domain attempted to capture the data shown in Table 2-1.

Table 2-1. Synthetic Metrics used by Portrey (2006).

Individual aircraft data	Range at missile launch Mach at missile launch Loft angle at missile launch Altitude at missile launch Percentage maximum at launch Escape G at launch G-load at missile launch Distance of miss Clear avenue of fire
High order data	No. of enemy strikers reaching target Closest distance of strikers No. of F-16 mortalities No of enemy strikers killed before base Total no. of enemy threats killed 'Top Gun' summary scoring scheme

[4] Support flying are the aircraft required to make up the formation and opposition over and above that used for the trainee.

Portrey's metrics required 24 months to program into the simulator's software (Portrey, Keck and Schreiber, 2006, p. 11) but undeniably capture the most comprehensive list of measurements relevant to front-line training to date and provide a ready method to compare teams of peers against each other. This is the clear benefit of testing within the synthetic domain where data can be captured from databus and used to provide the figures above. As argued previously, however, the relevance and applicability of these results to combat assumes that all of the simulator's models, e.g. missile flyout, impact of weather on radar, jamming etc. contain no differences to the live environment and a corresponding assumption that a transfer of training to live flight would take place.

In the live-flight environment SME Sortie Report Forms (SRFs) are used by all major air forces to ascertain if a pilot meets the minimum standard. The Typhoon SRF, see Figure 2-3 is developed from work by Dstl which defined the Mission Essential Competencies (MECs) required to be resident in each pilot (Dstl, 2009) and represents a typical example of such a form. Each task is graded between 0 and 5 with 3 being assessed as a pass standard. Example definitions for each task and the definition of each of the numerical grading can be found in Appendix A'. The SRF's weakness is that it provides little by the way of quantitative measurement, unlike the synthetic environment, however its use in the live environment does definitively prove transfer of training against an experientially-based standard.

SORTIE GRADESHEET	Sortie	Formation Number	Duration	Date
	COMBAT(FLY)01	2 of 2	1:15	28 Apr 2011
Name	Course	Aircraft Block	Instructor	
GAZ	OPCON(AD)5	5 (4.2)		

TYPHOON FEB 2011 / OPCON(AD) COURSE / MISSION ELEMENTS	UNKNOWN	GRADE 0	GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5
1. Mission Prep	☐	☐	☐	☐	☐	☐	☐
a. Threat of the day	☐	☐	☐	☐	☐	☐	☐
2. SUTTO	☐	☐	☐	☐	☐	☐	☐
a. RATD	☐	☐	☐	☐	☐	☒	☐
b. B-Scope lead / lag	☐	☐	☐	☐	☒	☐	☐
c. B-Scope targeting	☐	☐	☐	☐	☒	☐	☐
d. Fade / clean procedures	☐	☐	☐	☐	☐	☒	☐
3. Tactical Domestics	☐	☐	☐	☐	☐	☐	☐
a. WSC	☐	☐	☐	☐	☐	☒	☐
b. Combat fence checks	☐	☐	☐	☐	☐	☐	☐
c. G-warm	☐	☐	☐	☐	☐	☐	☐
4. Gunex	☐	☐	☐	☐	☐	☐	☐
a. Resolve POM	☐	☐	☐	☐	☐	☒	☐
b. Resolve Range	☐	☐	☐	☒	☐	☐	☐
c. Resolve Lead	☐	☐	☐	☐	☒	☐	☐
d. Energy management	☐	☐	☐	☐	☒	☐	☐
e. Effective gun employment	☐	☐	☐	☐	☒	☐	☐
5. WVR 1v1	☐	☐	☐	☐	☐	☐	☐
a. Turn circle recognition	☐	☐	☐	☒	☐	☐	☐
b. Lag BFM execution	☐	☐	☐	☐	☒	☐	☐
c. Kill or BFM	☐	☐	☐	☐	☒	☐	☐
d. Maintenance of the offensive	☐	☐	☐	☐	☒	☐	☐
6. MPH	☐	☐	☐	☐	☒	☐	☐
a. Lead turn execution	☐	☐	☐	☐	☒	☐	☐
b. LV placement (& maintenance)	☐	☐	☐	☐	☒	☐	☐
OVERALL GRADE (MINIMUM GRADE 2 REQUIRED FOR PASS)					3		

SORTIE OUTCOME			
DCO	☒	☐	INCOMPLETE
DNCO STUDENT	☐		
DNCO OTHER	☐		

Remarks (continue on reverse side)

ACCOMPLISHED: AS per syllabus

OVERALL: 1 x Bubble bust on gunnex

SPECIFICS: RATD. Good radar handling and communication with the lead, starting to bring in the Elev page without prompting. Very minor - needs to be closer to the 1nm range rather than 2 as this will allow a more rapid transition to battle via the WSC.

WSC - satis
Fence - needs to respond to lead with ' XXX2 fenced in, code 0'
G warm - correctly flown and capacity indicators atarting to expand ie. recognising and commenting on amoung of g being pulled.

GUNNEX - POM is correct with small error corrections being used, still missing but by less of a margin and using correct techniques. Range - needs to control his range by starting his turn towards earlier/later as required.Presently the range on the first split is correct but the second is at 6000' vice <3000'. (This is partially a function of the student lead presently being utilised).
Management of lead is correct during the 'ease' and GAZ is starting to understand the concept of energy management as he approaches the Gun ranges as demonstrates by his throttle control. To improve the effectiveness of his Gun work GAZ needs to recognise the moding the gun is in and move up the weapons tree where available.

1v1 - GAZ is able to recognised the correct turn circle on the first iteration but having taken a F2 and eased his return to the circle is late resulting in misalignment. This is caused by a recognition that he is on the turn circle followed by a pull - thus leaving him outside. If GAZ alters this to 'approaching the turncircle' followed by the pull he will arrive on the circle with less work to do.
Still fights the tendancy to roll out rather than ease but his eases are better and within the 10-12 alpha limits.
Kill or BFM recogntion is satis as is his LV placement.

Instructor Signature	Film Review	Stud Initials	Sup Initials	Next Instr

Figure 2-3. Example Sortie Report Form

The above examples show a number of accepted high-level metrics used by previous work. The method of more detailed data collection is determined by the testing environment chosen and varies between quantitative measurement and an observed assessment of ability. As the research question is targeted at understanding how much of Eurofighter's training can be ported into synthetics the only way of determining if the transfer of training has occurred is to test in the live environment. In making this decision the use of SME assessment through the use of SRFs is the only method that has historical and accepted precedence.

2.6 DEVOLVED THEORIES AND PRACTITIONER USE

The literature examined so far has almost exclusively concerned itself with whether simulators can provide any valuable contribution to the training of the pilot. This, however, provides only a single perspective on understanding how the research question should be tackled. The challenge of finding the 'extent' to which simulation can be used requires investigation of the limits of implementation to date. Thus an examination of present practical implementation was conducted. The implementation of the field's research within a practical environment to increase the LSB has been pioneered by the F-15 and F-16 communities allied to the US Air Force Research Laboratory (AFRL). The present practitioner's standpoint is represented by Figure 2-4 below.

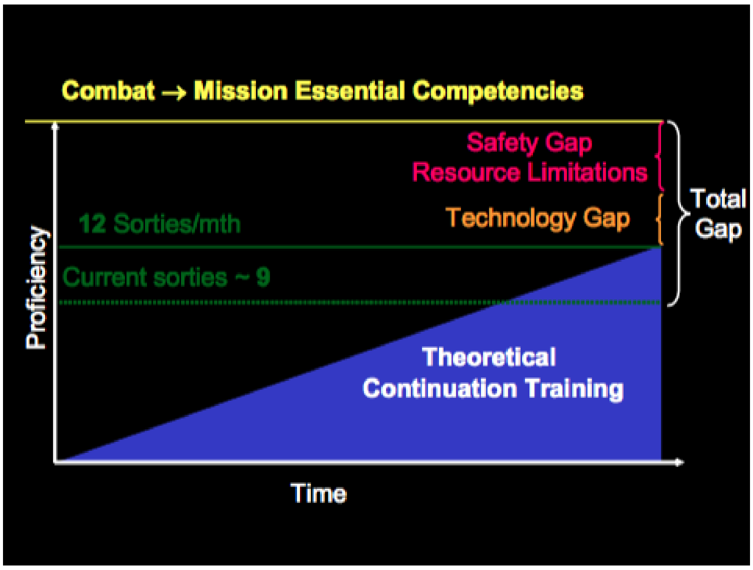

Figure 2-4. Total Current Training Gap. Reproduced from McGrath (2005)**.**

The view is that simulation should fill the gap between Continuation Training (C-T) i.e. day to day training and a theoretical maximum proficiency that combat operations would demand. The gaps are generated via a lack of aircraft, the need to preserve peacetime safety margins, such as preventing highly dynamic manoeuvres close to other aircraft, or increasingly the need to train pilots to understand latest software or hardware enhancements. McGrath further highlights the effects of future funding cuts with the number of monthly sorties available to a pilot to be cut from 12 to 9. The result is that the available peacetime flying hours are spent servicing the training requirements that are within these bounds and simulation is targeted to achieve 'gap' training.

A major tenet of the theoretical maximum proficiency is arrived at by application of theories examining knowledge and skills and their links to mission competency (Symons, France and Bell, 2006). Mission Essential Competencies (MECs) begin by sourcing the strategic level aims for air power and works downwards determining the competencies required to achieve these aims, with these competencies further broken down into knowledge and skills required by a pilot. Finally the experiences required to achieve, or at least provide an opportunity to achieve, these requirements are determined. In the USAF F-15 community this work led to a reassessment of the training syllabus and pushed elements from the 'live' environment to the synthetic resulting in a LSB ratio of 3:1 for Basic Mission Capable pilots (Baldwin, 2008). The need for this reassessment of the LSB and the desire to increase efficiency of the F-15 RAP (Ready Aircrew Program) Syllabus was driven by auditor observations of over-flying of certain sorties (Baldwin, 2008, p. 3), 'nebulous allocations of flying hours' (Ibid, p.24) and indiscriminate cuts in the flying budget and training programme (Ibid, p.4). The F-15 fleet, facing similar problems to the RAF, have recognised the need for increased use of simulation in order to increase the value of the live flying hours (Ibid, p.34) and have amended their simulation syllabus to do so. The work was mirrored for the Typhoon Fleet (Dstl, 2009) in order to highlight possible gaps in training but failed to result in any increase in the synthetic proportion of the LSB.

The F-15 issues illustrate the difficulties of the level of technology comparable with that of the Typhoon. The US' experiences with the much more advanced F-22 provide an understanding of possible vectors of these issues. The F-22's experiences are relevant as, like the Typhoon, the need for simulation is focussed by the high cost per flying hour - $68,362 (Thompson, 2013). Ausink et al. (2011) found that the lower number of training events accomplished per month was due, primarily, to having to fly red-air (enemy opposition) missions in order to provide training value for those pilots receiving missions that did qualify as training events, an issue also recognised in the Typhoon and F-15 fleets (Lockwood, 2006, p. 36; Marken *et al.*, 2007, p. 6).

Ausink et al’s solutions for reducing the gap and comments of their applicability to the Typhoon context are given below.

a. **Increase both live and simulator hours available**. This requires and depends on funds being made permanently available increasing in line with the number of pilots that require training as the fleet grows.

Figure 2-5. Dassault DA20 Falcon - 'Red Air'.

Figure 2-6. BAe Hawk - 'Red Air'

b. **Increase amount of red air being outsourced to other units**. The aircraft being used for this role in the RAF are the externallycontracted Hawk and DA20 Falcon (Figure 2-5 and Figure 2-6). Both have strengths but neither is capable of replicating the level of performance required as they are primarily a training aircraft and business jet respectively, consequently Typhoons continue to be used to provide a realistic adversary. Simulation, however, provides the opportunity to fight against threats with the correct flight envelope replicated with high fidelity.

c. **Develop simulation capabilities to provide the desired threats, or link to other simulators that can**. This requires software to provide a limited artificial intelligence to the enemy fighters or additional manned simulators that are linked by a network. Whilst ASTA is unable presently to link externally to other non-Typhoon simulators (Wong, 2010) it does have a basic tactical capability referred to as 'the doctrines'.

d. **Develop the Live Virtual Construct (LVC) capability**. The ability to have live and simulated aircraft simultaneously broadcasted within each other's environment. Live pilots will see simulator or constructed entities on their sensors whilst the simulator pilots will see computer generated images of friend and foe. Comparative work to develop this idea has been conducted in the UK by Dstl (Anderson, Walls and Read, 2011). Practical application has been prevented by the need for the pilot to see realistic representation of the fake entity on the sensors in order to prevent incorrect decisions and motor skills being formed. This is possible but requires access to a section of the aircraft software that would require significant financial investment. As such further development of this area has stalled.

Of Ausnick et al.'s observations the only option that will reduce the wastage of assets used for 'red air' and that does not require significant financial investment or require a leap over a substantial technical hurdle is the provision of a realistic threat environment through simulation. This provides a reasoning for the natural focus on simulation over the last few years as defence is made more accountable for its spending through the regular defence reviews.

Thus, with reference to the research question, the simulator should be able to demonstrate competence in the provision of enemy air replication rather than the previously tested tasks of just landing or instrument flying of the historical experimentation. This will ensure the 'extent' of simulation use referred to in the question carries forward into the postgraduate training of the Combat Ready Work Up and will affect the LSB of the Front-line Squadrons and their day-to-day training.

2.6.1 LIMITATIONS OF SIMULATION

All of Ausink et al's observations of how to reduce the training 'gap' have an element of simulation within them however employing simulation in these areas with no understanding of where the limits of training in simulation lie would be counter productive and financially wasteful. Fortunately work done by the F-15 Fighter community sought to make explicit these limits (Seaman, 1999):

a. **The Cultural Limit**. Seaman quotes General Richard Hawley, the commander of Air Combat Command as saying that their simulators were used "to learn some basics about the weapon system, learn to start the motor, how to employ the radar […] but [not to] learn the essence of the business, which is team combat." (Seaman, 1999, p. 2). In 2010 this same methodology was still being used by the RAF (29 Squadron, 2011a).

b**.** **Simulator Design Limitations**. Poor image projection, incorrect cockpit layout, non credible threats and overly accurate information provided by the Fighter Controllers (FCs) all contributed to limit the employment of the simulators; as originally found by Houck et al. (1991). In contrast to these limits Payne's work argues that the potential to provide negative learning because of these limits can be negated simply by highlighting the differences between the real and simulated environment to the pilot and results in a valid transfer of training to take place (Payne, 1982).

c. **Physiological Issues**. Seaman theorises that it is not possible to reproduce the same level of physical exertion or stress as the live environment (Seaman 1999, p.33). This is partially contested by Magnusson's (2002) research which examined similarities and differences in psychophysical reactions between simulated and real air to ground missions and found 'little difference in the reaction patterns on the psycho-physiological variables between simulated and real flight', although there was significant difference in the mean level of each variable, such as heart rate (Magnusson, 2002, p. 59). Magnusson hypothesises that the reasons for this are that the pilots either react to the mental workload or their bodies react as though the mission is real rather than simulated.

Just as understanding limits will help avoid the positioning of simulation within areas it is unlikely to add benefit, an understanding of its strengths will facilitate the placement and maximise affect. Houck (1991) questioned 87 F-15 combat pilots and found the events shown in Table 2-2 were better when trained in the simulator.

Table 2-2. Simulation Strengths. Adapted from Houck (1991, p.10).

Simulation Better than Live	Valuable Simulator training but Not Better than Live
1. Multibogey, Four or more enemy 2. Reaction to SAMs 3. Dissimilar Air Combat Tactics 4. All weather employment 5. ECM employment 6. Communication Jamming 7. Low Altitude Tactics 8. Chaff / Flare Employment 9. Escort Tactics 10. Working with FCs	1. All Aspect Defence 2. Beyond Visual Range (BVR) Employment 3. Radar Sorting 4. Missile Employment 5. Egress Tactics

Portrey (2006), however, provides quantifiable evidence of improvement of the tasks in the right hand column of Table 2-2 measured in simulated scenarios given to 76 teams of F-16 pilots. This improvement over the 1991 results could be down to the different aircraft type or, more likely, the improvement of simulation technology in the intervening years.

In sum previous research has identified both strengths and weaknesses of simulation when used in a tactical environment. This guided the present methodology of research, allowing efforts to be focussed on areas that had been historically poor in simulation.

2.7 LITERATURE REVIEW SUMMARY

The literature review has sought to provide an understanding of the background to the problem from both a high level (Parliament and Air Ranking Officers) and a lower level (tactical) standpoint. The RAF's understanding of the abilities and intentions of simulation has been found through examination of papers and reports leading up to and including the period of the SDSR 2010. These have led to an appreciation of the motivation surrounding the desire for increased employment of the technology. At the same time the factors of concurrency, functionality, reliability and culture were identified as barriers to the increase of the proportion of synthetics in training.

Study of the extent of simulation use in both the civil and military sector has found the maximum proportion of simulation used in the civil world, excluding airlines, to be an LSB of 2:3 (Cessna training in Section 2.5.1) within the military this reduces to 1.7:1 (Rafale training in Section 2.3). As a rule these simulation sorties have followed the cultural pattern of pre-flight preparation rather than being used as an environment where *all* required training can be given.

All the experiments into the effectiveness of simulation within the military sector discussed here have used 8 pilots in each group and using limited exposures to specific areas of training, for example instrument flying, landings or BFM. These exposures have been for short time periods and used less than ten simulator sorties during the training or multiple runs of a task of less than 5 minutes, with the most live sorties used being 6 per student (1975 US Navy study quoted in Seaman (Seaman, 1999)). The purpose of these sorties was to train the students in the techniques required for live flight so that the studies might search for an improvement in student quality. None of the research viewed simulation as an entire replacement for the live environment, only preparation for it; to this end the desire to establish if the output was better than live training is understandable. Sample sizes for the research have varied but as the complexity / cost has increased the sample sizes have fallen, with research using live testing in Front-Line combat aircraft using a minimum of 6 (Magnusson 2002) and a figure of 8 for both Pohlman and Reed (1978) and Gray and Fuller (1977). The testing environments, both live and synthetic, have also been discussed with advantages and disadvantages found for both. Incorporated within this was the method of metric measurements for both environments as well more overarching equations to determine the effectiveness of the training. The final section looked at the experience of practitioners when employing the research to date: the limiting issues, maximum LSBs attained and assessment (by current combat pilots) of areas that were better trained in the simulator.

Chapter 3. TRIAL PANDORAS BUZZARD

3.1 OVERVIEW

Trial PANDORA'S BUZZARD intended to determine if the utilisation of a heavily synthetic syllabus was even feasible when training fighter pilots to fly and fight the Eurofighter Typhoon. Unlike previous research that targeted specific skillsets over a few missions this work sought to replace the entirety of Operational Conversion Unit training with simulation, thus ensuring the simulator was the only location a student would have gained the knowledge used in live flight. Allied to the strengths and weaknesses, the trial intended to expose the results were to inform the RAF which areas appeared valid to exploit further when increasing their synthetic proportion of the syllabus. Finally the trial hoped to expose second-order effects such as squadron structure and the impact on instructor hours that had hitherto been concealed.

3.2 KEY LITERATURE REVIEW DIRECTION

Having presented the historical issues within the literature it was clear that a number of key decisions needed to be articulated in order to establish direction and thus determine the research methodology, the largest of these being the testing environment i.e. was the success or failure of the trial to be judged by assessment of standards within the simulator (quasi-transfer) or by witnessing standards in live flight (true transfer), see section 2.5.2. After consideration the live environment was selected for testing, the major reason aligning with Bell and Waag's 1998 view that live testing was the only true way of validating a transfer of training. Had this research intended to train within the cultural norms of live flight preparation, or *not* wished to determine 'limits' then quasi-transfer training would have offered clear advantages. The intention to train student pilots entirely within the simulator for each phase of an entire Operational Conversion Unit of a front-line combat aircraft represented a new exploration of limits, and as such validating that training in anything other than live flight would not be acceptable to either academic or military peers;

particularly if the research was intended to act as a catalyst for a cultural change.

A consequential issue forced by the election to test in the live environment was the difficulty of generating a sufficiently large sample size as cost and complexity of the skillsets increase and available pilots decrease; discussed at length in section 2.5.3. The literature review highlighted that, whilst Magnusson (2002) may have used only 6 pilots, a figure of 8 pilots in the trial group was more normal and acceptable to the research field, thus requests to Air Command for trial pilots must at least achieve this figure.

3.3 CONTEXT

The trial was intended to be multi-disciplinary in nature, incorporating empirical data and statistical analysis of existing data, as review of the literature had indicated that the effects of practical integration of simulation into training was not limited to the simulation devices themselves. To this end the research was to be undertaken with a ‘systems’ mind set, the high-order areas and aims of which are shown in Figure 3-1.

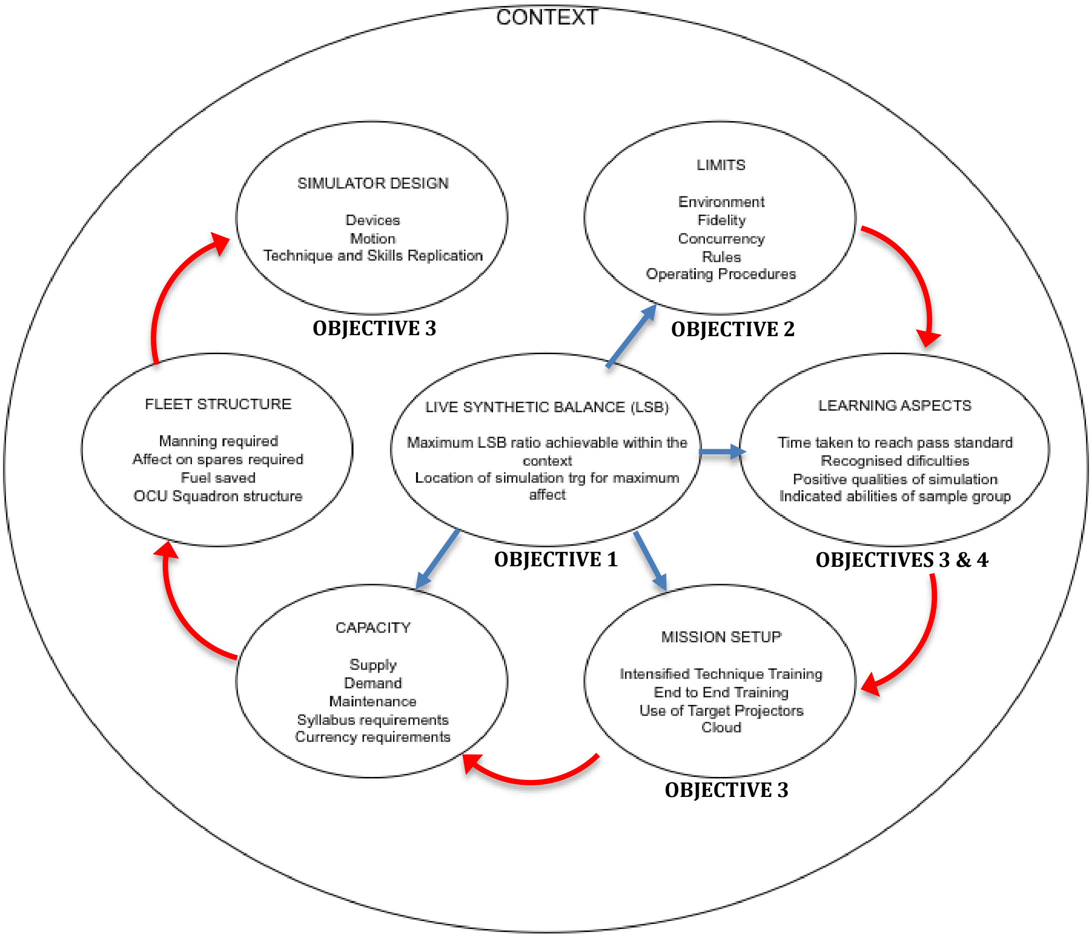

Figure 3-1. Systems Approach High Level Intentions

Multiple interactions between the groupings was anticipated, a logical and coherent thread throughout the research was therefore necessary to provide initial findings before any further iterative cycles could be considered, thus the objectives referred to in Figure 3.1 refer back to those identified during the dissection of the research question leading to Sub-Question 1 (section 1.8.1), to ensure the link to the research question remains explicit. The flow of the research followed the logic arrows in Figure 3-1; an understanding of the maximum achievable LSB on the OCU would lead to an appreciation of areas where simulation could provide maximum effect, the failures in these areas informing the 'limits'. Learning aspects and mission setup observations will be sourced from the Sortie Report Form narrative and in combination with the determined LSB the required capacity of the present simulation devices can be found. This in turn will provide a desired fleet structure and help inform the weaknesses in present simulator design such that they may be addressed in future procurement.

In order to bound the problem and provide a conceptual limit Objective 1 of Sub-Question 1 identified that the initial research should consider only discreet training packages within the training of Typhoon pilots rather than day-to-day training done by Combat Ready (CR) pilots on the front line. The use of discreet training was intended to take advantage of defined packages with stated aims and objectives, extant scoring matrices and an established end-of-training pass standard. These features provided a conveniently comparable ‘end-of-phase’ standard for test subjects to meet and, importantly, a number of pre-trained assessors that have been demonstrated to adhere to a standardised level of assessment via regular STANEVAL (STANdard and EVALuation Flight) assessments of the instructional level and ability of each of the Instructor Pilots (IP).

3.4 AIMS

The discreet training packages targeted for the research within Trial PANDORA’S BUZZARD and the intended aims for each of the packages are listed below:

a. **Phase 1**. Day conversion to the aircraft; the training of students up to and including first solo.

i. Determine if results of the Instrument Rating Test (IRT) are affected by a fully synthetic training lead in and test when compared with results from the standard syllabus, tested in live flight.

ii. Determine the feasibility of achieving a successful first solo from a fully-synthetic syllabus.

b. **Phase 2**. Combat – Basic Fighter Manoeuvres (BFM); short range fighting within the visual arena.

i. Determine if the Combat Trial Syllabus should be recommended for inclusion into the core syllabus with particular reference to the likelihood of a fully synthetically trained student passing the End-of-Phase (EoP) check ride when compared to a student trained using the standard method.

ii. Identify any limiting factors preventing the further exploitation of simulation to train within the BFM environment.

c. **Phase 3**. Counter Air (CA); long-range fighting outside visual range using longer range missiles.

i. Determine if the Trial Syllabus should be recommended for inclusion into the core syllabus with particular reference to the likelihood of a fully-synthetically-trained student passing the End-of-Phase check ride when compared to a student trained using the standard method.

ii. Identify any limiting factors preventing the further exploitation of simulation to train within the Counter Air environment.

d. **Phase 4**. Quick Reaction Alert (QRA); the enduring role of the Air Defence pilot is to protect the homeland during peacetime and provide an airborne “police force”. This ride is conducted against a low slow aircraft simulating a civil aircraft lost below cloud.

i. Establish the probability of a fully-synthetically-trained student passing the QRA End-of-Phase check ride when compared to a student trained using the standard method.

ii. Identify any limits or risks associated with the QRA phase.

3.5 DETERMINING THE RELEVANT FACTORS

Before commencing the trial the measures to verify its success, or otherwise, and factors that affected that success needed to be determined. Having chosen ‘true transfer’ of training using live flight as the test environment the measure of success would be determined by SME marking of the sorties as argued in section 2.5.4. As this was the scoring scheme normally used within Typhoon training this methodology would allow the historic records to be investigated to provide the benchmark that the Trial group would need to meet.

Access was gained to all the Typhoon training archives which sourced 7 years of training records, back to 2004. Investigation of the syllabus and training methods found that between 2004 and 2006 the syllabus and techniques were not suitably similar to that intended to be used for the trial group’s syllabus, thus these years were discounted. From 2006 onwards the archives showed the syllabus to maintain a stable and comparable format to the trial. This provided results for all the pilots that had passed through the OCU over the preceding 5 years – a population of 57. Scores for each phase for all 57 students were recorded, however the first 27 had only completed the first 3 phases; the 4th phase, QRA Low Slow, being introduced in 2009. Thus the records contained 3 phases with 57 results and 1 phase with 27. The descriptive statistics for each of the phases and the sum of the marks for those that conducted 3 and 4 phases respectively can be found at Appendix B.

Now that the style of measurement had been decided further investigation was required to determine any correlating variables that would provide key predictors that drove the resultant marks. In doing so it was intended that this would inform the type of students that would be requested for the trial itself. Having already conducted multiple selection procedures, such as officer selection and elementary, basic and advanced flying training, individuals meeting the entry requirements to be posted to the Typhoon were a relatively homogeneous population. It was postulated, however, that the predictors likely to correlate to the scores were those that affected an individual’s experience and ability:

a. **Total flying hours.** A clear measure of experience in aviation. It was expected that more-experienced pilots would perform better than less-experienced ones.

b. **Rank**. Promotion being based, in part, on ability in the air meant that senior officers such as Squadron Leaders and above were expected to achieve higher grades than the Junior Officers of the rank of Flight Lieutenant and below.

c. **Age**. At first glance this complements the measure of Total Flying Hours above, however it had often been commented that the younger pilots known as the 'Playstation generation' had an advantage over their older colleagues through the similarity of computer gaming to the highly computer-driven Typhoon. Thus this metric was included to incorporate a generational observation.

3.5.1 BINARY LOGISTIC REGRESSION: RANK VERSUS 3 PHASE SUM

Binary logistic regression was used to determine if there was a significant relationship between the sum of the marks given for the final flight of each of the 3 phases of IRT, BFM and Counter Air (to be hereafter known as the 3 Phase Sum) and the Rank of the pilot: Senior Officer (SO) or Junior Officer (JO).

```
Link Function: Logit


Response Information

Variable  Value  Count
Rank      SO        12  (Event)
          JO        42
          Total     54
```

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Constant	-3.62547	3.01349	-1.20	0.229			
3 Phase Sum	0.210940	0.263988	0.80	0.424	1.23	0.74	2.07

Log-Likelihood = -28.271

Test that all slopes are zero: G = 0.667, DF = 1, P-Value = 0.414

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	2.58692	4	0.629
Deviance	3.52163	4	0.475
Hosmer-Lemeshow	2.22172	3	0.528

Table of Observed and Expected Frequencies:

(See Hosmer-Lemeshow Test for the Pearson Chi-Square Statistic)

Value	Group 1	2	3	4	5	Total
SO						
Obs	2	0	4	3	3	12
Exp	1.3	0.9	3.4	4.0	2.3	
JO						
Obs	7	5	12	13	5	42
Exp	7.7	4.1	12.6	12.0	5.7	
Total	9	5	16	16	8	54

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	233	46.2	Somers' D	0.15
Discordant	157	31.2	Goodman-Kruskal Gamma	0.19
Ties	114	22.6	Kendall's Tau-a	0.05
Total	504	100.0		

The measures of association above are measured between -1 (when all pairs of variables disagree) and +1 (when they all agree and therefore are associated). The results show values between 0.05 and 0.19 hence it can be said that the poor measures of association are poor. Similarly the test for a relationship between the rank of the pilot and their '3 Phase Sum' did not return a statistically significant result (P value = 0.414). Maximum Likelihood Methods however, can be biased towards small samples and as such an Individual Value Plot was generated to provide a intuitive visual check of the lack of association (Figure 3-2). Again this plot highlighted no definitive difference in scores between the 2 groups, thus the variable of rank was considered not to be a factor.

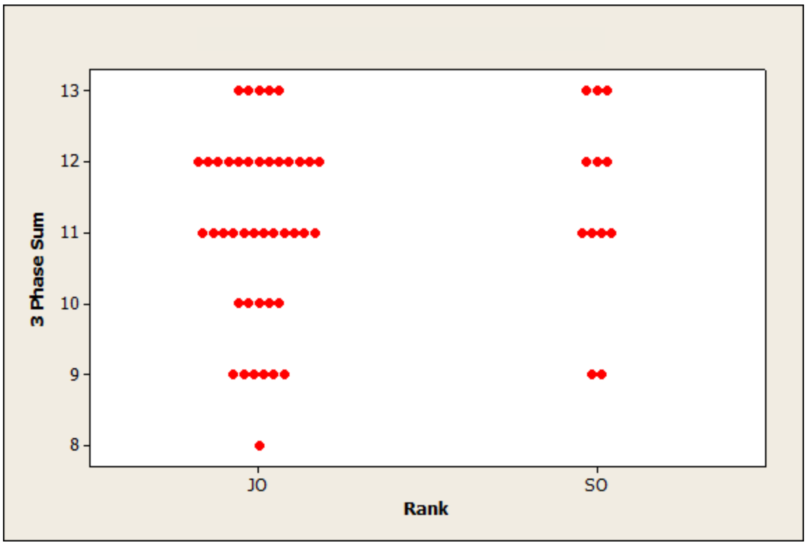

Figure 3-2. Individual Value plot for Rank vs 3 Phase Sum

3.5.2 REGRESSION ANALYSIS: 3-PHASE SUM VERSUS TOTAL FLYING HOURS AND PRESENT AGE

After the removal of rank as a factor, Regression analysis was used to determine if there was a relationship of age or total flying hours to the 3 Phase Sum (sum total of IRT, BFM and Counter Air phases). Minitab software identified 3 unusual observations and examination of these results found two that were unusual: one had had significant family issues during the course whilst the other had left the fleet after the OCU and prior to joining a front-line squadron. Thus both these results were removed and the analysis re-run (below), this gave a higher 3 Phase Sum result for a given flying hour and age.

```
The regression equation is
3 Phase Sum = 12.8 + 0.124 Total Flying Hours (00s) - 0.109 Present age

Predictor                      Coef  SE Coef      T      P
Constant                     12.810    1.948   6.58  0.000
Total Flying Hours (00s)    0.12370  0.05638   2.19  0.033
Present age                -0.10858  0.08338  -1.30  0.199


S = 1.27178   R-Sq = 11.1%   R-Sq(adj) = 7.6%


Analysis of Variance

Source          DF      SS     MS     F      P
Regression       2  10.326  5.163  3.19  0.049
Residual Error  51  82.489  1.617
Total           53  92.815

Source                    DF  Seq SS
Total Flying Hours (00s)   1   7.583
Present age                1   2.743
```

This new analysis demonstrated a significant correlation with total flying hours, suggesting that the value of ‘3 Phase Sum’ increased by 0.124 for every 100 flying hours the pilot had amassed (P = 0.03, thus statistically significant at the alpha level of 0.05). Age however returned a value of the coefficient that was not statistically significant, P = 0.19, and as such the age was removed from consideration as a factor. Thus selection of students for the trial group was made with a consideration of the variable of ‘Total Flying Hours’.

3.6 3 TESTS TO MEASURE SUCCESS

As previously highlighted the design of the research was bounded by both academic and practical limitations, in addition there were large financial, political and risk implications associated with recommending a transfer out of live flight into the simulator, therefore the intent throughout the research was to favour the status quo of training in live flight. Thus where any marginal decisions on the measure of success were to be made the metric will always be adjusted to favour training in live flight, in this way training in the simulator will have to provide significant and robust results to show in favour of simulation.

The resultant regression of the factor of ‘Flying Hours’ returned the linear equation shown in Figure 3-3. Whilst it is evident that this equation does not have any reasonable predictive power given the weak R-Sq values the result is valuable in that it shows the effect of the only significantly correlated factor on the sum of the scores across the 3 phases. Given that the sample contains every pilot on record that passed through the OCU since 2006 this relationship has been proven to be acceptable to the Typhoon Instructors when graduating students. Thus, given the small sample size likely to obtained for the trial, the trial subjects should not only prove themselves within each phase but be seen to match this historic relationship. Whilst it is possible to pass the course with lower marks, if the trial is to be considered successful by peers then the 3-phase results should be no less than the integer beneath this line (sorties are marked in integers, a decimal result thus not being possible).

<u>TEST 1</u>: The total of the Trial pilot's results across all 3 phases should fall no less than the closest integer beneath 10.33 + 0.06* Total Flying Hours(00s)

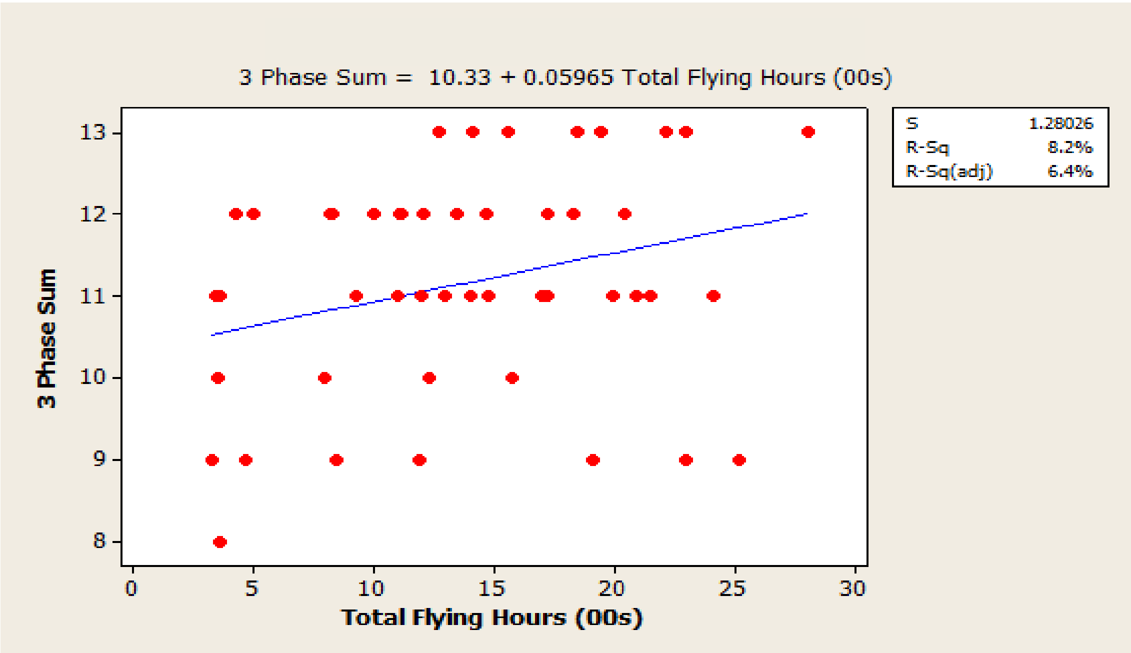

Figure 3-3. Linear Regression of Total Hours Vs Sum of the Scores of the 3 Phases

Given that the removal of the unusual results increased the gradient of the equation the intent that the research should favour the status quo has been met. However, as previous students have passed the course with lower marks the need to satisfactorily meet Test 1 is primarily a cultural requirement that would promote acceptance of the Trial. The stated requirement to pass each phase of the course is a score of 3 – satisfactory, in each examination sortie at the end-of-phase. Given the small sample size Test 2 was set such that *all* trial subjects must pass the End-of-phase check ride in order for that phase to be considered for recommendation for inclusion in the core syllabus as a synthetic phase.

<u>TEST 2</u>: Each of the test subjects must pass (score 3 or higher) the EoP check ride for the phase to be considered for recommendation for inclusion in the core syllabus as a purely synthetic phase.

Finally and in keeping with the intent to favour the status quo the mean score for the control group should not be shown to be higher than that of the trial group. Given the small sample expected and the supporting evidence from Tests 1 and 2 the P-value for Test 3 was set at 0.1.

TEST 3:

The hypothesis (H_0) - 'The mean value of the EoP Score for the standard course is greater than the Trial course' should be disproved.

3.7 EXPERIMENTAL DESIGN

3.7.1 EXPERIMENTAL SUBJECTS - TRIAL SAMPLE

As identified in the literature review (see section 2.5.3) all comparable previous work has used a small sample size, typically of 8 trial and 8 control. Although the present work is intended to cover a larger timescale and skillset it was felt that any less than an equivalent trial size would not be able to be justified. Larger samples, whilst highly desirable, were unlikely for reasons of cost and resource efficiency; as an illustration each pair of pilots took approximately 5-6 months to progress through the trial syllabus.

As the sample size was small the criteria for the pilots intended for the simulation syllabus were selected carefully to match those on a standard live syllabus. The method for selection was chosen as a stratified random sample (Bryman, 2008, p. 173) with groupings of the only correlated factor – Total Flying Hours. RAF 1 Group (HQ) and Human Resources were requested to provide the Trial with the numbers of students required for each flying hour grouping (see Table 3-1).

Table 3-1. Sample Groupings

Hours Groupings	Population Sample (n=57)	Trial Sample (n=8)
1-1000	20	3
1001-2000	26	4
2001-3000	11	1

This method was chosen as it addressed a number of pragmatic factors:

a. The primary researcher would have control over the requirements of the sample but would not be able to select individuals that were known to him. In this way the possibility of selecting individuals with a high probability of passing was removed.

b. Unlike a laboratory trial the subjects available had already passed through 3 years of training that removed all those unsuitable for fast jet training, thus the numbers exiting the training pipeline and available for Typhoon training during the trial's period were small. This sampling method allowed the sought-after attributes to be designated without unrealistically asking for all 8 trial subjects to come from a single grouping or restricting 1 Group's career plans for individuals. This pragmatic addressing of the politics secured the command chain's support.

The student pilots (to be known as Trial Subjects) provided by 1 Group (HQ) were commensurate with the strata requested; the pilots' hours matched those required by Table 3-1. Additionally all pilots had been deemed to be of a suitable standard for Typhoon Training by the Training Board at HQ 1 Group, RAF High Wycombe and, like all their colleagues, all pilots were of a high average standard as stated in their F5000 (Flying Appraisal Reports). Thus entry standards matched historic norms.

Once the conversion phase of the trial was complete for the Trial Subjects 1 Group HQ would go on to sanction the use of the synthetic syllabus to train, were necessary, standard course students (these would become known as ‘The Augmentees’). Although this was not planned for at the commencement of the trial the small number of Augmentee results are included in the Conversion phase results as the exposure up to the test point was exactly the same as the Trial Subjects. In this manner the sample size, for the Conversion phase only, was increased.

Prior to enrolment on the Trial all students (both Trial and Augmentee) underwent the same pre-employment training as their standard syllabus colleagues. This consisted of a week at RAF Henlow receiving lectures on aviation medicine and undergoing practical hypoxia and G-straining training, in the Hypobaric Chamber and the Centrifuge respectively, before experiencing high G in modified Hawks at RAF Boscombe Down. Upon arrival at RAF Coningsby the subjects undertook the standard 5 week ground school, learning the technical details and systems of the aircraft. Upon completion individuals were tested using the standard ground school examination. Consequently Trial Subjects and Augmentees commenced their flying with the same base level knowledge as all pilots entering the flying phase of Typhoon training.

3.8 EQUIPMENT

3.8.1 DEVICES

The simulators used were the Aircrew Synthetic Training Aids (ASTA) delivered to the RAF in 2006. The system comprises of four devices: 2 Full-Mission Simulators (FMS) and 2 Cockpit Trainers (CT). Figure 3-4 shows an external view of the FMS demonstrating the scale required to achieve a realistic spherical projection. Figure 3-5 shows a similar view of the CT – the 230 degree projected surface can be seen, coloured sky blue behind the ironmongery, with the cockpit in the centre. An example of the cockpits used in both the FMS and CT is shown in Figure 3-6. All four devices are capable of operating independently within their own synthetic environment or alternatively being linked together in any combination and sharing a single synthetic world. All devices contain a high-fidelity cockpit, with respect to the switches contained therein, of the same design as the live aircraft with the single design exception of the Head Up Displays (HUD) of the Cockpit Trainers. Rather than a holographic HUD they contain only standard glass, the HUD display being projected onto the visual scene.

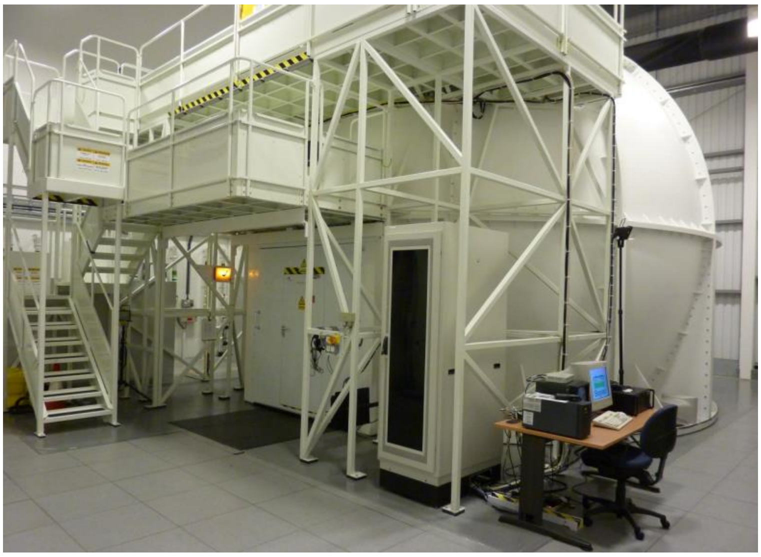

Figure 3-4. FMS Dome External View

Figure 3-5. CT External View

Figure 3-6. FMS and CT Cockpit

The visual world is taken from the same database but has 2 marked differences across the devices. The FMS visuals are projected on to a 360 degrees, 8 metre-high dome whilst the CT uses a 240 degrees dome approximately 4 metres wide. In addition, one of the CTs used higher quality visuals than the rest of the devices (although the second CT was upgraded to a matching visual standard during the period of the research). This provides a noticeable improvement in quality and clarity and allows the software engines to display cloud and other environmental issues in a more realistic manner.

3.8.2 PILOT CLOTHING

The design of the CTs provides for a lower level of pilot immersion than the FMS, with the CT pilot wearing his day-to-day flight suit and a headset (Figure 3-7). Contrast this with the FMS pilot in full summer flight clothing, Life Survival Jacket (LSJ), G-suit and helmet (Figure 3-7). The seat in the FMS is the enabler for this difference as it supports the ability to blow air into the G-suit and LSJ as well as provide breathing air for the pilot mask and a communications interface for the standard helmet. The FMS set is also capable of G-cueing, the ability to provide small movements to indicate either the feeling of aircraft rumbling along a taxiway or the onset of G. Combined with G-suit inflation the device is capable of indications only of G onset. Neither the FMS or the CT have motion capability and thus the FMS devices represents the limit of ASTA's G-cueing.

It is for these reasons that the FMS was chosen for the majority of the sorties undertaken by the students, the only exception being the first instructional sortie teaching the students to defend against hostile missiles within visual combat, known as Forward Quarter Missile Defence (FQMD). In this case the CT was used as the better visual representation allowed the pilot to see any inbound enemy missiles.

Figure 3-7. Pilot Clothing CT (left) and FMS (right).

3.8.3 SOFTWARE

The ASTA devices of RAF Coningsby use simulation rather than the emulation devices present at RAF Leuchars. Simulation utilises re-hosted aircraft software for all the computational and display tasks that are done normally within the physical airframe, whilst emulation is software written to provide a shortcut directly to the intended displayed output of the task. Emulation has been witnessed to have better stability than simulation, however its 'simplicity' means that it does not provide the pilot necessarily with what the aircraft 'does do' only what it is 'supposed to do', see Table 3-2 for a list of major differences between the two types. As a result emulation in the present Typhoon fleet provides for a lower level of fidelity than the ASTA and it is for this reason that the devices at RAF Leuchars were discounted from use within this research.

Table 3-2. Comparison Major Simulation and Emulation Differences

Item of Interest	ASTA Simulation Devices	DCT Emulated Devices
Cockpit	All switches present with look and feel of aircraft.	Number of switches missing. Touch screen devices replace multi functional displays of real aircraft.
Clothing required	Helmet radios, inflatable jacket and trousers supported.	No flying clothing supported, headset only.
Radar	Use of actual aircraft software.	Perfect emulation – no loss of tracks or environmental deterioration.
Scenario generation	Capable of complex scenarios with multiple enemy entities acting with 'intelligence'.	Highly simplistic and controlled entirely by the operator. This leads to a practical limit of 3 manoeuvring enemy entities.

Both FMS and CTs were loaded with software 3.1.X, which equates to an aircraft PSC (Production Software Configuration) of 4.3; concurrent to that used on the Typhoon OCU during the trial period. Upgraded software did become available part way through the trial and was installed at a similar time to the aircraft program. All students were trained on a software load that was the same as that flown in the live flight.

3.8.4 EXTANT LIMITATIONS

The flight and systems models used in the ASTA give rise to a number of limitations recognised prior to the commencement of the trial. Because of these limitations neither testing or training would be carried out within these disciplines:

a. **Close Formation**. When in close formation the lead aircraft jumped in each axis by an estimated 2 metres on a cycle of approximately 1Hz. Thus it was not possible to train close formation in simulation or test in live flight.

b. **Low Level**. Representation of the terrain at low level (250' and below) lacked the fidelity to train students for live flight. Owing to the Typhoon's poor forward visibility at low level from the Instructor Pilot's seat the primary researcher was not willing to risk life testing in this regime.

c. **Air-to-Air Refuelling (AAR).** The Air-to-air refuelling modelling was not complete; the hose was unable to make contact with the basket, thus there was no way of demonstrating the correct technique. Whilst software could be written to make this a functional environment it would be difficult to model the intricate, interlinked air flow patterns around the basket, tanker and receiver.

d. **Night Flying**. In addition to the ASTA limitations the real world deployment of aircraft and pilots to OPERATION ELLAMY required savings to the training schedule to be found. Thus the night flying portion was removed in toto.

3.9 METHOD

3.9.1 INTRODUCTION

A total of 8 Trial Pilots were taken into the simulator portion of the trial and underwent the training for each phase of the OCU entirely within the simulator devices. The performance of the students would be compared with that of students on the standard syllabus that used both simulation and live flight. The training techniques, syllabus taught and training aims of each sortie were the same as those for students on the standard syllabus. Similarly the Sortie Report Forms (SRF) and guidance used to ascertain scores for each of the teaching points would remain the same between the Live and Trial students. At the end of each of the phases the Trial students would be tested with a single End-of-phase flight in the live environment.

3.9.2 PREPARATION

As it was intended that all instruction received by a student pilot was to be given within the simulator there was a need to create a syllabus that provided a learning environment that covered all the situations that could be encountered on a live sortie but with a particular reference to safety. Thus, the standard live syllabus was used as a template but each mission was crafted to train each student on elements that could possibly be encountered in live flight. As this was to be the first time, in Europe, that first solo from simulation was to be done, and there was no intention to provide a safety chase aircraft as done in the US or to restrict the weather limitations any further than the standard live syllabus, the elements were closely linked to the risk register generated prior to the generation of the syllabus; see Risk Register Example at Appendix C.

In order to ensure consistency in delivery each sortie of each phase used a standard set of briefing slides that ensured each student received the same information prior to each sortie (see Appendix D). Within the phases of the course that used, as hostile opposition, Computer Generated Forces (CGFs), templates were used for each run of each sortie. Thus ensuring a uniform level of difficulty for each of the students.

As the trial involved human volunteers the primary researcher sought ethical clearance from the MoD Research Ethics Committee (MoDREC), which was granted prior to the first live flight by a trial subject. Approval was also sought to utilise 1 Group's resources; simulators, live aircraft, pilots, engineers and financial backing. This also included approval to fly students without instruction on their first ever live sortie. Each phase was separately requested and example paperwork along with ethics approval can be found in Appendix E.

The ethical discussions raised some important and difficult issues peculiar to the military context. Firstly the perception of a student being ordered to enter onto the trial rather than it being an individual's choice. This was countered by provision of a briefing sheet for all students that stated clearly that they were not being ordered. This was accompanied by a consent form containing a similar statement that they signed to demonstrate this. Secondly the perceived negative impact of a failed sortie on an individuals career. Again this was addressed in the pre trial briefing and the consent form, the students were informed that failure at some point was to be expected but that this could be equally to do with simulator limitations as ability. Re-training would be undertaken using the primary instructor and a subsequent failure would only result in them being placed on the extant flying syllabus with no detrimental effect on their career.

3.9.3 INSTRUCTIONAL STAFF

The instructors used to train and assess the students were all Qualified Typhoon Instructor Pilots (QPI) assigned to 29 Sqn Operational Conversion Unit. Each QPI was current and thus was assessing student standards daily; to this end the assessment of the trial student's performance would be set within the context of current performance levels. The majority of the simulator instruction was provided by a single instructor in order to maintain an instructional constant. This instructor was the primary researcher who had 12 years experience as an RAF instructor and 6 years in the Typhoon program and was also a qualified Typhoon flying and groundschool instructor.   Any conceptual failures in tactics or procedures, with this QPI as the source, should therefore have been resident in all students; facilitating identification. The exceptions to the use of this instructor were the live test sortie and the preceding simulator ride which ensured that the primary simulator instructor had not taught any incorrect techniques and insulated the primary researcher from influencing the transition into live flight. This was to ensure the standard was on a par with those entering this phase of the live course, such that the resources allocated to the live flight would not be wasted.

To provide training in the Combat and Counter Air phases there was a need for the student to act as a part of a formation led by a competent formation lead. In the Counter Air phase this was provided by ex-RAF Simulator Instructor Pilots (SIPs) who had been specifically retrained to provide this service. For the Combat phase, however, there was no suitably trained SIPs available, thus formation lead was provided by another Trial Subject pilot. To ensure consistency of presentations with these inexperienced formation leaders a prebrief and highly procedural directions were given for each sub task along with careful monitoring of the leader during the runs themselves.

Use of this instructional methodology allowed consistent instructional levels to be maintained over the 2 years of the trial and ensured that assessments of the Trial Subjects was always undertaken by qualified and competent staff that were training peers daily in the live environment, thus being familiar and current with the application of the assessment standards.

3.9.4 LIVE TEST SORTIE

The live flights were independently assessed by OCU IPs of B1 standard (experienced) or higher. To maintain this independence no test flights used Sqn Ldr Allsop, the primary researcher, in any airborne role. The testing pilots were briefed that the Trial Subject should be treated in the same manner as 'live' syllabus students and the course standard to be achieved was not to be compromised. All test sorties used the same profiles and content as the corresponding standard syllabus test.

3.9.5 SAFETY

Electing to send students solo directly out of simulation is a European first and thus contained considerable risk, primarily the students could have been over aroused, resulting in forgetting key checks or techniques. To ensure the live flights were conducted within a safe environment the student was provided with a 'Student's Friend', an allocated QPI with experience of the Trial. This IP accompanied the student at all stages of the live flight up to engine start, at which point they would position to the Air Traffic Control (ATC) tower to act as Duty Pilot (DP). The purpose of the position was primarily to offer answers to any last minute questions the student might have and to ensure the student had not forgotten any element of strapping in due to nerves. Once in the tower the Student's Friend communicated on an allocated radio frequency, making sure the student had:

a. 4 points into the harness Quick Release Buckle, including 2 arm restraints.

b. Oxygen and Personal Survival Pack connected.

c. Ejection seat ARMED, 2 maintenance pins stowed.

d. Completed an Emergency Brief pre take off.

On the flight itself a 'ghost' pilot was positioned in the back seat. This qualified pilot was present in the event of risk to life only and was instructed not to assist the student in any way, verbal or physical. This was ensured by switching off the intercom between cockpits and all control ability forwarded to the trial pilot's cockpit. Interaction with the student was also restricted on the ground – with no discussions permitted, even to the extent of the ghost pilot crewing in before the trial pilot had even walked for the aircraft. In this manner safety was maintained without affecting the autonomy of the trial pilot.

All emergencies during flight were to be actioned by the student initially, using the DP in the tower as required. The ghost pilot would only be used should the emergency be immediately life threatening. Articulation of emergency indications across the radio between the student and the DP was practiced in the simulated sorties. Flying instruction provided by the DP was limited to the visual circuit and was restricted within this to gross error recognition. To compensate the student was trained to identify and fault find errors themselves during the simulated sorties and provide an assessment as to the reasons for any errors. These student assessments were then compared to the clues available to the DP, to ascertain if the course of action proposed was sensible and safe. The final actions of the DP was to ensure that upon landing the student applied correct braking action to achieve the deceleration required; the Typhoon brake system has a short delay between the request from the pedal and the application of the brake. On a first solo the student would have very limited experience of this delay and incorrect/insufficient application of brakes was identified as a potential risk.

Finally, upon completion, cockpit recordings were examined to confirm that no verbal assistance had been provided from the ghost pilot and due to the flight control system of the Typhoon the rear stick is disengaged when the front pilot has control, as such no helpful guidance on the controls could be provided. These measures were intended to separate the ghost from the trial pilot to ensure that they were isolated from any assistance from within the aircraft.

3.9.6 MEASUREMENTS, METRICS AND FAILURES

All sorties were marked using the scales and intent used on the standard syllabus to ensure parity with all historic records; examples and decodes of these are given in Appendix A. There was no consideration given to the fact the Trial Subject had completed less time in the air than his peers. As stated previously this trial supports and utilises the SME-opinion metric for a number of reasons: academia has yet to agree on a set of quantitative metrics; it is the measurement method used in the majority of Air Forces in Europe and the US; and finally the SME opinion contains implicit tacit knowledge gained over years of flying that is unable, presently, to be articulated explicitly.

3.9.7 FAILURES

In the event of a test flight failure the student and QPI would be debriefed to determine if the failure originated in the conceptual, simulator representation or student ability domains. In the event of an identified issue in simulator representation a period of synthetic retraining would be undertaken before the test flight was reflown. This re-flight would contain substantially different profiles to prevent any student from 'learning' the test sortie.

3.9.8 PROCESS

After ground school all students underwent training in 4 phases conducted in a linear fashion. Each of the phases and their specific training aims are described below.

Conversion Phase. The conversion phase teaches the student to be able to start, taxi, take off and land in all weathers, source information from the cockpit displays and ensures they would be able to return the aircraft safely even if a malfunction occurred. The key elements of the phase are Instrument Rating Test (IRT), a solo flight and formation training, the standard syllabus for which is shown in Table 3-3.

Table 3-3. Standard Syllabus - Conversion Phase

Sortie	Time (h:mm)
Sim 01: Normal Ops/ Emergs / Instrument Flying (IF)	1:15
Fly 01: Normal Ops / General Handling	1:15
Fly 02: Normal Ops / IF	1:15
Fly 03: High Level Handling	1:15
Fly 04: Low Level	1:00
Fly 05: Solo Check	1:15
Sim 02: Pre IRT Practice	1:30
Fly 06: IRT	1:15
Fly 07: Solo - GH / PD (no instructor required)	1:00

It can be seen from Table 3-3 that the control subjects passing through the standard syllabus underwent 2 simulation and 7 live sorties. Comparatively the trial syllabus contained 10 simulation missions and a single live sortie. The self-help sorties, whilst always flown in the standard syllabus, were never captured, the trial syllabus formalises these events for later analysis of resources. The trial pilot commencing the solo flight was instructed in each aspect of the standard syllabus, the critical difference being that all the trial student's knowledge had demonstrably come from simulation, thus any failures or safety-critical issues witnessed in live flight would be as a result of a deficiency of the simulator environment.

Table 3-4. Trial Syllabus - Conversion Phase

Sortie	Time (h:mm)
Sim 01: Normal Operating Procedures; TO / land techniques.	1:45
Sim 02: Airframe / engine handling / Circuits in variable wind	1:45
Sim 03: Nav system / IF / Practice Diversion(PD) / Circuits	1:45
Sim 04: Self help – Circuits / nav kit (no instructor required)	1:15
Sim 05: Heavyweight Single engine / ML handling / PD	1:45
Sim 06: IRT practice	1:45
Sim 07: IRT	1:35
Sim 08: Self help – circuit and nav kit (no instructor required)	1:15
Sim 09: Emergencies	1:30
Sim 10: Solo practice	1:30
Fly 01: Ghost first solo	1:15

Combat Phase. The Combat Phase trained the student to the end-of-course standard of 1v1 high-aspect BFM combat against a similar type. Instruction in this phase used a student pilot in both the lead aircraft and the wingman position. This is not the practice in the standard live syllabus; the student would be led by an IP but the profiles to be flown were within the capability of the student pilots and the shortage of IPs available made the method a necessity. The advantage was that the student was exposed to formation leading and its considerations, as well as gaining experience of the BFM environment without the pressure of assessment. The increased training time was nevertheless captured for analysis and comparison to the standard live syllabus. As in the conversion phase the live test flight was flown with a ghost pilot in the rear seat who bore the same restrictions as the previous phase.

The Trial syllabus and allocations are provided at Table 3-5. This syllabus was foreshortened after approval from the RAF, the omitted elements being shown in Table 3-5.

The reasons for this foreshortening were twofold; firstly Dissimilar Air Combat (DACT) against other types such as the F-15 was not included as it was recognised that this would trade increased exposure in the live environment with little proof of simulator capability over the normal Typhoon v Typhoon test. Secondly the 2v1 Air Combat Training (ACT) was trained entirely in the simulator also with the aim of reducing live flight exposure; it was recognised that this element would be tested in the following Counter Air Phase and as such the exposure in the live environment would be duplicated.

Table 3-5. Combat Phase Syllabus

STANDARD Live Combat Syllabus			**TRIAL Combat Syllabus**		
Sortie	Purpose	Time	Sortie	Purpose	Time
Sim 1	Turn circle theory, wpns handling	1.15	Sim Cbt1 (2 x FMS)	Turn circle theory, wpns handling	1.15
Fly 1 (Dual)	1v1 Offensive perch	1.15	Sim Cbt 2 (2 x FMS)	1v1 Offensive perch	1.45 (inc. start and taxi)
Fly 2	1v1 Offensive high aspect	1.00	Sim Cbt 3 (2 x FMS)	1v1 Offensive high aspect	1.30 (inc. start and taxi)
Fly 3 (Dual)	1v1 Defensive perch	1.00	Sim Cbt 4 (2 x FMS)	1v1 Defensive perch	1.30
Fly 4	1v1 Defensive high aspect BFM	1.00	Sim Cbt 5 (2 x FMS)	1v1 Defensive high aspect BFM	1.30
Fly 5	1v1 Neutral combat	1.00	Sim Cbt 6 (2 x FMS)	1v1 Neutral combat	1.30
Fly 6	1v1 Neutral combat	1.00	**Cbt Fly 1 (Ghosted)**	1v1 Neutral combat	1.00

Table 3-6. Omitted Combat Syllabus Sorties approved by Command

STANDARD Live Combat Syllabus			**TRIAL Combat Syllabus**		
Sortie	Purpose	Time	Sortie	Purpose	Time
Fly 7	1v1 DACT	1.00	Sim Cbt 7 (2 x FMS, 1 CT)	2v1 ACT	1.15
Combat SIM 2	2v1 ACT	1.15	**Cbt Fly 2 (Ghosted)**	2v1 ACT	1.00
Fly 8 (Dual)	2v1 ACT	1.00	**Cbt Fly 3 (Ghosted)**	1v1 DACT	1.00

Counter Air Phase. The Counter Air Phase taught long range Air Defence through to ACT, the end of course standard being 2 v 2 multigroup (multi GP) against a given threat aircraft and missiles. The syllabus can be seen in Table 3-7 and was flown as approved by Command. As stated above the enemy red-air profiles flown against both sets of students were the same and as in the previous live flights the trial students were ghosted with an experienced pilot in the rear seat. It can be seen from Table 3-7 that the simulator syllabus followed the standard as closely as possible.

Table 3-7. Counter Air Phase Syllabus

Standard Live Counter Air Syllabus			Trial Counter Air Syllabus		
Sortie	**Purpose**	**Length**	**Sortie**	**Purpose**	**Length**
Course Sim	Fundamentals Geometry	1.15	Course Sim	Fundamentals Geometry	1.15
Sim 01A	1v1 ID Stern Geom	1.15	Sim 01A	1v1 ID Stern Geom	1.15
Sim 01B	1v1 QRA	1.15	Sim 1B	1v1 QRA	1.15
Fly 1	1v1 ID Stern	1.15	Sim 2	1v1 ID Stern	1.15
QRA Sim 02	2v1 ship QRA Ops	1.15	QRA Sim 3	2v1 ship QRA Ops	1.15
QRA Fly2	2v1 QRA ID and Inter	1.15	QRA Sim 4	2v1 QRA ID and Inter	1.15
Sim 3	1v1 Skate Banzai	1.15	Sim 5	1v1 Skate Banzai	1.15
Fly 3	1v1 Skate Banzai	1.15	Sim 6	1v1 Skate Banzai	1.15
Sim 4A	2v1 Skate	1.15	Sim 7	2v1 Skate	1.15
Fly 04	2v1 Skate	1.15	Sim 8	2v1 Skate	1.15
Sim 4B	2v2 Skate	1.15	Sim 9	2v2 Skate	1.15
Sim 5	2v2 Banzai	1.15	Sim 10	2v2 Banzai	1.15
Fly 5	2v2 Banzai	1.15	Sim 11	2v2 Banzai	1.15
Fly 6	2v2 Skate Banzai	1.15	Sim 12	2v2 Skate Banzai	1.15
Sim 7	PH 4 1v1 2v1	1.15	Sim 13	PH 4 1v1 2v1	1.15
Fly 7	1v1 PH 4 2v1	1.15	Sim 14	1v1 PH 4 2v1	1.15
Sim 8	HFF FQMD	1.15	Sim 15	HFF FQMD	1.15
Fly 8	FQMD, ACT	1.15	Sim 16	FQMD, ACT	1.15
Sim 11A	2vX Multi Gp	1.15	Sim 17	2vX Multi Gp	1.15
Sim 11B	2vX Multi GP Hostile Bogey Mix	1.15	Sim 18	2vX Multi GP Hostile Bogey Mix	1.15
Fly 11	2v2 Multi Gp	1.15	**Fly 11**	2v2 Multi Gp	1.15

Operational Flight Training (OPFLY) Phase. The OPFLY phase had been traditionally used to provide the student with semi scripted but more difficult enemy presentations. Post completion of the phase by the second trial student however, Command ordered that elements of this phase of the standard live course be cut in order to increase the throughput of student numbers. To this end a corresponding cut, highlighted, was made in the trial syllabus (Table 3-8). The remaining trial students thus only flew one live sortie - QRA LOW SLOW. This sortie entails high levels of manoeuvring against a light aircraft, whilst manipulating a radar and flying in formation through cloud down to heights of 500'. As such it was considered the hardest sortie on the course, nevertheless the ghost pilot received the same set of restrictions as all previous phases. Thus as both standard and Trial syllabuses contain the same sorties the only two differences were the lack of instructor prompts from the rear seat and the student's lack of live flying exposure.

Table 3-8. OPFLY Phase Syllabus

Standard Live Combat Syllabus			Trial Combat Syllabus		
Sortie	**Purpose**	**Length**	**Sortie**	**Purpose**	**Length**
OP Sim 01	Low/slow QRA Intercepts	1.15	OP Sim 01	Low/slow QRA Intercepts	1.15
OP Fly 01	Show proficiency in QRA LOW SLOW	1.15	**OP Fly 01**	Show proficiency in QRA LOW SLOW	1.15
OP Fly 02	High Risk Point Def	1.15	OP Sim 02	High Risk Point Def	1.15
OP Fly 03	1Multi Gp – Pre EoCC	1.15	OP Sim 03	1Multi Gp – Pre EoCC	1.15
OP Fly 04	EOCC	1.15	OP Fly 02	EOCC	1.15

3.10 RESULTS AND DISCUSSION

3.10.1 CONVERSION PHASE

A total of 13 students (8 trial subjects and 5 augmentees) completed the Conversion Phase in the simulator and subsequently undertook their first flight as a ghosted solo. No IP reported having to switch their microphone on or take control of the aircraft. A further 3 of the most experienced students (all ex Harrier pilots) completed their flight as a true live solo i.e. not ghosted. The tapes of these flights were reviewed and showed a high degree of adherence to the technique and no safety points. All flights contained 2 practice diversions to RAF Marham and RAF Wittering, General Handling (GH), area familiarisation and a recovery for circuits at RAF Coningsby. Following the closure of RAF Wittering sorties included a supersonic run in the North Sea in order to burn down the fuel prior to the RAF Marham diversion.

During these 16 live events a number of real-time issues were encountered, these were normal occurrences of problems that were seen on the fleet's aircraft at the time. There was a single instance of a double CSG (Computer Symbol Generator) failure that restored, an SPS (Secondary Power System) Computer failure and loss of SEP (Specific Excess Power) bars. All of these issues were dealt with using the Trial Subject and the DP using the methods taught in the simulator. There was no assistance from any IPs. Of particular note, however, was the instance of unforeseen poor weather; a 1000' cloudbase forecast to improve became 350' whilst 2 trial subjects were airborne. It was elected to continue making approaches until the base raised to 400', the student's legal minima, at which point they landed safely, their first solo sorties being subsequently reflown. These incidents and problems demonstrated that the trials pilots were able to deal with real emergencies without on-board assistance using knowledge gained solely within the simulator and in doing so further informed the risk register.

3.10.2 TESTING THE CONVERSION PHASE

The lack of IP input and the ability of the trial subjects to deal with emergencies showed that solo flight direct from simulation was indeed possible and de-risked the use of the methodology should it be taken into the core syllabus. Clearly, however, the nature of a solo flight prevents assessment of a student pilot's abilities other than satisfying the macro items of safe takeoff, landing and in-flight navigation. Whilst it was not able to provide a useful comparison of performance between control and trial groupings, the Instrument Rating Test (IRT), conducted 2 sorties before the solo flight, did provide a noteworthy comparison between training and testing methodologies that asked questions of the quasi-transfer tests referred to in the literature review. The quasi-transfer tests referred to in much of the meta-analysis reviewed used SME opinion to provide an overall mark for individual trainees. This assumed that the mark an SME gave in the synthetic environment was directly comparable to that given in the live. More specifically that key events within the subtasks, such as a safety or procedural violation, had the effect of reducing the overall mark by the same amount in the live and the synthetic environments. Trial PANDORA'S BUZZARD sought to compare the IRT results of those trained and tested in the live environment with those in the synthetic in order to search for any abnormalities in the scoring between the two environments. This was the first test of its type as the reason for quasi-transfer tests had been the high cost of live training comparison in the first place; Trial PANDORA'S BUZZARD therefore offered the first known ability to search for this affect. Significant differences in the grading effects would make the admission of Quasi-transfer Trial evidence referred to at section 2.5.2 difficult.

Thus the results of 23 subjects and augmentees that conducted their IRT in the simulator, following the Trial PANDORA'S BUZZARD (PB) syllabus were compared to all available historic records (n=53) of pilots that conducted training using the standard syllabus and undertaking their IRT airborne. The tasks set in the simulator were exactly the same as those set in the tests completed in live flight and were flown to the same standards as laid down in Group Air Staff Orders (GASOs).

3.10.3 CONVERSION PHASE RESULTS

The population was initially investigated to determine if there were any particular groups that might skew the results. As the skill set tested was common to all pilots regardless of background the only remaining possible factor was that of rank. Just as in section 3.5 the population data was filtered to determine if rank, an indicator of experience, had an affect on overall scores, although in this section only the IRT scores were considered rather than the sum of the marks for the 3 phases. Flight Lieutenants (F) were compared with more experienced ranks; Squadron Leaders and above (E – Experienced). A subsequent 2 sample t test demonstrated a difference of 0.46 to a significance of P=0.02 (see Figure 3-8). Thus, the two groups were considered separately for the remainder of the conversion phase and, as the focus of the study is those pilots of Flight Lieutenant rank it is this sub group that will be investigated primarily, with the experienced group commented on by exception.

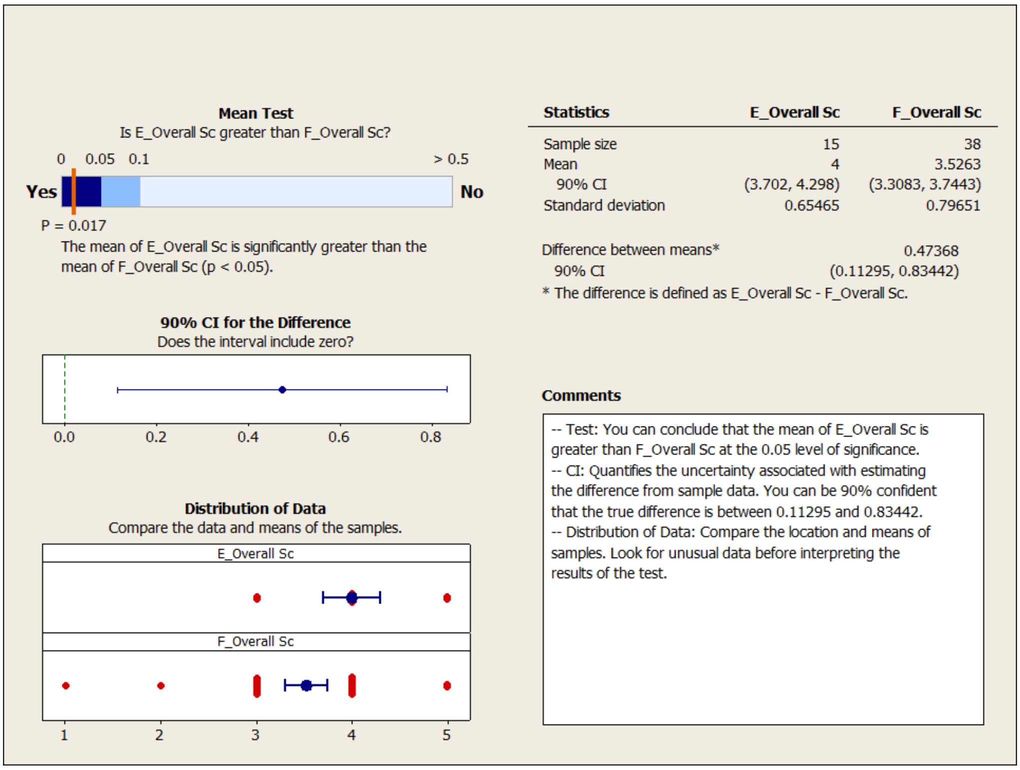

Figure 3-8. 2 Sample-t Results: Experienced (E) v Flight Lieutenant (F) with respect to IRT Overall Score on the Standard Syllabus

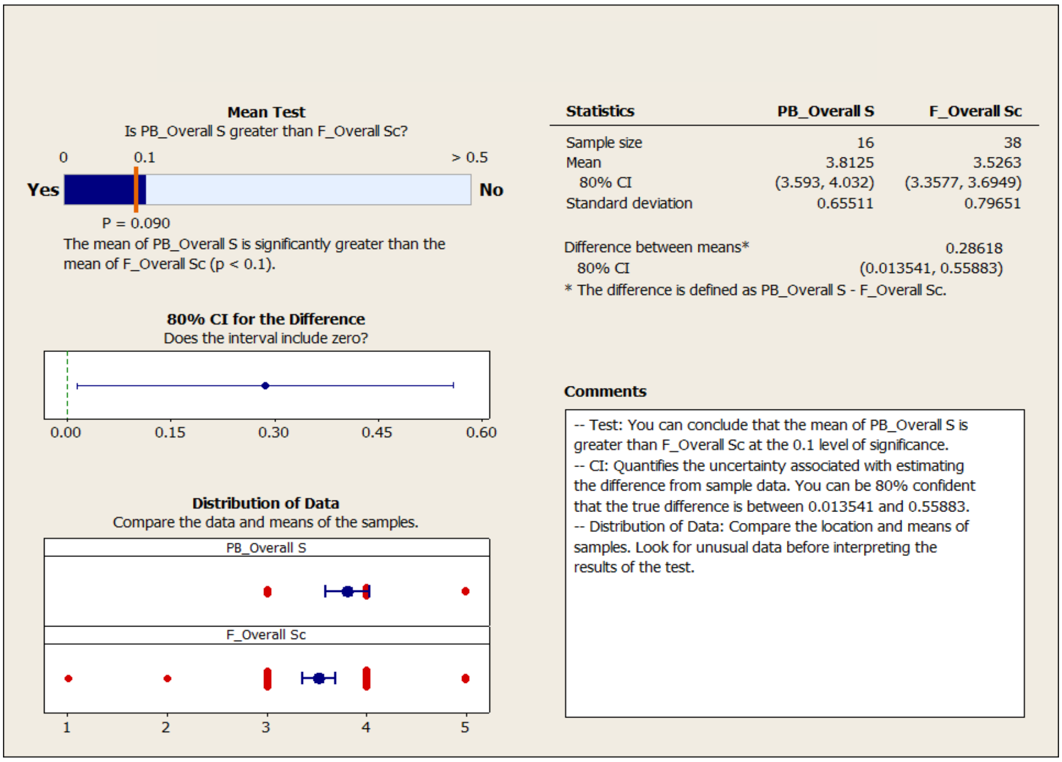

Figure 3-9. 2 Sample-t: Comparison of Means 2: IRT

Figure 3-9 provides a comparison of the means of the Flight Lieutenants tested using the Trial PB and the Standard syllabus. It shows the Trial PB syllabus have a significantly higher (P=0.09) mean than the Standard Syllabus. This appeared to indicate that training and testing in the simulator therefore produced a better performance than training and testing in live flight: supporting the same observation from the Quasi-Transfer trial results discussed in the literature review. This result, however, could have been a function of the simulator providing a better training environment or that the SME's subjective metric baseline altered whilst within the simulator environment.

In order to investigate this question the subcategories of SME scoring and their relationship with the overall mark given were examined. Figure 3-10 shows a broad examination of each of the subtasks relationship with the overall mark for both the simulator and the live (standard) syllabus. The difference between the gradients of the 'sim' (trial) and 'live' (standard) demonstrates the difference in the relationship to the Overall Score ie. 'Is a score of 3 in Mental Performance likely to return the same overall score, whether the sortie was flown live or in the simulator?' Of note from Figure 3-10 is the sub-group of 'Safety', which indicates noticeable gradient differences. These categories were then regressed to determine their relationship to the overall mark.

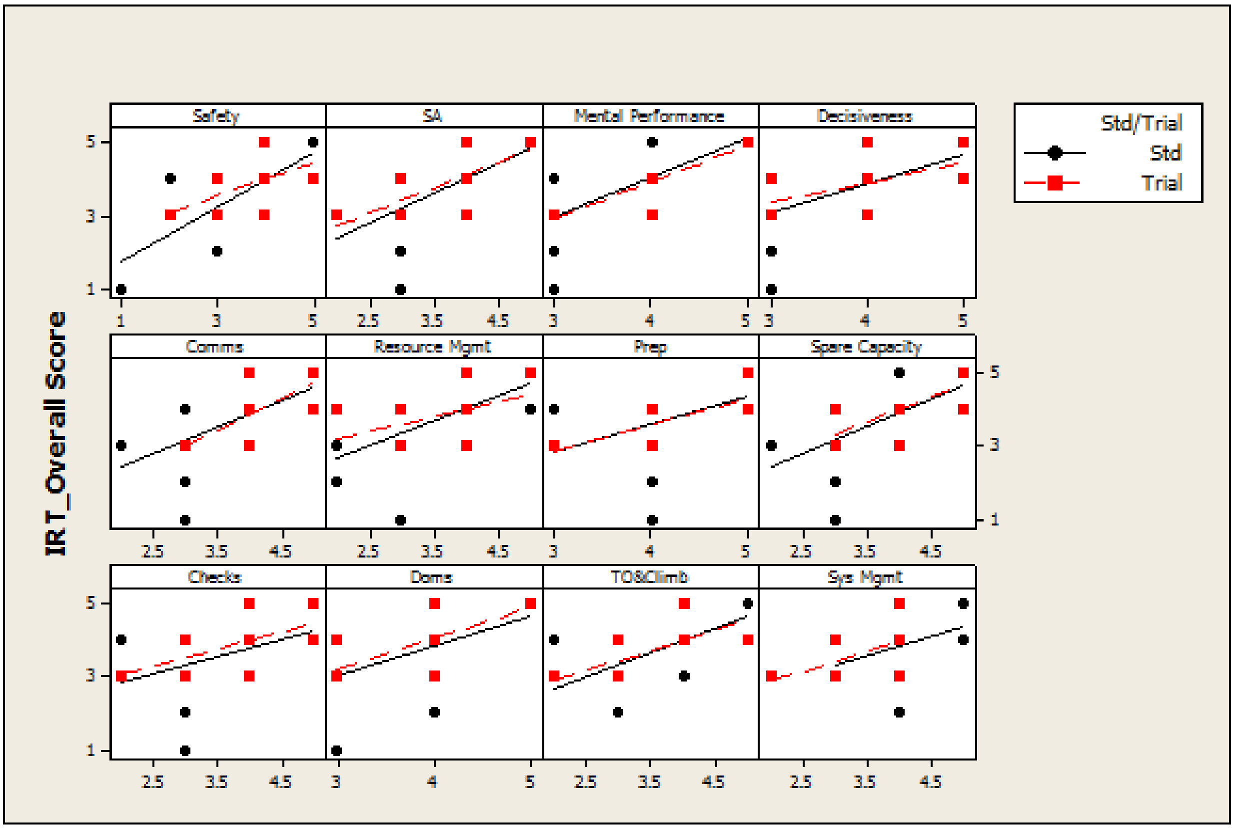

Figure 3-10. Comparison of Sub Category Relationships vs Overall Score, IRT

The results for Safety returned a linear regression of:

- IRT Standard Syllabus Overall Mark = 1.01 + (0.7317 x Safety Score)
- IRT PB Trial Syllabus Overall Mark = 2.215 + (0.4359 x Safety Score)

All results show statistically-relevant results where $p<0.1$. The percentage of the results explained by the model are:

- IRT Std Syllabus v Safety, R-sq (adj) 57.0%
- IRT PB Trial Syllabus v Safety, R-sq (adj) 17.3%

These figures indicate that when considering 'Safety' a score of 2 for the subcategory is likely to return a lower mark for the overall score when the IRT is conducted within live flight versus that conducted in the simulator. Whether this is to do with the physical separation of the IP and student pilot in the simulator compared to the shared exposure experienced within live flight it is not clear.

The result should also be seen within the context of the low R-sq (adj) figure for the PB syllabus, nevertheless it is right that this apparent discrepancy is highlighted in terms of possible risk should a fully synthetic syllabus be considered.

The number of grades that do not meet the expected standard for the individual skill being tested; graded 2 or less can be labelled in statistical terms ‘defects’. Thus a comparison of defects per student was made between live flight (‘before’) and simulation (‘after’). The results showed a small rise from 3.37% to 3.77% with the 95% CI for the after group showing a small increase in the upper limit to 7.03% versus the previous 5.10% (see Figure 3-11).

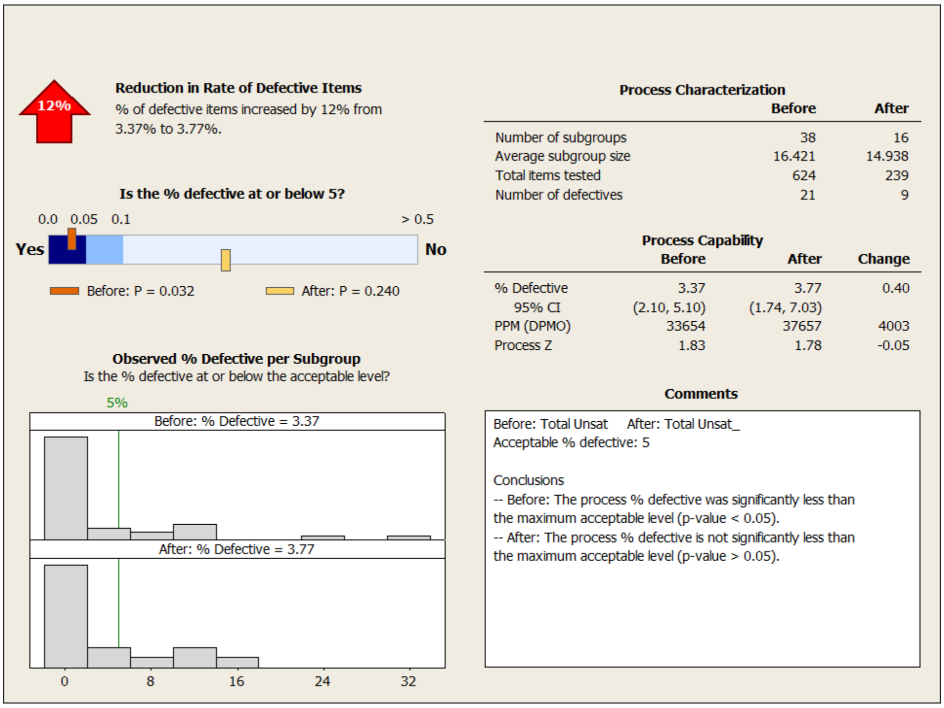

Figure 3-11. Effect on 'Defects' when IRT is Tested within the Simulator

3.10.4 <u>CONVERSION PHASE SUMMARY</u>

The first aim for Phase 1 – Conversion, laid out in Section 3.4, was to determine if the results of the IRT were affected by a fully-synthetic training package and test versus the standard syllabus and the test in live flight. The results above have shown that the trial group performed at least as well as the control. It should be caveated, however, that whilst this would demonstrate simulation to be a suitable location to test in, the RAF should be aware of the differing 'leverage' of safety issues between live and synthetic testing. i.e. A safety critical error in the simulator would appear to effect the overall mark less than if the testing environment was in live flight.

The second aim was to determine the feasibility of achieving a successful first solo from a fully synthetic syllabus. Conversion Phase results have shown that it *is* possible to send student pilots solo directly out of simulation, with all 16 pilots undertaking the flight successfully and completing their mission without intervention, despite some difficult technical and weather-related issues. Although the sample size is recognised to be relatively small it is large enough to recommend to the RAF a continuation of the methodology to further increase the sample size before inclusion into the core syllabus.

Finally within the first theme of the dissected research question Objective 5 of Sub-Question 1 required the generation of a risk register to determine the risks during the first solo flight. This was completed and is presented in Appendix C.

3.11 COMBAT PHASE

Within the Conversion Phase (IRT) only the Flight Lieutenant ranks were considered, with the higher ranks filtered out, in order to provide a common experience level. This was permitted as instrument flying is a common skill across all platforms, barring idiosyncrasies of each particular aircraft type, this is due to the much slower speeds, non-dynamic manoeuvring and rigid adherence to external, laid down profiles. Thus, experienced pilots were more likely to have an advantage over the inexperienced. In the Combat phase, however, no advantage exists as the Typhoon produces a performance that exceeds all previous RAF aircraft by significant margins. The Typhoon is capable of 9g, has a thrust-to-weight ratio exceeding 1:1 and carefree computer-controlled handling, in comparison all previous RAF aircraft were capable of no more than

5g, had thrust-to-weight ratios of less than 0.5 and pilot-controlled flight envelopes. To this end no pilot could be classed as 'experienced' in the Typhoon's combat environment and as such all pilots regardless of rank were considered for the statistical analysis.

The standard syllabus (live training) results, i.e. the control group were taken from contemporary records between late 2010 and 2012 as before and after these dates the Combat syllabus underwent significant changes with respect to tactics and techniques, thus preventing a true comparison to be undertaken. The sample sizes reflect these restrictions with only 22 live (control) and 8 trial students passing through the course during these periods; using the syllabus provided in Section 3.9.8.

3.11.1 <u>COMBAT RESULTS</u>

Figure 3-12 shows that the mean of the Trial group was a significant 0.47 less than that of the students trained under the standard syllabus. Primarily this was due to a failure of a single trial pilot (who scored an overall mark of 2). Thus the combat phase fails both tests 2 and 3.

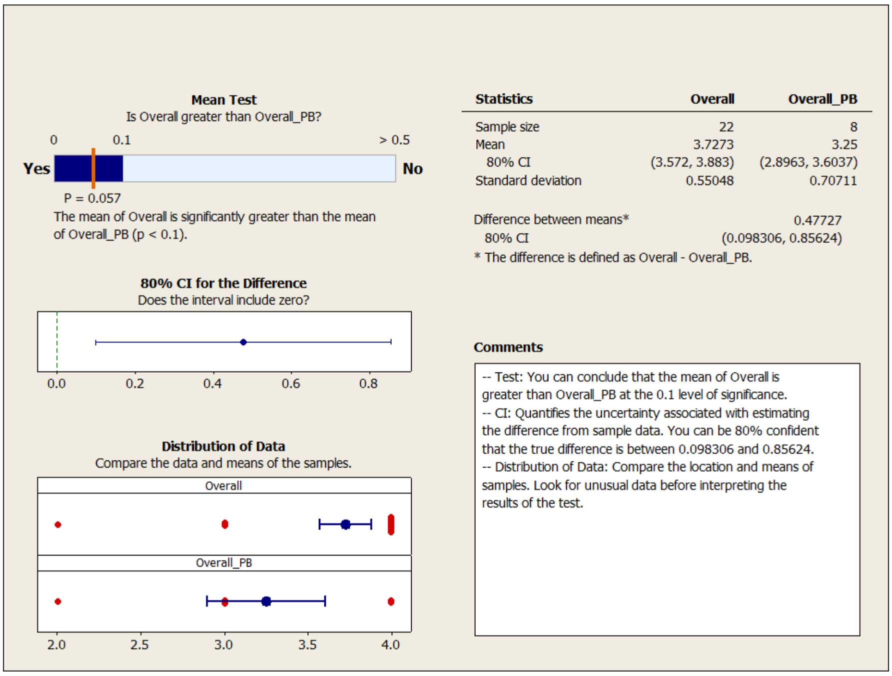

Figure 3-12. 2 Sample t Test, Standard v Trial Syllabus.

The overall results show that transfer of training from simulator to live flight was not capable of producing a comparable student to that of the standard ‘live’ methodology, with the synthetically-trained students producing a mean that was lower than that of the live-flight trained. The underlying reasons for this poor transfer were not immediately clear. In order to determine influencing factors the sortie reports were broken down into skillsets and examined against the corresponding standard syllabus areas. Table 3-9 details the individual skills within each skillset and the reason why that skillset is important. The mean of the skillset was calculated from all available individual skills within the set. Figure 3-13 shows the comparison of the standard and trial syllabus for each skillset.

Table 3-9. Skillsets and their Relevance.

SKILLSET	INDIVIDUAL SKILLS	REMARKS
Domestics (Doms)	• Preparation (Prep) • Startup, Taxi, Take Off (SUTTO) • Tactical Domestics (Tac Doms) • Weapon System Checks (WSC) • Border Crossing Checks (Fence) • G Warm up	Demonstrates the ability to safely conduct training and fly the aircraft to and from the training area.
Guns Exercise (GunEx)	• Resolve Plane of Motion (POM) • Resolve Range • Resolve Lead • Energy Management (NRG) • Effective Guns Techniques	This skillset is concerned with close in weaponeering (within 3000ft) and is characterised by highly dynamic manoeuvring at close range.
Within Visual Range (WVR)	• Plan Execution • Advantage Recognition • Turn Circle Recognition • Lag BFM Technique • Kill or BFM decision making • Maintain the Offensive • Defensive BFM • Infa Red Decoy Dispensing	Positioning of the aircraft to achieve a missile kill. Ranges from 2 miles to 3000ft. Influenced by the highest g forces and students understanding and employment of counter manoeuvres to enemy positioning.
Max Performance Handling (MPH)	• Lead Turn Recognition • Execution of the Break Turn • Lift Vector Placement • Energy Management	Shows the students ability to fly the aircraft to achieve the max performance from the airframe.
Weapon System Handling (WSH)	• Weapon Engagement Zone Recognition (WEZ) • Hostile WEZ Recognition • Weapon Tree Execution • Weapon Employment • Gun Combined Error Technique (CET) • Validity of weapon solutions	Demonstrates student awareness of the envelope of both their and the enemy weapons. Shows the correct technique in the employment of those weapons.
Resource Management	• Cockpit Resource Management • Fuel Management • Task Management • Situational Awareness (SA) • Mid Air Collision Avoidance • Adherence to Air Training Instructions (ATIs)	Shows the awareness of and demonstrates the ability to employ all the rules and regulations associated with combat training.

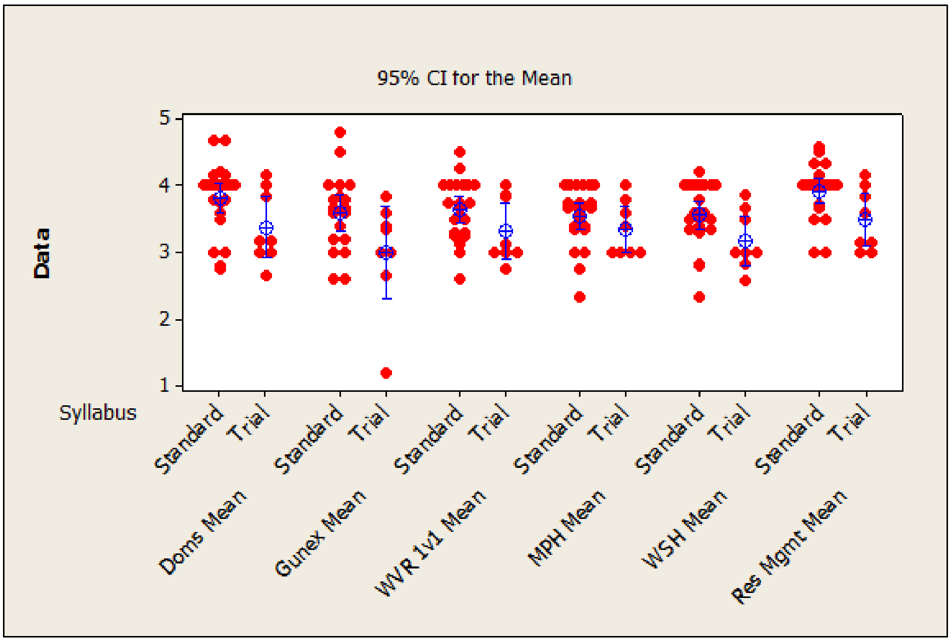

Figure 3-13. Overview of Combat Skillset Means v Syllabus Type

Of particular interest were the major skillsets that indicated a substantial difference in mean between the 2 syllabi. From a high-level viewpoint Figure 3-13 shows that there were two areas in which this may occur: ‘Gun Exercises’ and ‘Weapon System Handling’. However, in order to understand these differences at a more meaningful level and understand which skills did not transfer into live flight Table 3-10 provides comparison of the skills within the sets. This compares only those skills within the core of the syllabus and discounts any skills that had 6 or fewer trial syllabus results. Given the serious risks associated with recommending training within the simulator over that proven to be successful in live flight (and the large monetary figures concerned) the questions and settings favoured live flight. Thus, should a particular skill be recommended for practice within the simulator it can be done with as much reduction in risk as possible, given the small sample sizes. To this end the question asked was “Is the mean of the skill of the standard syllabus greater than the mean of skill of the trial syllabus?” rather than ‘equal’. With the same reasoning the level of Significance was set at 0.1 rather than 0.05, demonstrating an increase acceptance of risk in making the 'greater than' conclusion when it was not true.

Table 3-10 is highlighted in red for those skills where the data show the standard syllabus to have a higher mean than the trial. For those results that did not return a P < 0.1 the power of the sample is provided to allow an understanding of the data's strength; the difference of interest being ¾ of a single mark. This level was set as a balance between the reality of the small sample size and the need to find a difference that would mark a clear deficiency in the transfer of training. 0.75 is, therefore, a figure that would lead to a practically-recognisable lower standard of student.

Table 3-10. Key Combat Skills Comparison Data

SKILLSET	TEST: IS MEAN OF SKILL (STD) GREATER THAN SKILL (TRIAL)?	N STD	N TRIAL	MEAN STD (80% CI) [SD]	MEAN TRIAL (80% CI) [SD]	SIGNIFICANCE	POWER % FOR DIFF OF INTEREST 0.75
Guns Exercise	Resolve POM	19	7	3.79 (3.57,4.01) [0.71]	3.00 (2.37,3.63) [1.155]	0.07	59.5
	Resolve Range	19	7	3.47 (3.20,3.75) [0.90]	2.71 (1.96,3.47) [1.38]	0.01	49.4
	Resolve Lead	19	7	3.47 (3.29,3.66) [0.61]	3.14 (2.77,3.52) [0.69]	0.15	87.3
Within Visual Range (WVR)	Advantage Recognition	19	7	3.54 (3.34,3.73) [0.52]	3.57 (3.28,3.86) [0.53]	0.55	94.9
	Turn Circle Recognition	22	7	3.82 (3.71,3.93) [0.39]	3.29 (3.02,3.55) [0.49]	0.02	98.7
	Lag BFM	20	8	3.65 (3.51,3.80) [0.49]	3.38 (3.12,3.63) [0.52]	0.11	98.3
	Kill or BFM	22	8	3.45 (3.27,3.64) [0.67]	3.25 (2.90,3.60) [0.71]	0.21	89.1
	Maintain the Offensive	22	8	3.68 (3.50,3.87) [0.65]	3.38 (3.12,3.63) [0.52]	0.10	97.1
	Defensive	8	8	3.25	3.13	0.33	90.5

	BFM			(3.02,3.48) [0.46]	(2.80,3.45) [0.64]		
Max Performance Handling	Handling	8	7	3.63 (3.37,3.89) [0.52]	3.57 (3.28,3.86) [0.53]	0.42	91.7
	Lead Turn Recognition	9	8	3.56 (3.31,3.80) [0.53]	3.25 (3.02,3.48) [0.46]	0.11	96.0
	Break Turn Execution	21	8	3.2 (3.33,3.72) [0.68]	3.5 (3.23,3.77) [0.53]	0.46	96.1
	Lift Vector Placement	21	8	3.52 (3.38,3.67) [0.51]	3.38 (3.12,3.63) [0.52]	0.25	98.2
Weapon System Handling	Energy Management – BFM	22	8	3.55 (3.38,3.71) [0.60]	3.13 (2.80,3.45) [0.64]	0.07	93.3
	WEZ Recognition	20	8	3.70 (3.56,3.84) [0.47]	3.00 (2.73,3.27) [0.53]	0.01	98.0
	Hostile WEZ Recognition	12	8	3.42 (3.21,3.61) [0.51]	3.38 (3.12,3.63) [0.52]	0.43	96.5
	Weapon Tree Execution	21	8	3.48 (3.30,3.65) [0.60]	3.38 (3.12,3.63) [0.52]	0.33	97.5
	Weapon Employment	19	8	3.32 (3.05,3.59) [0.89]	3.00 (2.62,3.38) [0.76]	0.18	81.4
	CET Gun Technique	14	8	3.57 (3.39,3.76) [0.51]	3.00 (2.62,3.38) [0.76]	0.04	86.8

3.11.2 COMBAT ANALYSIS

Table 3-10 shows that, when tested in live flight, the skills of Turn Circle Recognition, Maintain the Offensive, Energy Management and WEZ (Weapons Engagement Zone) Recognition all returned a better performance when using the standard syllabus than the trial. Additionally, despite having a sub-optimal power, the skills of Resolve POM (Plane of Motion), Resolve Range and CET(Combined Error Technique) Gun Technique also returned a similar observation.

These observations are notable as, with a single exception – energy management, they all contain a single root requirement - the ability to distinguish immediately a change in the enemy's roll or pitch angle and action an immediate counter move. Thus, the inability of the trial pilots to transfer their training into live flight was hypothesised to have two possible causes within the synthetic training. The first was that the pilot was applying the wrong counter to the enemy's action; this was discounted as the sortie report forms from both the primary instructor and the instructors conducting the pre-live flight simulator mission displayed no evidence of this, indeed the students' conceptual understanding of BFM had appeared to be a strength. The second possible cause was the student delaying the counter action; a delay of 2 seconds could equate to the enemy aircraft gaining well over 30 degrees on the student – a position difficult to recover from. Successful recognition of a change in the enemy's state needed to be recognised and actioned in about 1 second to ensure that parity in the fight was retained. Further examination of Table 3-10 showed that tasks requiring the student to extrapolate a line of flight, such as Resolving Lead when employing guns or Advantage Recognition, were conducted satisfactorily. Thus, it was concluded that the student had been unable to detect the immediate roll or pitch change in the simulator and was applying the counter when the change became apparent through an established change of direction or sight line. Figure 3-14 provides context through the provision of HUD footage taken from live flight. Rounds from the gun will always be fired down the line of the gun Fixed Aim Cross, however the actual flight path of the bullets will be along the bullet fall line as a result of the aircraft's flight path and gravity. The fall line shown in the Figure is being continuously updated, resulting in a dynamic problem for the pilot. In order to successfully kill the enemy the fall line must be held over the enemy at exactly the range of the enemy fighter. If the enemy fighter moves the new flight path must be assessed, the vector of the friendly fighter altered and the newly computed fall line placed over the enemy at the correct range. The first indication that the enemy pilot is changing his flight vector is that the angle of bank of the target aircraft will change, and in doing so the lift vector will be reorientated resulting in the change of flight path. Figure 3-14 shows just how difficult spotting the indicators change in bank will be. In reality this change is reduced to a simpler variable – a perceived change in the distance of the two wingtips. If that distance increases the target is rolling towards the friendly, if it reduces it is rolling away. What the research has shown is that all the skills that require this variable to be recognised quickly are poor. Thus the inference is that this variable has difficulty being recognised in a timely manner.

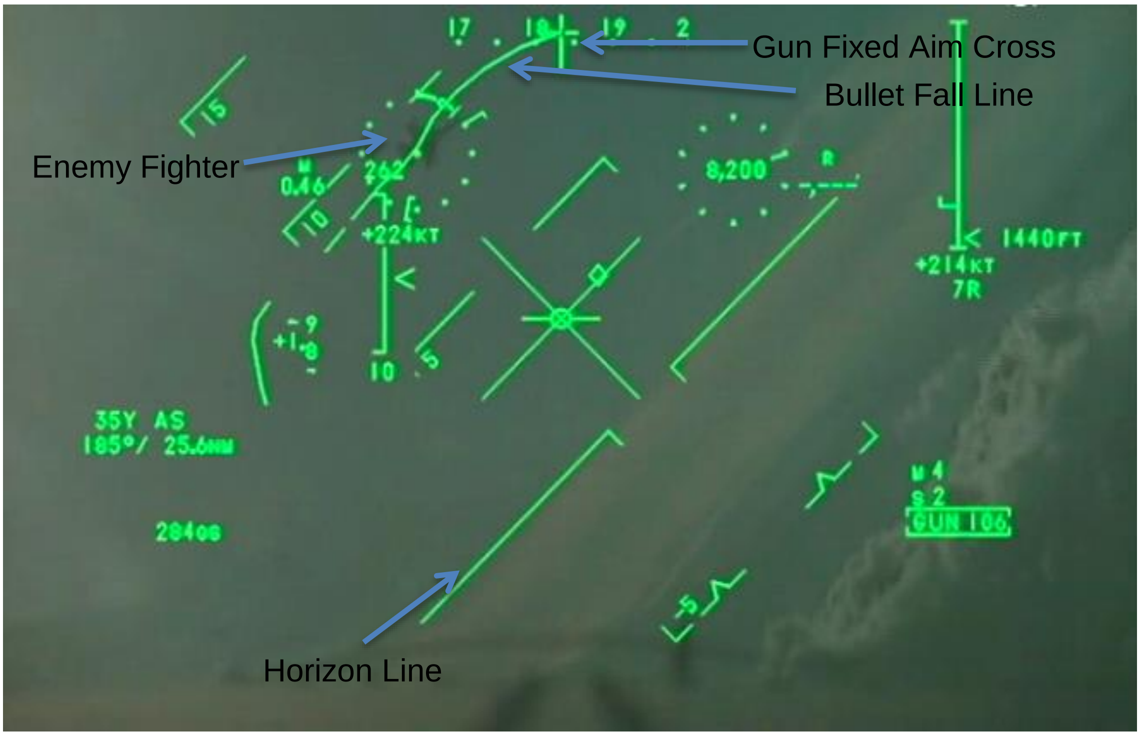

Figure 3-14. Example Guns Footage from Head Up Display. Unclassified Source.

The issue was negated within the simulator as the other pilot was faced with the same disadvantages, the 'motor program' had then been taken airborne into the test flight resulting in the delayed reaction. The inability to recognise the change in pitch or roll of the simulator is due to the fidelity of the visual scene within the simulator. Despite using state-of-the-art projection, aircraft at all ranges were too dim to be able to pick out these variables. The inability to increase the contrast of nearby aircraft had resulted in an addition of 6 Target Projectors that projected an image of the nearest aircraft no matter its position to the host aircraft. With only a single task these projectors boosted the contrast to what had been believed to be sufficient levels and had been used to their full extent in the trial. The results of the trial indicated that these had only been partially successful in resolving the issue and thus were listed as a limitation to the transfer of training into live flight within the regime of Combat.

The skill that differed in its root cause was that of Energy Management. This concerns the physical handling and 'feel' of the aircraft. The inability of this skill to transfer to live flight from simulation would infer that the performance handling of the simulator differed from that of the live aircraft. This could be a function of the lack of a number of factors: the physiological component, airframe buffet, incorrect audio cues or an incorrect performance model within the simulator itself. Operation close to the edges of the performance envelope would combine all of these factors and thus isolation of a single factor was not possible either in experimental design or in the examination of the results.

3.11.3 COMBAT PHASE SUMMARY

The aims set out in section 3.4 require a recommendation as to the inclusion of the 'Phase 2 – Combat' trial syllabus for inclusion into the core syllabus, with particular reference to the likelihood of a fully synthetically trained student passing the EoP check ride compared to a student on the standard syllabus. Additionally the trial was asked to identify any limiting factors preventing further exploitation of simulation in the BFM environment.

Whilst the probability of a student passing the trial course was unable to be determined categorically with the small sample size, the failure of the Phase to pass Tests 2 and 3 (3.6 3 Tests to Measure Success) provided sufficient evidence *not* to recommend this element of the trial syllabus. This was also in line with Bock et al's (2002) use of small samples to make recommendations (see section 2.5.3). This recommendation was accompanied by two key limitations, firstly that the target projectors needed to be switched on for combat training but their lack of fidelity when reproducing the movements of enemy aircraft increased the likelihood of the student learning to identify incorrect visual cues for their decision making. Secondly the lack of ‘G’ within the simulated environment was unquantifiable but the regular use of 9G per second onset rates and sustained 9G fighting increased the likelihood of a black out; this increased risk to life would be unable to be mitigated within a true solo flight, thus further supporting the recommendation not to include Phase 2 within the core syllabus.

Despite this negative recommendation a number of positives were to be taken from this phase. Although skill of weaponeering with guns did not transfer to live flight it was taught effectively within the simulator. Thus all dogfighting concepts can be taught synthetically with the latest simulation capability, the transfer into live flight however needs restructuring to train the guns skills specifically. This would allow increased proportion of synthetics within the syllabus. Additionally, away from the guns skill, the ability of the simulator to train pilots in the other skills was not in question, thus the simulator offers a valid location to train students who have failed or are struggling to grasp the concepts of BFM in the Typhoon.

3.12 COUNTER AIR PHASE

The Counter Air Phase measured the performances of the 8 trial subjects in the Beyond Visual Range (BVR) through to the closer-range BFM.   Initially the phase taught radar handling and 1v1 operations, moving through differing forms of intercepts and culminating in a pair of Typhoons versus 2 enemy aircraft.

The Trial syllabus (Section 3.9.8) is a copy of the standard syllabus with live flights replaced by a flight in the simulator covering the same aims and objectives as in the air. Enemy game plans, dictating their heights, speeds, formation and shot ranges, were taken from the standard syllabus ‘Red Air’ profiles so that Trial Subjects saw the same enemy presentations as their live counterparts.

It was observed that in live flights of the standard syllabus only 3 runs per sortie would be achieved before the students would be out of fuel. It was found that a corresponding sortie in the simulator would achieve 6-7 runs before recovery owing to the ability of the simulator freeze and reset to a particular start point rather than flying back. This provided increased exposure to employment of the techniques and allowed the student to demonstrate some consolidation.

The standard syllabus introduced enemy radar emissions from the 10th sortie of the syllabus onwards. These radar emissions were produced not by other Typhoon aircraft, as there were simply not sufficient numbers, but by an external contractor flying a Dassault DA20 aircraft. Use of these aircraft consumed a fixed budget and thus for the majority of the live training flights the RAF preferred to use the cheaper and more numerous Hawk aircraft which had two distinct disadvantages: they were not fitted with radars and thus were unable to stimulate the student aircrafts' Defensive Aids Sub-Suite and secondly the flight envelope was significantly inferior to all except the very oldest threat aircraft. Uninhibited by these constraints the simulation used in the trial syllabus was able to utilise representative threats from the outset. Their early introduction allowed students to become accustomed to the sounds and indications of the threats over a number of sorties, without having to react to them until sortie 10 onwards. Consequently debriefs were able to discuss considerations pertaining to these threats much earlier in the syllabus.

It was recognised that the Trial Subjects would have to perform in live flight within the same weather limitations as their standard syllabus peers. Whilst this stopped short of fighting in full cloud it did include transits of up to 5000' and the ability to fly tactically in the presence of layered cloud. The simulator provided excellent representation of these environmentals but stopped short of providing broken cloud due to visual issues. As cloud was able to be built as the sortie progressed the later runs of each sortie were fought in the presence of ever increasing cloud layers. Whilst this could be experienced during any live sortie the simulator's ability to elect to insert cloud ensured all students had been exposed to a minimum level of cloud whilst the best students could be pushed further.

3.12.1 COUNTER AIR RESULTS

During the first live flight test of the first batch of 2 Trial subjects the student pilots failed their sortie (score 2 - low average) for sacrificing tactical formation in order to handle the weapon system. This failure was assessed as poor prioritisation with respect to maintaining Visual Mutual Support (VMS) leading to an incorrect workcycle. The second pilot produced a course standard performance but it was clear that he had learnt from his colleague; maintaining tactical formation at half the standard range in order not to lose visual. In line with the failure process set out in section 3.9.7 all pilots and students were debriefed at length in order to determine the root cause. These debriefs determined that both of the students' visual scans between the internal displays and external visual formation position were not quick enough and that the resulting positional errors in formation were not being addressed early enough. Cameras were set up within the simulated environment to examine head movement and thus the internal/external scan rate (Figure 3-15). The ability to see inside the simulator during instruction made explicit the skewed nature of the scan. It was readily apparent that practically no external scan was being conducted. After some experimental adjustment of the visual scene was conducted the cause of the issue was found, as with the Combat Phase, to be the visual set up of the Target Projectors (TPs). The images provided by the TPs had been set at their maximum brightness, to address the recognised Combat Phase issues. However the unintended consequence was that the pilot could immediately determine sufficiently the aspect, range and most importantly location of the leader without the need to spend time searching the sky. The result was that the student had no need to consider the prioritisation of tactical formation because it was particularly easy. After further experimentation with staff pilots a suitable visual setup was determined that necessitated the Target Projectors and navigation lighting on all aircraft to be switched off. Once found the students re-flew their final simulator sortie and live test flight using scenarios that contained no similar enemy presentations to the preceding live test flight. In this manner they were unsighted on enemy intentions and could utilise none of their knowledge from the previous failed test rides. The resultant scores were 4 - high average and 5 - above average respectively. In lieu of these results the simulator visual setup was fixed for the remainder of the test subjects.

Figure 3-15. CCTV - Monitoring the Student Visual Scan

A further student failure (graded 1 - below average) was experienced by a trial subject in the third batch of students after a performance airborne that resulted in the abandonment of the sortie. The reason given by the IP was the real-world environmental conditions (low sun and 100kts of tailwind behind Red Air), coupled with chaff remnants in the airspace conspired against the student to reduce his tactical effectiveness. This aligned with the student's post-sortie interview which stated that the sun was directly in front of him making it impossible to see any cockpit displays. The records show the sortie was flown within an hour and a half of sunset and with a threat direction of west, which would have provided the environmental conditions reported. Additionally the unusually strong wind meant that the chaff laid by the previous users of the airspace gave effects the student should never have been exposed to – either under live or synthetic conditions. To determine if the student's performance was limited by ability or the environmental conditions he did not undergo any retraining but was resubmitted for the live test sortie with differing (and harder) scenarios with a time and location set to avoid the previous issues. The result was a 'high average' indicating that the students' original score was a consequence of the environmentals rather than ability.

The overall marks for the Trial were compared to historic results, shown in Figure 3-16. It is evident that the Trial average (CA Trial PB: representing the subject's first effort at the test sortie) was lower than that of the historic results of the standard syllabus (CA Std Cse), this was accompanied by a large confidence interval caused by the sizable spread of marks in the small sample. CA PB Post Vul (Vulnerability) Training shows the overall grading of the subjects once the issues with the 3 failed students were retrained or removed as described above. The trial average is seen to increase above the historic average after retraining but more importantly the confidence interval is comparable, implying that synthetic training might return an equal or higher mean to the present syllabus.

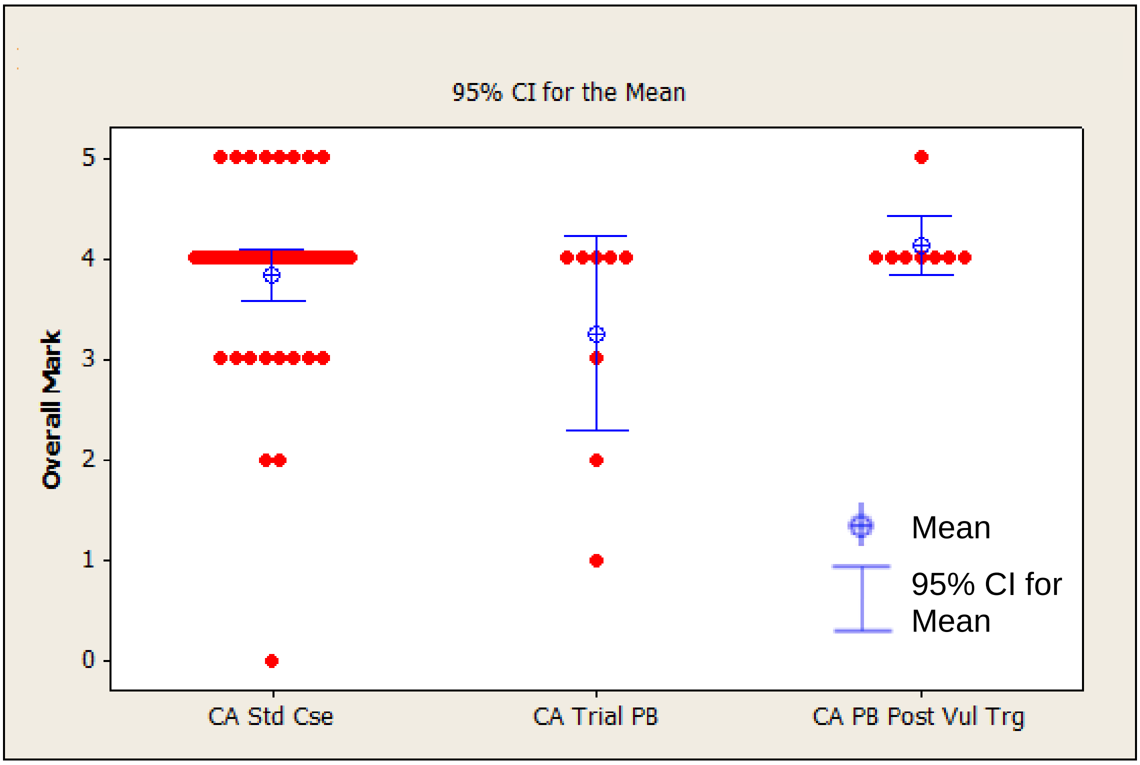

Figure 3-16. Comparison of Overall Marks for the Counter Air Phase between historical data (Std Cse), initial Trial Data (Trial PB) and revised trial data (PB Post Vul Trg).

3.12.2 COUNTER AIR SUMMARY

The results above pass tests 2 and 3 of Section 3.6 and thus the Counter Air Trial Syllabus is recommended for initial inclusion within the core syllabus to further expand the sample size. Whilst the successful nature of the trial phase indicated its suitability for inclusion the limitations found should be incorporated within the training process to prevent assets being wasted:

a. Contrary to the lessons identified in the Combat Phase, TPs and navigation lights should be turned off within the simulation environment to force a visual scan rate that transfers to live flight.

b. Environmental conditions, such as sun, that cannot be represented within the simulator should be avoided within the live environment. Similarly items that would not normally be expected within that part of the live syllabus, such as extremely strong winds or chaff, should either be avoided in live test flights or incorporated within synthetic training.

The observation of the effect of environmental conditions on performance was included within the risk register as, although identified within the Counter Air Phase, it was also likely to impact first solo flights where the risk to life was significantly higher (Sub-question 1, Objective 5).

In line with Objective 2, identifying factors that could affect the transfer of training, the Counter Air trial syllabus, unrestricted by the financial boundaries of training in live flight, was able to incorporate enemy radar waveforms and their impact much earlier. Subsequently the student's exposure to these was significantly higher than individuals on the standard course. Similarly, as identified above, the number of runs per sortie was more than double that of the standard course. Both these factors have the potential to affect the transfer of training by reducing the time and assets required in the simulator or training to a higher level in live flight.

3.13 QRA PHASE

The QRA phase consisted of a single simulator sortie followed by a live sortie, mirroring the standard syllabus, the only difference between the two syllabi was that the Trial did not permit any instruction from the rear cockpit within the live test. The simulator was flown as a 2-ship formation against a low, slow target (civilian Cessna aircraft) across the same airspace as the live flight would normally take place. The sortie included runs both with and without cloud although abilities to simulate the cloud experienced at lower levels were hampered by the simulator's inability to provide anything other than full cloud cover with no gaps, cloud normally experienced below 5000 feet being broken cumulus. Further issues were encountered when attempting to represent the effect of the ground and earthbound objects on the radar performance within the simulator. Radar performance will normally be degraded within the presence of ground clutter but, despite modifications, the simulator's modelling of this was not apparent, similarly the affects of the numerous wind turbines on the radar's false target rate was also lacking. The overall effect was to present a significantly clearer synthetic radar picture than was present in live flight. The final missing effect was that of traffic: below 10,000 ft overland increased civilian traffic clutters the radar picture and significantly increases communications on the frequencies the student would be listening to. Whilst the extra traffic in the area could be modelled the faultless performance of the radar allowed these to be broken out easily, additionally the increased amount of communication was also difficult to provide despite dedicating a single simulator operator to its replication.

3.13.1 QRA LOW SLOW PHASE RESULTS

Figure 3-17 provides the overall results for the phase. Even cursory analysis of the overall marks between the 2 syllabi shows a spread of grades much wider than the syllabus for the standard course. It can be stated that "the mean of the standard course *is* greater than that of the trial course for the sample used" where $p>0.1$.

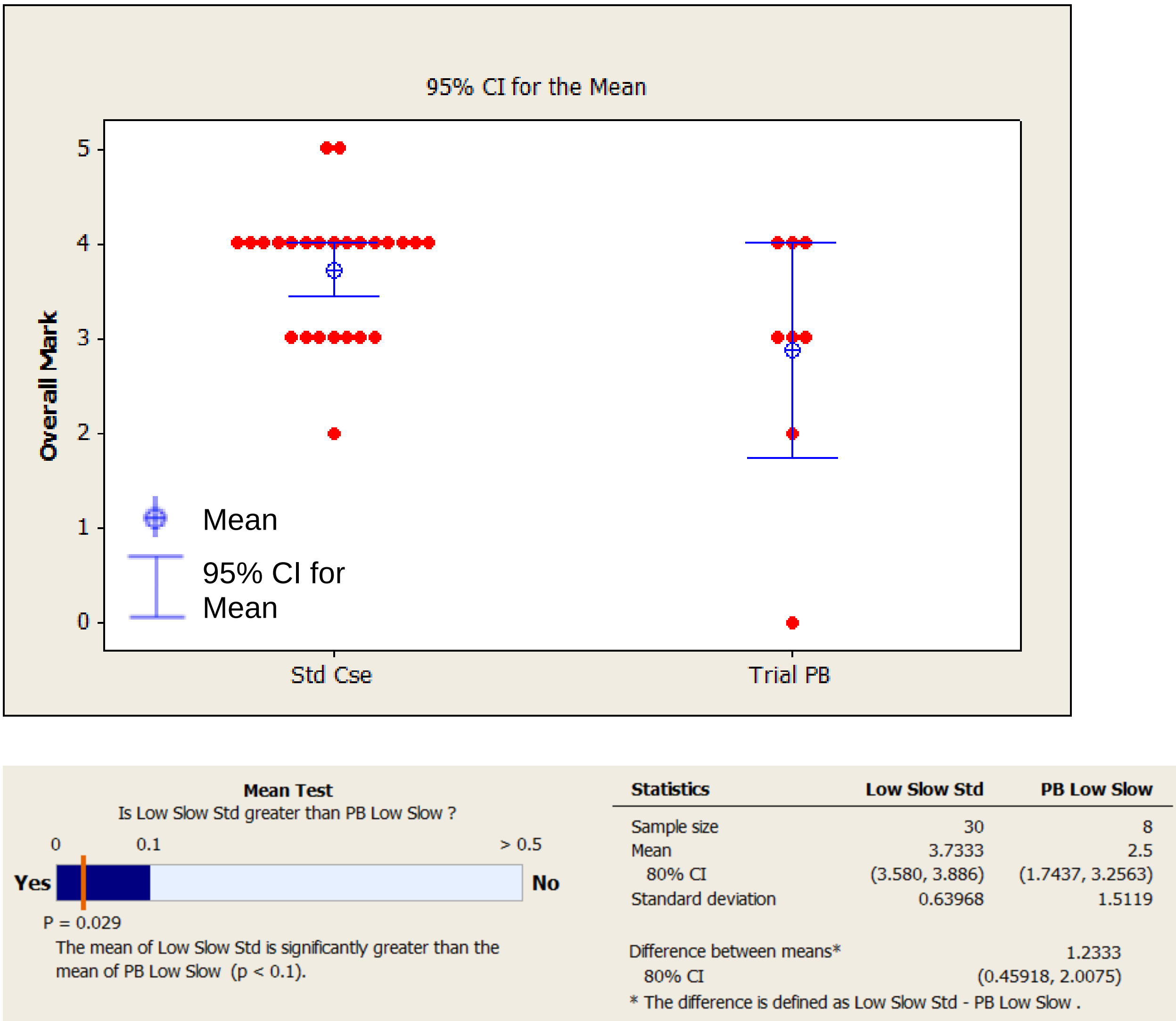

Figure 3-17. Comparison of Overall Marks of Standard (Std Cse) and Trial (Trial PB) Syllabus in QRA Low Slow Phase

Of the eight subjects to conduct this portion of the trial 4 failed; two scoring low average, one below average and one score of 0 - unsafe. The reasons provided by the IPs were closure to within minimum distances within cloud, a loss of Situational Awareness (SA) near other traffic and within cloud during manoeuvring, incorrect and dangerous formation positioning around cloud and poor weapon system handling. Additionally there was an incorrect assessment of a 200ft passing distance on the bogey aircraft, an inability to maintain SA within busy airspace and poor radar handling at low level in cloud.

Unusually one of the failed students also suffered airsickness towards the end of this sortie.

The reasons for failure can be distilled therefore as cloud, spacing estimation and SA. As discussed in the Counter Air phase the simulator was not capable of providing realistic cloud formations in anything other than stratus layers. This is more representative of the cloud encountered at height (>5,000 ft), as in the CA phase, where large cumulus cloud would simply be avoided. During QRA sorties at low level scattered and broken cloud is accepted as part of the working environment and the student must be able to anticipate its affect on the formation and the task. Thus, it is possible that the training simulation was not able to replicate this element to the level required to transfer to live flight.

The second issue of spacing demonstrates, as in the BFM phase, a lack of transfer of training with respect to understanding an opposing aircraft's range, aspect or orientation. Given the sensitivity of the visual setup in the previous phases it is likely that the correct setup had not been found for the QRA low slow phase.

The third problem, that of SA, is less tangible. On initial examination the fault may reasonably be levelled at a reduction in capacity through high levels of physical stimuli, brought on by heavy manoeuvring of the aircraft under high workload, in cloud at low level; the student's airsickness being a symptom. As the failure sample size is too small and the flight paths flown too dissimilar, it is not possible to ascertain this as the true cause. Additionally, however, the simulator itself was noted to provide poor training in a number of key areas for this phase. The first was that the target aircraft was able to be acquired on radar at a range much greater than to be realistically expected in live flight. This reduction in difficulty was compounded by a radar model of the terrain that provided no clutter. Some of these issues can be traced to the terrain database but the incorrect radar pick-up range and the tenacity of its lock are within the radar software. This software was an older version of that presently installed in the aircraft and whilst the ‘switchology’[5], handling and displays were the same this phase provided the first of the noticeable differences during training. The second issue was that of provision of realistic level of simulated traffic in the local area. Even after eight students had passed through the trial experienced simulator operators were unable to consistently provide traffic that tested the student at the correct time, nor were they able to give an associated level of communications chatter whilst running the other aspects of the lesson. The effect of these deficiencies was to provide the student with a much simpler environment to train in.

3.13.2 QRA LOW SLOW PHASE SUMMARY

It can be seen from the results that the simpler environment did not train half of the sample group to a level that tested their ability to work to a realistic timeline. Thus, in the air the student’s workcycle was altered away from that learnt in the training environment, which in turn affected the prioritisation of problems allocated by the student. The result in each case was poor formation coordination and a loss of SA culminating in overall poor training transfer. The failure of the phase to pass tests 2 and 3 of section 3.6 resulted in a recommendation *not* to include this phase within the core syllabus to increase sample size.

The technical limits to training (required by Objective 3 of Sub-Question 1) were: the inability of the simulated radar software to replicate live flight close to the ground, synthetic modelling of the ground clutter, poor ability to replicate communication chatter and cloud modelling.

3.14 EFFECT ON RESOURCES

In order to answer the requirement of Objective 4 of Sub-Question1 – determine the level of training transfer between the simulator and live environment and provide additional data for the ‘Capacity’ and ‘Fleet Structure’

[5] ‘Switchology’ is the colloquialism for the switch selections required to achieve sets of tasks.

elements of the Systems approach shown in Figure 3-1 a summary of the resources used was made (Table 3-11). The table shows a comparison of the assets used in each phase of the trial when compared to the standard syllabus.

Table 3-11. Resources Used

Phase	Standard Syllabus				Trial Syllabus			
	Simulator Hours (HH:MM)	Aircraft Hours (HH:MM)	IPs Req'd	External Assets	Simulator Hours (HH:MM)	Aircraft Hours (HH:MM)	IPs Req'd	External Assets
Conversion	3:45	8:15	8		15:50	1:15	8	
BFM	2:30	12:30	10		15:00	2:00	7	
Counter Air	31:45	20:00	32	10 Hawk 2 FRA	40:00	2:30	21	2 Hawk
QRA Low Slow	2:30	2:30	4	1 Tutor	2:30	2:30	4	1 Tutor
Totals	40:30	43:15	54	10 Hawk 2 FRA 1 Tutor	73:20	8:15	40	2 Hawk 1 Tutor

From Table 3-11's figures the Transfer Effectiveness Ratio can be calculated. Transfer Effectiveness Ratio (TER) is defined as number of hours training saved on operational equipment divided by number of extra hours on the simulator (see section 2.5.4). It provides a comparison of the effectiveness of an hour spent in the simulator when compared to spending that hour airborne. These results are given in Table 3-12 along with the savings of RAF Instructor Pilots and any external assets.

Table 3-12. Savings Achieved

Phase	Aircraft Hours Saved (HH:MM)	Extra Simulator Hours (HH:MM)	TER	IP Savings	Remarks
Conversion	7:00	12:05	57%	0%	Successful phase. Simulator times include time taken to start up, taxi and shut down. Recommended.
BFM	10:30	12:30	84%	30%	Partially successful phase. Not Recommended.
Counter Air	17:30	8:15	212%	34%	Successful phase. Recommended.
QRA Low Slow	0	0	0%	0%	Unsuccessful phase. Not Recommended.

Similarly it is possible to calculate the Support Flying Transferred by the implementation of the trial syllabus, where support flying are all aircraft not containing a student but that take part in each mission:

$$\textbf{Support Flying Transferred (SFT)} = ((Sp_c - Sp_x) / Sp_c) * 100$$

Where Sp_c = Support live flying required in the control group syllabus.

Sp_x = Support live flying required in the trial group syllabus.

Table 3-13. Support Flying Transferred

Phase	Standard Syllabus			Trial Syllabus			SFT (%)
	Total Live Hours inc Support (HH:MM)	Total Live Student Hours	Support Hours only (HH:MM)	Total Live Hours inc Support (HH:MM)	Total Live Student Hours	Support Hours only (HH:MM)	
BFM	12:30	6:15	6:15	2:00	1:00	1:00	92
Counter Air	35:00	11:15	23.45	5	1:15	3:45	84
QRA Low Slow	3:45	1:15	2:30	3:45	1:15	2:30	0

Table 3-12 and Table 3-13 demonstrate the effect of simulation on resource usage for the trial syllabus when compared with the standard. It can be seen that for the successful phases simulation frees significant resource for reinvestment elsewhere in the training system or to increase the numbers of pilots trained.

To determine if the Typhoon force has sufficient capacity to train using the new methodology an assumption of a new syllabus construct has to be made that replaces successful phases of the trial within the standard syllabus. Thus, a potential new syllabus could have the construct seen in Table 3-14. Given the simulators have 165 training hours per week and that the 60 IPs will require to conduct 1 emergency simulator sortie of 1hr 15mins per month the remaining training time available is 156hrs 15mins. There are only 2 Full Mission Simulators available and thus time available to inhabit those FMSs are 78 hours per week. Even at the surge rate, which assumes 2 x 1hr 30min event a day for a course of 4 students, only 60 of the 78 hours would be consumed, thus there is sufficient capacity within the simulation training plan to accept the increased demand.

Table 3-14. New Construct Syllabus

Phase	New Construct Syllabus			
	Simulator Hours (HH:MM)	Aircraft Hours (HH:MM)	IPs Req'd	External Assets
Conversion	15:50	1:15	8	
BFM	2:30	12:30	10	
Counter Air	40:00	2:30	21	2 Hawk
QRA Low Slow	2:30	2:30	4	1 Tutor
Totals	60:50	15:45	43	2 Hawk 1 Tutor

Assuming an average of 15 maintenance hours per live flying hour and 8 students every 4 month cycle, the new construct syllabus has the potential to save a total of 3,300 man maintenance hours every 4 months. This would permit a considerable repositioning of engineering manpower to more overstretched areas. Of note, however, is the new Live Synthetic Balance (LSB) of 16 hours live and 61 synthetic. Whilst the focus to this point has been on the student this new LSB would be difficult to apply for the instructors without dropping below the accepted safe live flight minimum per month of 12 hrs 30mins. Over a 4 month course the 8 students resident on the OCU will only require a total of 126 live hours. Even if every live flight were flown with an IP in the rear seat for safety the live flying available for the 20 IPs on the OCU would fall as low as 1hr 30mins per month. Whilst this would infer a reduction in IP numbers this route is prevented by the need to instruct within the simulator. Thus, it is the LSB and the effect on safe numbers of live hours that is the limiting factor within the present Typhoon Force construct.

3.15 CONCLUSIONS

Trial PANDORA'S BUZZARD sought to answer, through a series of objectives, the first of the research Sub-Questions: "What limits synthetic training in the instructional environment?" The research was undertaken within the structure of Figure 3-1. Systems Approach High Level Intentions', which intended to ensure that the objectives, whilst answering the first of the research themes, also fed towards usable output, namely information on the capacity and structure of the Typhoon force. Thus, each of the objectives' findings are summarised below before addressing the impact on the capacity and structure of the force.

3.15.1 OBJECTIVE 1- ASSESS THE SUITABILITY OF ASTA TO TRAIN EACH PHASE OF THE OCU SYLLABUS

The success of the Conversion and Counter Air phases should be tempered with the acknowledgement of the statistical disadvantages associated with small samples. In reality, however, practical considerations such as cost and time associated with training front-line pilots prevent large statistically-impervious trials from being undertaken. The aim of this work has been to examine the results from a number of different aspects to provide a view on whether fully-synthetic training is viable. In the instances of the successful phases, Conversion and Counter Air, the initial indications are that synthetic training in these areas should be maximised, allowing the RAF to expand its sample size and generate a corporate knowledge of the strengths and weaknesses of a student trained in this manner. The failed phases of Combat and QRA Low Slow, demonstrate areas that had substantial weaknesses, even given the small sample size, and thus should synthetic training wish to be furthered in this area considerable thought and syllabus redesign is required.

The Trial has recommended the inclusion of both the Conversion and Counter Air Trial Syllabus into the OCU's syllabus. The Combat phase, whilst capable of imparting conceptual understanding of the combat manoeuvring, has not been recommended due to issues with the visualisation of the enemy's dynamic manoeuvring caused by the Target Projectors, that were resident across a number of close-range skillsets. The QRA Low Slow phase was not recommended for inclusion for reasons of synthetic replication and safety whilst flying in formation in close proximity to the ground and other aircraft.

3.15.2 <u>OBJECTIVE 2 – IDENTIFY ANY VARIABLES THAT COULD AFFECT THE TRANSFER OF TRAINING</u>

The effect of the incorrect setting of the visuals impacted the transfer of training, the evidence being seen most clearly in the Counter Air Phase. The incorrect setting established a workcycle that did not correctly prioritise lookout, effectively allowing the student extra time to interpret the displays. Whilst this set up was recognised and corrected it infers a sensitivity to the workcycle that was not clear prior to the trial. These effects on workcycle were also observed in the QRA low slow sortie with extra time having to be spent interpreting the displays as the simulator lacked any effects of the ground on the radar model. Allied problems such as poor cloud modelling and incorrect traffic levels that threatened the formation are also likely to have skewed the workcycle, compounding the effects.

Environmental conditions such as the effect of sun or strong winds at the tropopause provided problems for trial students and hampered any transfer of their training into the live environment. Thus, the conditions of the day should be considered carefully by the IP to determine if there are factors the student has not witnessed within the simulator. Similarly syllabus construction should aim to provide the student with as many of these experiences as possible in the synthetic environment.

A variable that provided a positive effect on the transfer of training was the ability within the counter air phase to be able to increase the number of runs conducted per sortie. Allied with the ability to replicate correct threats, without financial considerations, from very early on in the syllabus ensured the student's performance reached course standard early in the phase's training.

3.15.3 OBJECTIVE 3 – IDENTIFY ANY TECHNICAL LIMITS PROHIBITING THE EMPLOYMENT OF HIGHLY-SYNTHETIC SYLLABUS TO TRAIN STUDENTS

As stated above correct visual setup was a technical restriction in the Counter Air phase and necessitated a particular setup to provide adequate training. In the preceding Combat phase the Target Projectors were particularly poor, preventing the correct visual cues from being identified. This issue has been addressed to a degree by superior visuals delivered in 2014 however it should be noted that a single setup that suits all training phases is still not available.

Great efforts were made to ensure that the software in the simulator matched that of the aircraft the Trial Subjects would be tested on e.g. certain radar functionality introduced into the aircraft midway through the training was not available in the simulator and consequently the students did not utilise the features in the air. The importance of maintaining software commonality follows from the observations of the sensitivity of the student's workcycle in Objective 1. As an example a simulator where the fidelity of the radar modelling allowed a detection range of the target at 10% greater than the norm would skew the student's workcycle away from the need to rapidly manipulate and interrogate radar tracks and instead provide them with plenty of time to lookout and maintain correct tactical formation. In the live environment, however, the student would not now have the skillset to conduct these manipulations in the time required, forcing them to focus heavily on the displays with a subsequent effect on their lookout and formation. A key point is the normalising effect of live flight (validating the decision not to seek a quasi-transfer solution); a lower fidelity simulator will be satisfactory if the pilot spends a greater proportion of his or her time flying the aircraft where motor skills are corrected to the true environment. If majority synthetic training is to be undertaken then all aspects affecting the workcycle need to have a fidelity as close as possible to that of the live environment.

3.15.4 OBJECTIVE 4 - DETERMINE THE LEVEL OF TRAINING TRANSFER BETWEEN THE ASTA AND THE LIVE ENVIRONMENT

The levels of training transfer and effectiveness in Table 3-11 and Table 3-12 show that for the approved phases there are considerable savings in terms of flying hours and IP time that could be saved. The most successful of the phases, Counter Air, returning a Transfer Effectiveness figure of 212%, 84% of the support flying transfered and a reduction in the numbers of IPs used by 34%. The LSB on the new construct syllabus that incorporated the successful phases was also increased to 80% synthetic and 20% live.

3.15.5 OBJECTIVE 5 - GENERATION OF A RISK REGISTER TO DETERMINE THE RISKS DURING THE FIRST SOLO FLIGHT, WITH PARTICULAR REFERENCE TO RISK TO LIFE

The risk register for first solo flight was generated and held at Typhoon force HQ and an example is repeated at Appendix C.

3.15.6 CAPACITY AND STRUCTURE

Figure 3-1 provided the flow of understanding for Trial PANDORA'S BUZZARD culminated in a requirement to comment on the capacity of the Typhoon force, and determine any possible impact on the structure should a syllabus using a high proportion of synthetics be incorporated. Section 3.14 found that limit of using a new construct syllabus would be the new LSB of 16 hours live and 61 hours in synthetics which, if employed in the current structure, risked placing the IPs below the minimum safe hours to be able to conduct QRA - 12.5 hrs (CINC-Air, 2009). By way of validation of these calculations in 2012, against advice, 29 Squadron, the OCU, incorporated an immediate inclusion of both the Conversion and Combat phases of Trial PANDORA'S BUZZARD.

Early observations of the Officer Commanding 29 Sqn were that IPs, now conducting proportionally more synthetic events, quickly dropped below the minimum safe live flying hours and the trial was abandoned. This validation demonstrated that whilst the trial moved the Live Synthetic Balance ratio from approximately 1:1 to 1:5 the inclusion of all-synthetic elements is not possible without other considerations:

a. An awareness of the live flying minima.

b. The possibility of saturating the number of simulator slots should the number of students be increased.

c. The recognition that the simulator is a critical asset and prioritise its concurrency with the aircraft software accordingly.

d. Consideration to increase the number of simulator sorties instructed by BAE staff (who presently only operate the simulator). This would require recognition, approval and standardisation of their skillset but would decrease the number of IPs required on the Squadron as a whole, thus increasing the number of live hours available per pilot.

e. Awareness of the impact on engineering manpower and structure with a much reduced task: a saving of approximately 9,900 man hours per year.

An alternative solution to the manipulation of the manpower structure for the OCU is to saturate the simulator usage by employing simulation evenly across the whole of the Typhoon Force and in doing so reduce the proportion of simulation undertaken on the OCU. This would necessitate a further trial to determine the applicability of simulation to front-line training, this trial was termed Trial PANDORA’S DIAMOND and is discussed in Chapter 4.

Chapter 4. Trial PANDORAS DIAMOND

4.1 OVERVIEW

This work was intended to provide insight and evidence into the blend of synthetic and live training acceptable to the Typhoon front line through the lens of a gathered sample opinion of RAF Coningsby and RAF Lossiemouth Typhoon pilots. Trial PANDORA'S DIAMOND is complementary to Trial PANDORA'S BUZZARD in that it seeks to extend the understanding of simulation, albeit using differing methodology, into the realms of front-line training. The trial explored cultural acceptance as well as specific event-based use, the intended end result being to provide the Typhoon Force Commander with an indication of the current optimal synthetic blend, whilst also asserting the accepted limit of that blend. In the process, proof was sought on emerging opinions as to the level of acceptance of simulation and the impact of the level of complexity.

4.2 KEY LITERATURE REVIEW DIRECTION

The literature review chapter provides an indication of the perceived importance of the Live Synthetic Balance with many articles arguing that the primary reason for determining this balance is financial (Schank *et al.*, 2002; Harper and Hillier, 2007; Kruzins, 2008; Wells *et al.*, 2009). This focus on the monetary aspect appeared to display little consideration for the actual tactical training it was to replace, however from late 2009 onwards an understanding of just how much simulation the new Joint Strike Fighter (JSF) would use began to make its way into open press (Schank *et al.*, 2009). By 2012 this was to become an accepted Live Synthetic Blend (LSB) ratio of 1:1. Nevertheless the RAF altered little with respect to the employment of simulation until the success of Trial PANDORAS BUZZARD indicated that simulation may have reached a turning point. A demonstration of the RAF's realisation was that the trial and its research were awarded the Central Flying School's award for Instruction Excellence, the 1 Group award for Operational Innovation and 22 Group and the Guild of Educator's award for Educator of the Year, demonstrating the RAF's willingness to embrace the topic of simulation. Having seen the potential for simulation at the tactical level Air Officer Commanding 1Gp approved a new trial to determine how the LSB was to be determined with respect to training and tactical capability, the results from the trial were to inform the creation of a new training syllabus. From this point onwards rhetoric within the RAF surrounding simulation began to develop from financial argument to one of tactical and operational capability. This was publicly witnessed during a briefing to XI Squadron personnel during the Advance Training and Leadership Course at the Al Dhafra Airbase in the UAE by Air Vice Marshall Waterfall, on 12 November 2014.

The review tracked the increasing importance of understanding the amount of training that could be conducted via simulation in order to balance the competing requirements of finance and training value. This showed a demand for information but failed to provide a clear picture of the possible. Trial PANDORA'S BUZZARD had demonstrated that employment of simulation lagged behind its capability. Thus there was a need to gain an understanding of the level of possible employment of simulation that would necessitate ensuring the sample of pilots questioned were au fait with all of its capabilities. As the research would not include a fully-synthetic training syllabus, simulation employment would be limited by cultural influences and beliefs rather than any exposure of practical or technical limitations.

4.3 AIM

The aim of Trial PANDORA'S DIAMOND was to discover cultural limits of, and subjective opinion on, the use of synthetics on UK front-line Typhoon pilot training using an informed audience. The Trial was split into 2 phases, the first intended to provide an assessment of the cultural environment and the RAF Typhoon pilots' views on the utilisation of synthetics within their day-to-day training. The second phase concerned additional, more academic considerations and provided 'fill-in data' that gave a more rounded and complete set of results to assist both with the research question and subsequent simulation procurement. Phase 1 of the Trial took place between 1st October 2013 and 1st March 2014 and used simulation assets within the Typhoon ASTA (Aircrew Synthetic Training Aid) at the Typhoon Training Facility (TTF) and a total of 48 pilots from 3(F), XI and 29 Squadron. Phase 2 extended the initial cultural questionnaire to the Scottish base of RAF Lossiemouth. This report amalgamates both phases.

4.4 OBJECTIVES

Leading on from Sub- Question 2 - ‘ What is the cultural limit of synthetic use?’ the objectives of Trial PANDORA’S DIAMOND were as follows:

a. **Objective 1**. Determine the present level of cultural acceptance of synthetic use in front-line Typhoon training and investigate if this is common across both Typhoon bases.

b. **Objective 2.** Identify any factors, such as experience levels, that correlate to the level of acceptance.

c. **Objective 3.** Prove or disprove some commonly-held beliefs that had started to form since the initial Trial PANDORA’S BUZZARD had increased synthetic use, namely:

i. Younger pilots are more accepting of simulation than elder pilots.

ii. Acceptance of simulation is based upon a pilot’s qualifications.

d. **Objective 4**. Determine the subjectively-assessed LSB for each of the required tasks.

e. **Objective 5**. Investigate the affect of threat complexity on the LSB.

4.5 DESIGN

4.5.1 QUESTIONNAIRES

All pilots conducted an Entry Questionnaire (Appendix F), the purpose of which was to collect background data and determine the respondent’s opinion of simulation across 3 themes using aggregated results from Likert scale responses (Likert scales require the respondent to specify their level of agreement to a series of statements on a symmetrical positive/negative scale). Thus, spread across the 18 questions were 6 questions on each of the 3 themes; those being:

a. **Cultural Lean.** The acceptance level of the amount of simulation presently used in day-to-day training.

b. **Near Future.** The acceptance level of the proposed near future developments of simulation that were intended to be applied to Typhoon training.

c. **Simulation can provide an Experience.** Einstein is attributed to have said ‘Only experience is knowledge. Everything else is just information’. This theme asked the pilots if they believed that the simulator could generate knowledge through the provision of a valid experience, when compared to that gained in live flight.

To ensure all of the audience shared the same level of awareness with respect to the capabilities of simulation each pilot underwent a series of three synthetic sorties, of increasing complexity levels, each with a Post Sortie Questionnaire (example in Appendix F). Each of the Post-Sortie questionnaires asked the respondent to provide the optimum LSB for each of the skill sets that the Tactical Air Command Task (TACT) syllabus stated were required at that level of complexity. This was achieved by the respondent providing the minimum number of times a skillset should be undertaken in live and in synthetic training each year. Finally a random sample were given an Exit

Questionnaire (Appendix H) to determine if their initial view of simulation had changed over the period.

Reliability Measures. The importance of reliability and validity of the questionnaires is highlighted in Bryman (2008, p.149); reliability being the 'consistency of a measure of concept'. Reliability is made up of two areas: stability and internal reliability. In order to demonstrate the stability of the trial the Exit Questionnaires asked the same questions of the respondents once they had completed all simulation sorties. The 'Cultural Lean' results were matched to the answers given in the Entry Questionnaire and assessed to see if the opinion of the respondent had been altered by participation in the trial itself, or whether their view had remained constant; indicating that all their post-sortie questionnaires would have been conducted with a 'stable' view of simulation. Internal reliability is concerned with the relationships between the indicators remaining coherent between respondents, as such a test of correlation was carried out to determine the level of coherence. In order to seek the best result for internal validity the questions were refined over a four month period May – August 2013 whilst the author was posted to Mount Pleasant in the Falkland Islands, using a total of 16 pilots that were cycled through the Flight during this time. The remote nature of this location allowed the subjects to contribute freely to the forming of the questionnaire without of any oversight of a command chain.

Validity. Bryman (2008, p.151) states that validity of a questionnaire is concerned with 'whether a measure of a concept really [does] measure the concept'. The methods of determining validity and the trial's method of testing them are as follows.

a. **Face Validity.** The measure is sensible and acceptable to experts in the field, in this case Typhoon Force Headquarters was shown the draft questionnaires, the measures intended and was invited to comment. After an interview that discussed the measures with the Typhoon Force Commander the trial was commissioned in his name.

a. **Concurrent Validity**. Introduction of a criterion that is known to differ and is relevant to the concept. Within Trial PANDORA'S DIAMOND this was met by the use of the RAF Lossiemouth personnel in Phase 2. These pilots fly the same aircraft but do not have access to the same level of simulation, thus the cultural acceptance of synthetics and the reliability of the questionnaires is expected to differ compared with the RAF Coningsby pilots.

b. **Predictive Validity.** The correlations between the level of cultural acceptance and variables determined in the RAF Coningsby pilots would be used to predict the level of cultural acceptance of the RAF Lossiemouth sample.

4.6 SORTIE COMPLEXITY

Each level of complexity (labelled 1-3) matched the classified threat levels that the Typhoon was expected to face, 3 being the highest. An unclassified illustration of threat levels being:

a. **Level 1**. Flown as a pair using Quick Reaction Alert profiles, basic geometrical intercepts and Air Defence up to 2v2.

b. **Level 2**. Flown as a four ship in an 8 v 16, working in a coalition air package testing air-to-air tactics and self-escorted weapons release in a hostile Electronic Attack (EA) environment and using GPS and / or Laser Designation Pods. High Value Asset Defence Surface-to-Air missile threats, the majority non-networked.

c. **Level 3**. Flown as a four ship at night in a 16 v 16, using F-22, F-15 and GR4s against the 'most threatening' and 'most likely' threats. Air threats armed with long range missiles and expert tactics combined with knowledge of their own networked Surface-to-Air Order Of BATtle (ORBAT).

4.7 SORTIE DESIGN AND CONSISTENCY

Each of the sortie scenarios were provided with Air Tasking Orders (ATOs) and SPecial INstructionS (SPINS). In order to provide a realistic yet unfamiliar training area the missions were set over the Exercise RED FLAG ranges at the USAF Airbase of Nellis near Las Vegas and to ensure consistency of the threat, enemy assets were assigned doctrines to provide them with governing behaviours, awareness and abilities. ASTA instructors were issued with a ‘run sheet’ providing a timeline of the intended problems to be solved by the formation, again providing consistency between the formations experiencing each level. The operators were also required to act as formation leaders and controllers of the allied aircraft. All operators participated in a test mission prior to instructing a trainee in order to ensure misunderstandings in presentations were solved and all subjects would receive a sortie as near identical to their colleagues as was possible.

4.8 EQUIPMENT UNDER TEST

Both Full Mission Simulators (FMS) and Cockpit Trainers (CTs) were used for all sorties, with software load 3.1.X, which equates to an aircraft PSC (Production Software Configuration) of 4.3 with Drop 1 upgrade. Whilst concurrent to some of the aircraft in the live fleet, the majority now have a Drop 2 upgrade to the displays and controls and a new issue of radar software – R2Q. During the trial period one of the simulators, FMS 1, was taken down for long term maintenance, this loss of 25% of the synthetic assets was to impact the throughput of pilots and necessitated an alteration to the trial design as detailed in section 4.10 below.

Whilst not used directly it is important to note that the simulation devices used by RAF Lossiemouth have considerably lower fidelity than those at Coningsby. Whilst this lower fidelity did include the visual domain it also extended to items such as the aircraft systems and the ability to replicate true enemy tactics, additionally there were only 2 linked devices at Lossiemouth compared to Coningsby's 4. These limitations mean that the Lossiemouth simulators were used almost exclusively for emergency training only, the more tactical training being left to the live domain.

4.9 TRIAL SUBJECTS

Only qualified Typhoon pilots were used from the RAF Coningsby squadrons for phase 1 of the Trial. 'Qualified' being defined as successful completion of the Typhoon Operational Conversion Unit (OCU). A total of 36 pilots undertook the trial from the front-line squadrons, representing 100% of their manpower. 29 Sqn provided a further 12 of their 25 pilots, these were selected at random by the programmer as leave and operational constraints were taken into account.

4.10 TRIAL CONSTRAINTS

As stated above phase one of the trial was intended to provide immediate advice for the intended introduction of the new Typhoon training syllabus thus, in order to receive political approval for the remainder of the trial the scope was reduced. 29 Sqn would complete only the second simulation sortie as well as the entry and exit questionnaires. The decision not to include level 1 was taken because of the Squadron's high synthetic exposure at level 1, a function of their specific training role. It was decided, therefore, that there was nothing to be gained from testing at a level the instructors taught synthetically day to day.

4.11 TRIAL RESULTS

In total 77 pilots participated in the trial; 48 from RAF Coningsby and 29 from the smaller northern main operating base of RAF Lossiemouth, representing over 85% of the Combat-Ready pilots of Typhoon. All completed the Entry Questionnaire, however the actual simulation sorties were open to RAF Coningsby pilots only as the intent to move pilots from Lossiemouth to Coningsby to participate was deemed too costly by the HQ and funding was unable to be secured. Of the 48 Coningsby pilots 47 completed the level 1 simulator and questionnaire, 36 the level 2 and 20 the level 3.

4.12 ENTRY QUESTIONNAIRE

The Entry Questionnaire was designed to examine the respondents' opinion of simulation across 3 themes: the acceptance of the present use of simulation in training, to be referred to as 'Cultural Lean'; 'Near Future'; and the use of simulation to provide an 'Experience'. Before assessing the results the questionnaire was examined for internal reliability and stability in order to prove its validity.

Cronbach's Alpha test (Laerd, 2015b) is a statistical method to measure to what degree the items on a scale (in this case a Likert scale) are measuring the same dimension. It was used to assess the internal reliability by calculating the average of all the split half reliability coefficients across the 6 questions for each of the 3 themes. An acceptable level of reliability was set at 0.7, 1.0 demonstrating perfect consistency between all the questions and 0 being no consistency, the level of 0.7 being deemed acceptable in common practice (Wikipedia, no date b; Kline, 2000, p. 13). The results for each of the themes were as listed below.

a. **Cultural Lean.** The analysis for the questionnaire taken at RAF Coningsby showed a satisfactory alpha level of 0.74. As such all of the associated questions can be said to be reliably measuring the same concept. At RAF Lossiemouth, however, Cronbach's Alpha was measured at 0.68%. Omission of statement 1, 'The amount of simulation in day to day training COULD be increased', raised the value to 0.76. Whilst it was possible to raise the Alpha of Coningsby's results this could not be done by omission of the corresponding questions for Lossiemouth, as a consequence the results from the two bases could not be directly compared as the understanding of the theme 'Cultural Lean' is subtly different between the two.

b. **Near Future**. RAF Coningsby returned a Cronbach's Alpha of 0.75 whilst RAF Lossiemouth's result was below the acceptable limit at 0.59. Examination of the results showed that it was not possible to omit a single question and raise the percentage above the designated acceptable level. Thus the Near Future theme could not be used for RAF Lossiemouth.

c. **Simulation as an Experience**. Although this theme had proved consistent during the initial investigation the results of the scaled up responses across both Coningsby and Lossiemouth proved to be much lower than expected and well beneath the acceptable level at 0.24 and 0.33 respectively. Unlike the Cultural Lean theme, however, there was no single question responsible for the low figure. As a result this theme was discarded from use.

There were 6 questions per theme, the Likert scale scores were measured between 1 and 5 (5 being the most positive of simulation), 3 being the null point of opinion. Thus a neutral opinion would be centred on a score of 18 for each of the themes tested at Coningsby. At Lossiemouth, however, having had a question removed to increase internal validity, the null point would be at 15.

The level of stability of the answers given in the trial can be assessed by the level of difference between the answers given in the entry and exit questionnaires for the theme of 'Cultural Lean'. Analysis of the results showed that there was no significant difference (at a 5% significance level) between the opinions of the respondents as they entered and exited the trial and thus the trial results for phase 1 can be described as stable, and thus valid.

4.13 PERCEPTION VERSUS INDIVIDUAL OPINION

The respondent's Entry Questionnaires required answers from both their own point of view and their perceived position of the Fleet as a whole. The intent was to determine if they considered their opinion to be in line with that of the Fleet *en masse* or if they saw themselves as occupying a more positive or negative position. This would allow an understanding of how opinion differed when the pilots were questioned as a collective versus an individual's actual stance. This was intended to inform the process of how the RAF and the sub-contractors went about their information gathering in this new field.

A paired-t test (Figure 4-1) was carried out to determine if there was a statistical mean difference between the answers given for the individual (IND) and their perception (PER) for the 'Cultural Lean (LEAN)' category at Coningsby and Lossiemouth and the 'Near Future' category at Coningsby only. The paired t-test uses the same individuals tested at two different points in time on the same dependant variable and in doing so has an increased 'power' to look for the differences between the means. The null hypothesis for the test (H_0) was as follows:

H_0 = The pilots' INDividual opinion is the same as their PERception of the Force's opinion.

Thus, it follows that the alternate hypothesis (H_0) was:

H_A = The pilots' INDividual opinion differs from their PERception of the Force's opinion.

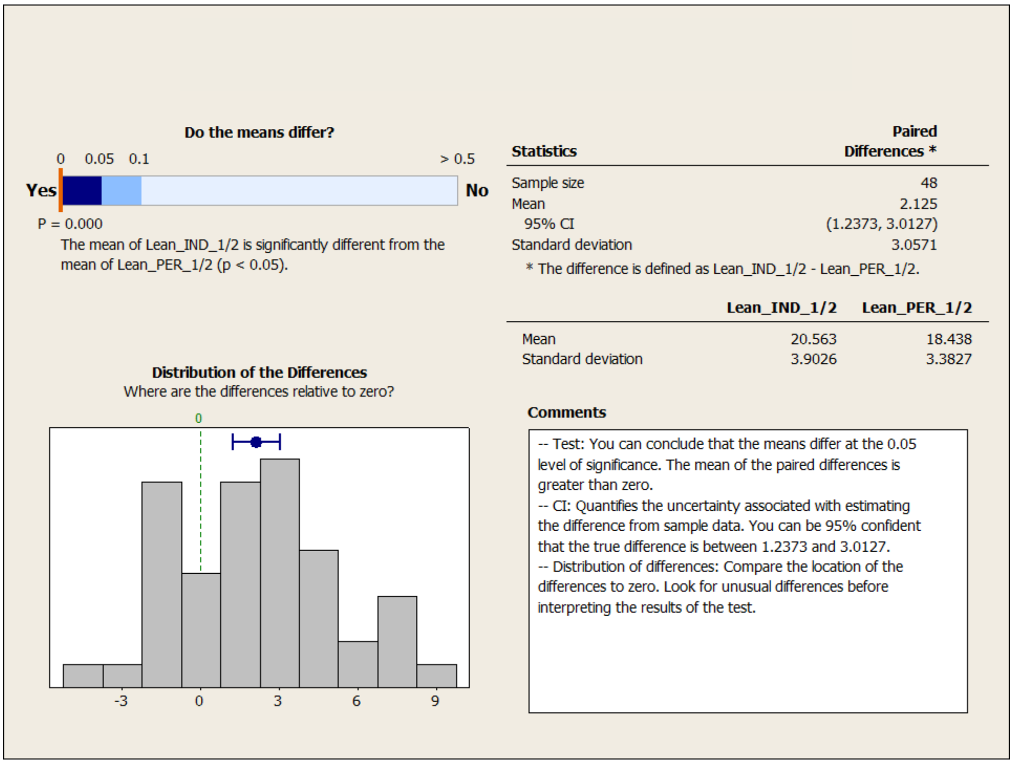

Figure 4-1. Summary of Paired-t Test at RAF Coningsby for Cultural Lean

Figure 4-1 shows the results for cultural lean at RAF Coningsby. There were no unusual points detected and the assumption of normality was not violated as assessed by a Ryan-Joiners test (p=.995) which examines normality of data by looking at the correlation between the test data and normal scores of the data. A correlation co-efficient close to 1 infers that the population is likely to be normal (Minitab, 2012c).

Pilots believed themselves to be more positive about simulation ($\bar{x}$=20.563, σ_x=3.90) than their perception of the Force as a whole ($\bar{x}$=18.438 σ_x=3.38), a value greater than 18 being a positive view of simulation. Thus, a statistically significant difference of 2.125 was demonstrated (95% CI, 1.237 to 3.013), t(47) = 4.82, p<0.001, d = 0.70. This presentation of results is used throughout the remainder of the thesis and has therefore been broken down at Appendix I.

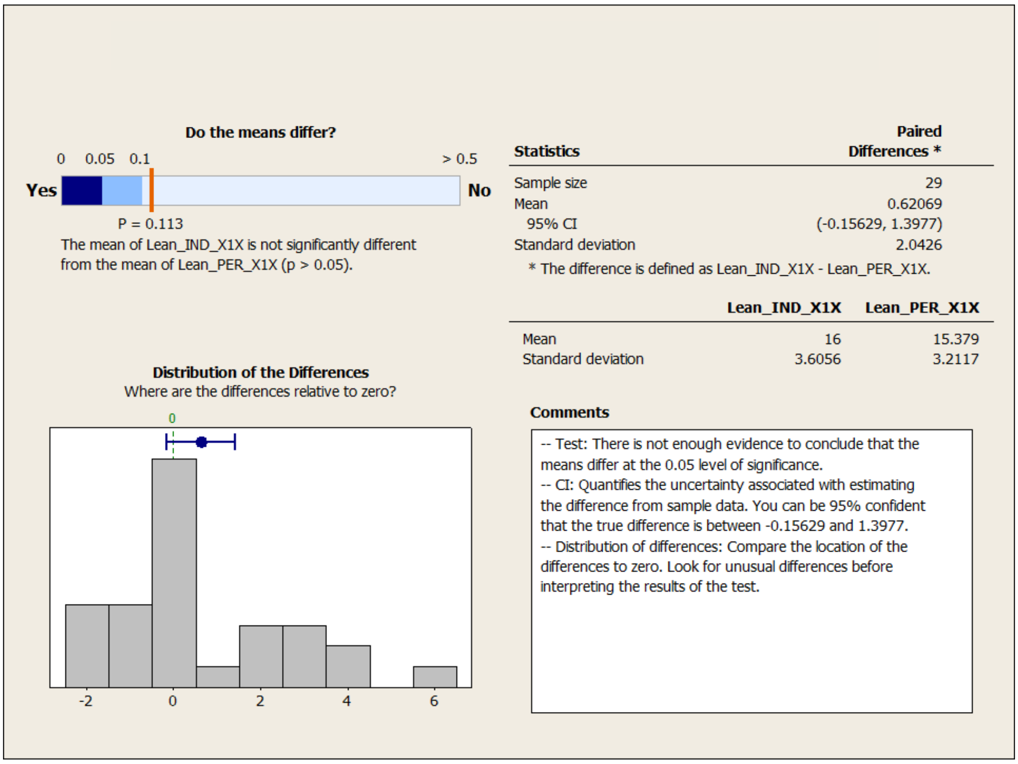

Figure 4-2. Summary of Paired-t Test at RAF Lossiemouth for Cultural Lean

The issue of Cultural Lean returned a subtle difference at RAF Lossiemouth (Figure 4-2). As with the test for Coningsby there were no unusual points detected and the Ryan-Joiner test returned a value of p=0.985. There was, however, *no* statistically-significant difference between the pilots' individual opinions of simulation ($\bar{x}$=16.0 $\boldsymbol{\sigma}_x$=3.61) and their perception of the whole Force's opinion ($\bar{x}$=15.379 $\boldsymbol{\sigma}_x$=3.21). The difference found being 0.62 (95% CI, -0.156 to 1.400), t(28) = 1.64, p = 0.113, d = 0.30. It should be noted that the Cultural Lean results for RAF Lossiemouth do not include Statement 1 of the questionnaire which was found to be inconsistent with the theme when applying Cronenbach's Alpha, as discussed above, thus a neutral opinion is found at the value of 15 rather than the 18 of Coningsby. Re-running the test however, with the statement included returned an even less-statistically-significant figure, thus the result would have remained unchanged.

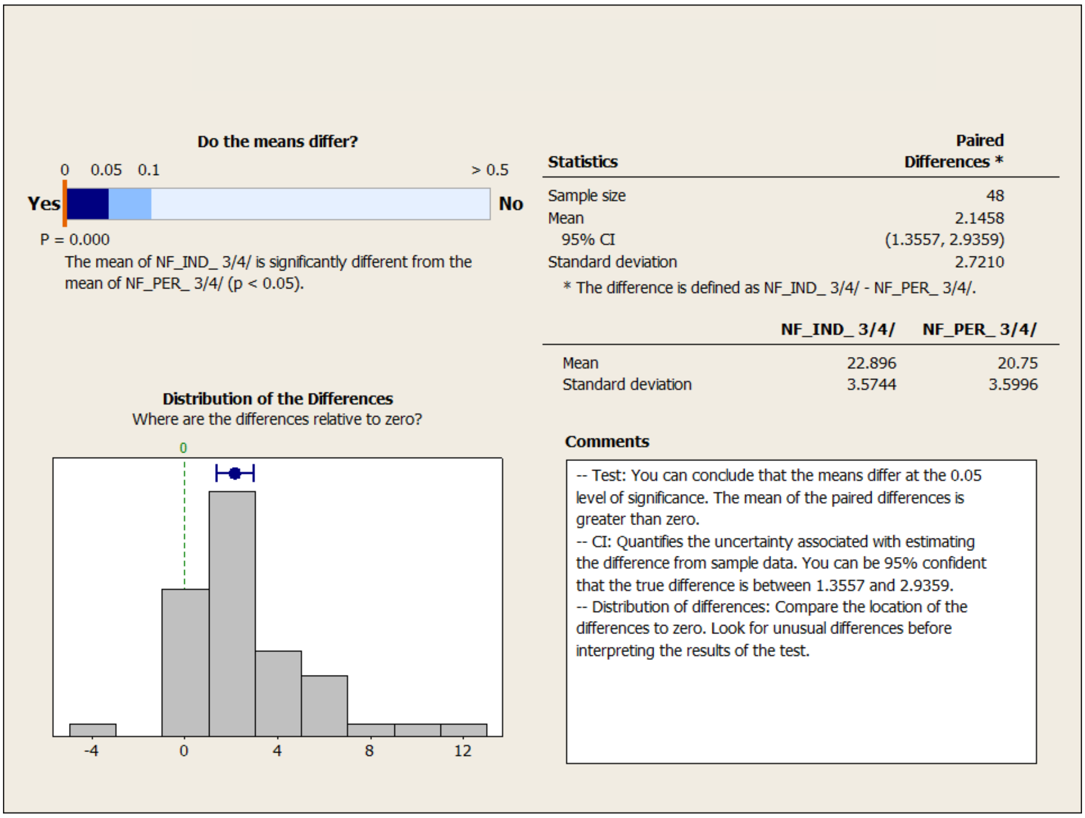

Figure 4-3. Summary of Paired-t Test at RAF Coningsby for 'Near Future'

The paired-t test for the theme of ‘Near Future’ was run for RAF Coningsby only, the theme having been found to be inconsistent at RAF Lossiemouth. There were two pairs with unusual differences, however inspection of their values did not reveal them to be extreme and they were kept in the analysis. As with the tests above the assumption of normality was not violated (Ryan-Joiner, p = 0.98). The individual opinion of the pilots was more positive ($\bar{x}$=22.896 σ_x=3.57) than the corresponding opinion of the Force ($\bar{x}$=20.750 σ_x=3.60), there was a statistically significant difference of 2.146 (95% CI, 1.356 to 2.936), t(47) = 5.46, p<0.001, d = 0.79.

The results show that for a value of alpha=0.05 both the categories of ‘Cultural Lean’ and ‘Near Future’ have a significant difference between an individual at RAF Coningsby and their perception of the rest of the Typhoon fleet’s opinion. In both cases the individuals considered themselves to be more positive about each than their colleagues collectively. At RAF Lossiemouth, however, the individuals considered themselves to be no different to the Force’s opinion in the topic of ‘Cultural lean’– no significant difference having been found. Reasons for this difference have not been definitively determined and could be due to the differing simulator types, methodology of use, simulator staff or a combination of these. What the research does show however is that when asking for feedback on the abilities of simulation in the past the answer has likely been affected by the airbase at which it has been asked and whether it has been asked of an individual or a group.

4.14 THE EFFECT OF QUALIFICATION

As pilots progress through their tour they gain qualifications that illustrate the areas in which they have a particular depth of knowledge. It was hypothesised that the qualification level of a pilot would have an affect on their answers to the themed questions based on what they needed simulation to provide for their particular area of expertise (Sub-Question 2, Objective 2). Respondents had been asked to indicate their qualification level in three ways. Firstly Combat Ready (CR) qualification level, the most basic being level 1 and the most advanced level 3. Secondly to state if they held the Electronic Warfare Instructor (EWI) qualification which is given, via a 3-week course, to promising pilots towards the end of their first tour. Finally Type of Pilot – Line (L), Qualified Pilot Instructor (P) or Qualified Weapons Instructor (W); Line pilot being the standard pilot, Pilot Instructor being typically a second tourist with 2-5 years experience on the aircraft and has completed the Central Flying School’s Pilot Instructor course, and finally Weapons Instructor who represent the top 5% of pilots in terms of ability, the qualification being awarded after a year-long challenging course.

A one-way ANOVA (see Appendix I) was used for the CR qualification and Type of Pilot variables and a 2-sample t-test for the EWI qualification (as it was a dichotomous variable) against the themes of ‘Cultural Lean’ and ‘Near Future’ in order to determine if there was an effect due to experience, based on qualification level. As a positive result in this area was likely to be immediately exploited by the RAF the alpha level was set at 0.05. The hypotheses being tested were:

H_0: all group population means are equal.

H_A: at least one group population mean is different.

4.14.1 QUALIFICATION VS 'CULTURAL LEAN' AT RAF CONINGSBY

Examination of the 'CR Qual' variable found there to be no outliers, as assessed by a boxplot (see Appendix I). The data were also normally distributed for each group, as assessed by a Ryan-Joiner test ($p>0.10$) and there was a homogeneity of variances, as assessed by Levene's test (Appendix I) for quality of variances ($p = 0.786$). The effect of the Combat Ready Qualification on the theme of 'Cultural Lean' was not statistically different, $F(2, 45) = 2.80$, $p = 0.07$. The theme's score for each CR Qual being; Level One 19.86 +/- 3.27, Level Two 22.23 +/- 3.67 and Level Three 19.57 +/- 4.36, where the data are presented as mean +/- standard deviation.

Testing the required assumptions for the variable 'Type of Pilot' returned no outliers (boxplot), a normal distribution (Ryan-Joiner $p > 0.1$) and a homogeneity of variances (Levene's test $p = 0.52$) across the categories. The 'Type of Pilot' variable was not found to have a statistically-significant effect on the theme of 'Cultural Lean', $F(2, 45) = 0.73$, $p = .49$. The return for each qualification being Line Pilot 21.15 +/-4.34, Qualified Pilot Instructor 20.00 +/- 3.67 and Qualified Weapons Instructor 19.50 +/- 2.45.

Determination of whether the EWI qualification was a variable that affected the 'Cultural Lean' theme was investigated through the use of a 2-sample t-test. The initial search for outliers highlighted a single data point of concern in the grouping without the qualification (Figure 4-4). Upon investigation it was found that the point was neither a data entry or measurement error but a genuinely unusual value. Consideration was given to removal of the data point, however it was elected to modify the data by replacing it with 1 less than the next largest value (Laerd, 2015b). Both Yes and No groupings were found to be normally distributed, p > 0.1. Levene's test with the modified outlier returned a p-value of 0.73.

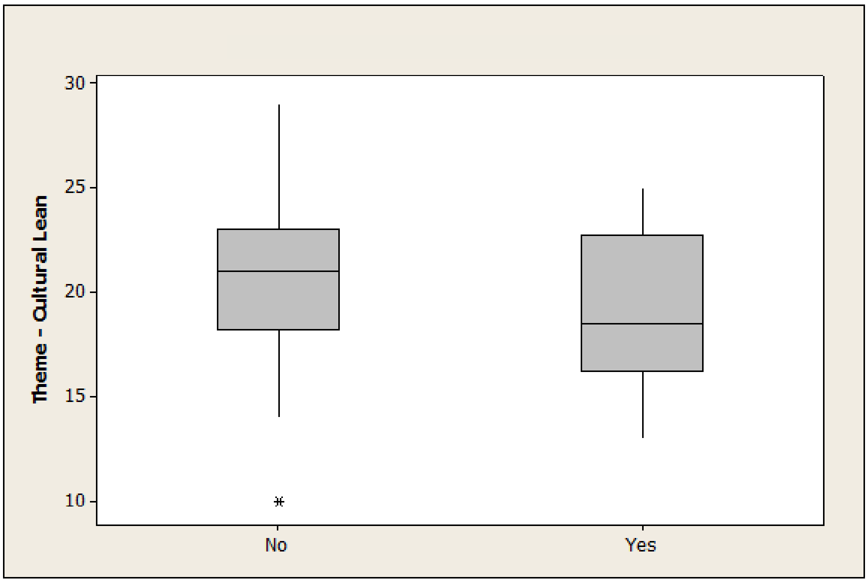

Figure 4-4. Outlier Determination for EWI Qualification

The result was not statistically significant with those with an EWI qualification scoring a 'Cultural Lean' mean value of 19.17 +/- 3.66 and those without 21.11 +/- 3.70. The difference being 1.94 (95% CI, -0.62 to 4.51), t (19) = 1.59, p = 0.13.

4.14.2 <u>QUALIFICATION VS 'NEAR FUTURE' AT RAF CONINGSBY</u>

The affect of qualification on Near Future was then sought, using the same methodology as that described above for each of the three experience variables, the results being tabulated in Table 4-1.

.

Table 4-1. Results of Experience on the Theme of ' Near Future'

Variable	Outliers Y/N (Boxplot)	Normality (Ryan-Joiner)	Levene's Test	Test conducted	Results 1	Results 2 (Mean +/- Std Dev)	Remarks	Hypothesis Upheld
CR Qual	N	P > 0.1	P = 0.157	One Way ANOVA	F(2,45) = 2.28 p = 0.115	Level 1: 22.31 +/- 3.11 Level 2: 24.28 +/- 2.89 Level 3: 21.93 +/- 4.27		$\mathbf{H_0}$
Type of Pilot	Y	P > 0.1	P = 0.208	One Way ANOVA	F(2,45) = 2.78 p = 0.073	Line Pilot: 23.70 +/-2.84 Qualified Pilot Instructor: 22.85 +/-3.7 Qualified Weapons Instructor: 20.50 +/- 4.472	Outlier modified	$\mathbf{H_0}$
EWI Qual	N	P > 0.1	P = 0.01 FAIL				Non equal variances	Non determined

4.14.3 THE QUALIFICATION VARIABLE AT RAF LOSSIEMOUTH

It was expected that comparison of qualification levels against themes at RAF Lossiemouth would be complicated by the homogeneous nature of the qualification sets present at the base. For example, of the 29 pilots questioned 22 were of the Line Pilot type and of the remainder there was only one QPI. This was expected as figures on the exact breakdown had not been collected before and the base is pure front-line rather than a mix of training and operations.

Only the CR qualification provided the barest minimum numbers of each grouping to be tested: Level 1, $n = 7$; level 2, $n = 8$; and level 3, $n = 14$. The boxplot returned no outliers, normality, using the Ryan – Joiner test, was found to be $p > 0.1$ in each case and Levene's test showed the variances to be homogeneous. However, no statistically-significant differences between the CR levels was found: $F(2, 28) = 0.28$, $p = 0.76$. The theme's score for each CR Qual was; Level 1 15.29 +/- 4.11; Level 2 15.75 +/- 2.77; and Level 3 16.5 +/- 3.94.

4.14.4 SUMMARY

Despite the general theory that acceptance of simulation may be related to the qualification level of the pilot no evidence to support this proposition was found at RAF Coningsby or at RAF Lossiemouth.

4.15 THE EFFECT OF THE SQUADRON

Whilst the Squadron was recorded on the questionnaire this was originally done for collation of records rather than a method of grouping results. No consideration of the Squadron as a group was originally considered as the perceived wisdom was that it was qualification that was the driver of acceptance of simulation training. This was logical at first glance; all pilots had received identical training prior to arrival on their squadron that contained a constant level of exposure to synthetic training, the squadrons had a near-identical-make up of qualifications and lastly they underwent a common, mandatory number of synthetic hours of training per month. During the course of investigating the results a pattern began to emerge that suggested that the squadron a pilot was assigned to did indeed have an effect on their perception of simulation. To ascertain the validity of this assertion the effects on the themes 'Cultural Lean' and 'Near Future' were sought with respect to each of the squadrons at RAF Coningsby: 3(F) Squadron (n=19), XI Squadron (n=17) and 29(R) Squadron (n=12).

A boxplot for each of the squadrons found a single outlier on 3(F), see Figure 4-5. In common with the outlier found when testing against the variable of qualification (see section 4.14) the outlier value was reduced to 1 less than the next data point. All groups passed the Ryan-Joiner normality check and Levene's test returned a value of p = 0.88.

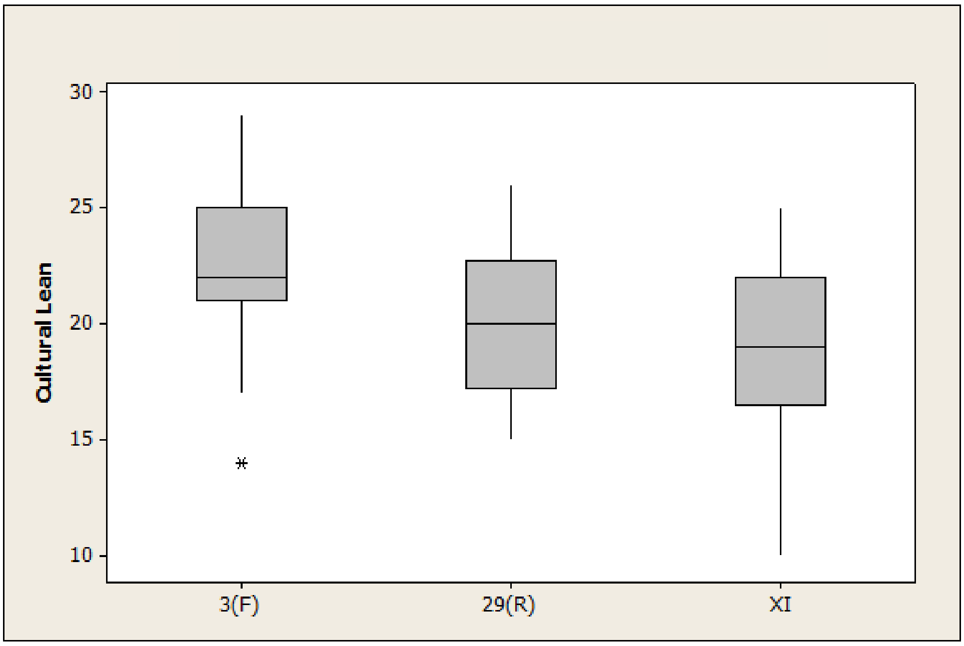

Figure 4-5. Boxplot of Squadron v Cultural Lean

A statistically-significant difference between the squadrons was found, $F(2, 45) = 4.71$, $p = 0.01$. The theme's score for each squadron being: 3(F) Squadron 22.47 +/- 3.42; XI Squadron 18.89 +/- 3.92; and 29 (R) Squadron

20.08 +/- 3.26. Thus, it can be said that the acceptance of simulation varied by the squadron, with 3(F) being the most positive.

The analysis was continued for the theme of 'Near Future' with the boxplot again finding a single outlier (Figure 4-6), which was modified in the same manner as above. Normality was found in all groups and variances were homogeneous (Levene's p = 0.56). Again the result was found to be statistically-significant; F(2,45) = 6.85, p = 0.003; 3(F) Squadron 24.68 +/- 2.52; XI Squadron 20.88 +/- 3.43; and 29 (R) Squadron 23.25 +/- 3.42.

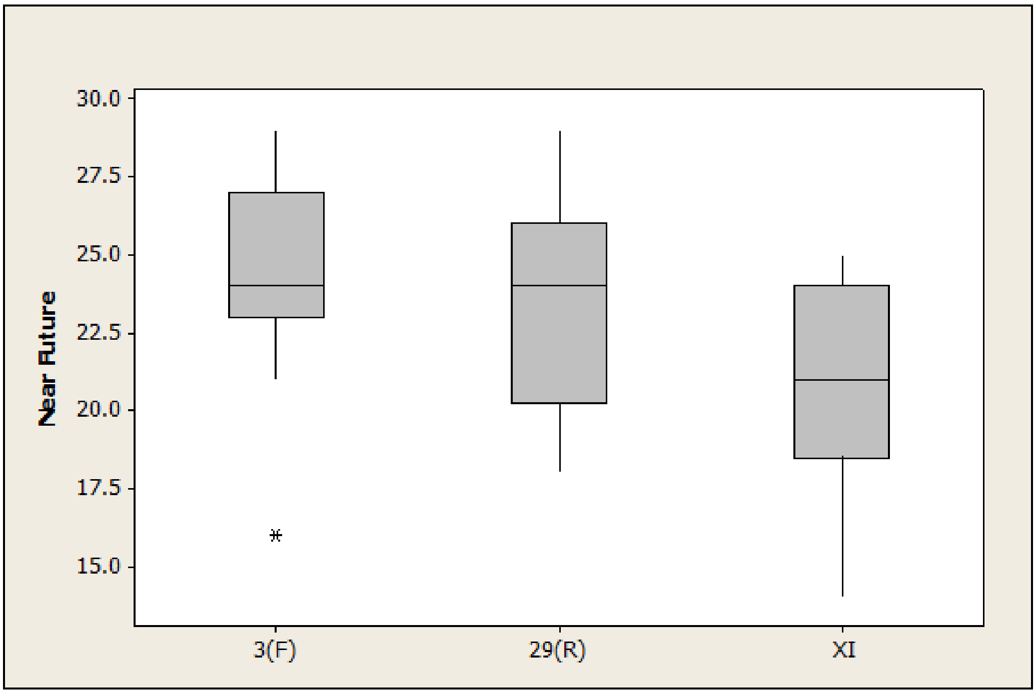

Figure 4-6. Boxplot of Squadron v Near Future

In order to determine if this was common at both bases the RAF Lossiemouth Squadrons were also examined, against the theme of 'Cultural Lean' only. Figure 4-7 shows the boxplot for 6 Squadron (n = 13) and 1 Squadron (n= 16), two outliers being present in the 1 Sqn results. Both outliers were modified in the manner previously identified, whilst this may be controversial for both outliers the 2-sample t-test would be rerun without a change in the overall result. Normality and variance homogeneity were satisfied. The result, however, was not statistically significant: 1 Squadron scoring 17 +/- 2.71 and 6 Squadron 15.15 +/- 3.83. The difference being 1.85 (95% CI, -0.78 to 4.47), t (19) = 1.47, p = 0.16.

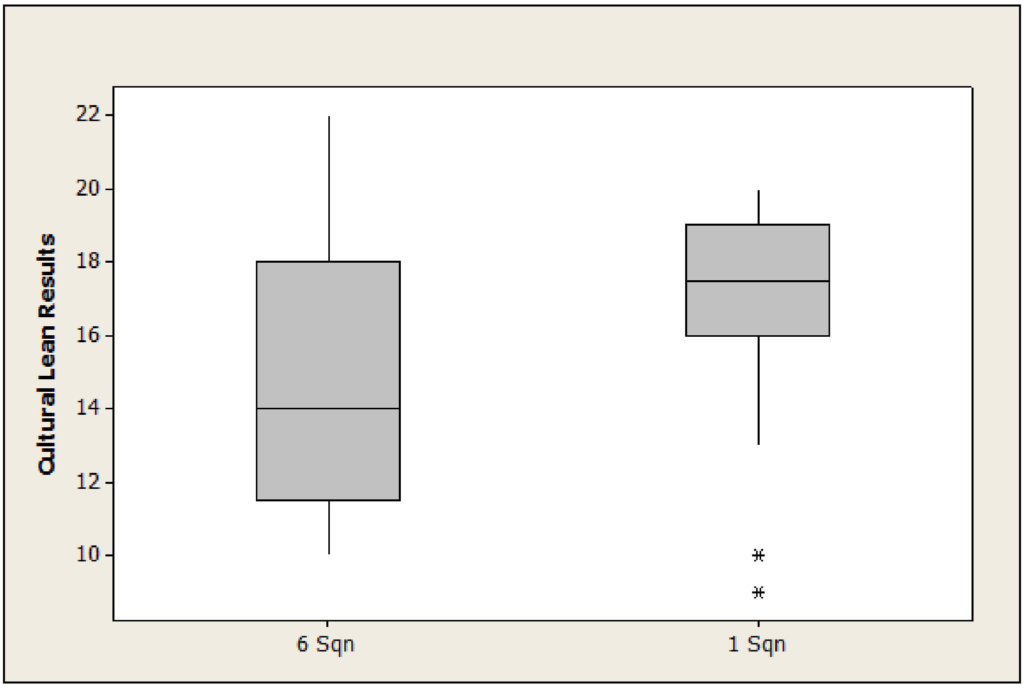

Figure 4-7. Boxplot of Lossiemouth Squadrons v Cultural Lean

4.15.1 SUMMARY

The effect of Squadron on the acceptance of simulation at RAF Coningsby was unexpected and thus the results were rerun leaving the outlier in position (as this would product the most unfavourable result). The overall results remained un-altered, however, and thus it can be said that the acceptance of simulation differs depending on which front-line squadron the pilots occupied. This result is unexpected as of all the squadrons at RAF Coningsby it is the two front-line squadrons (3(F) and XI) that at first glance would be considered the most similar, undergoing very similar training and numbers of hours in the simulator. In contrast 29(R) Squadron pilots, who spend longer periods training at a lower level of complexity were found to sit between the front-line squadrons in terms of results.

4.16 REGRESSION OF THE THEMES

Regression was used to determine if there were any factors, such as age or experience, that correlated to the level of acceptance (Objective 2). Given the discovery that the results for the themes were grouped by squadron meant that the regression of the themes would also be grouped by this variable. The continuous variables explored were:

a. Age.

b. Number of Typhoon hours.

c. Total flying hours (differs from Typhoon flying hours in that it includes all aircraft types an individual has flown).

d. Average number of 2-ship simulators flown in a 2-month period.

e. Average number of 4-ship simulators flown in a 4-month period.

During this section Objective 3i will also be investigated to determine if the commonly held belief that younger pilots are more accepting of simulation is in fact true. As this belief has come to the fore within 2014 the evidence for its genesis must be present in the personnel presently on the fleet during that period. As 100% on the front-line squadrons responded to the initial questionnaire the validity of this belief must be found within their responses. As such, to answer Objective 3i, comparison to a wider population is not required.

Pearson's product moment correlation (see Appendix I) was used to assess the relationship between the theme of 'Cultural Lean' and 'Near Future' and each of the chosen variables, the hypotheses tested are shown below. Given the exploratory nature of the work, the lack of historical precedence and the non-flight-safety nature of the results a significance level (α) was set at 10%.

H_0 : ρ=0; the population coefficient of the chosen variable is equal to zero.

H_A : ρ≠0; the population coefficient of the chosen variable is not equal to zero.

4.16.1 AGE AND 'CULTURAL LEAN' AT RAF CONINGSBY

Upon examination of the scatterplot for 29 Squadron (Figure 4-8) a single point stood out as an outlier, this being the single 53-year-old Typhoon pilot, the next nearest age being 39. After deliberation this point was removed from the analysis as, although it was a valid data point, there was not the density of data around this point to provide a weight of evidence for the age or that surrounding it. There were no outliers for XI and 3(F) squadrons.

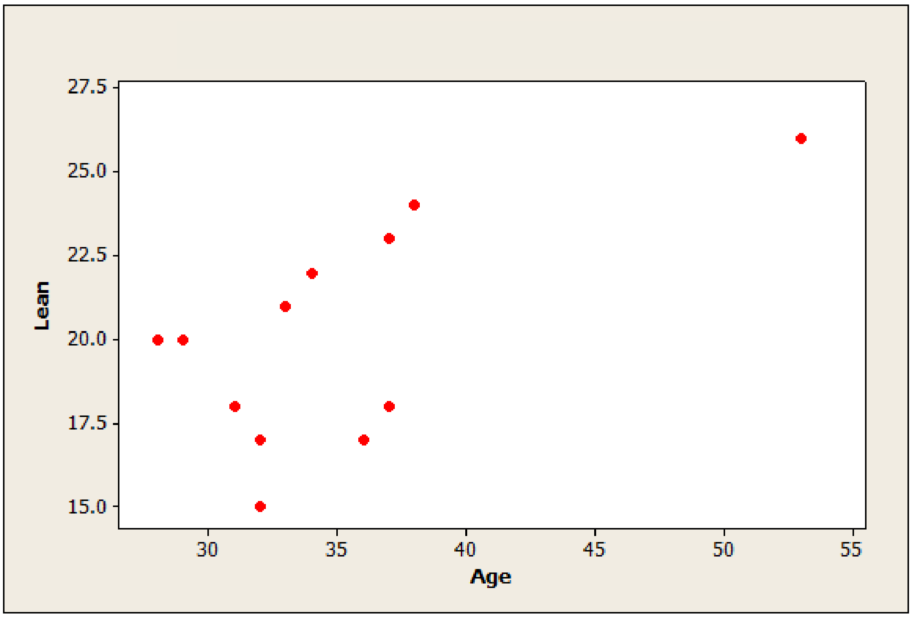

Figure 4-8. Scatterplot of Cultural Lean vs Age for 29(R) Squadron

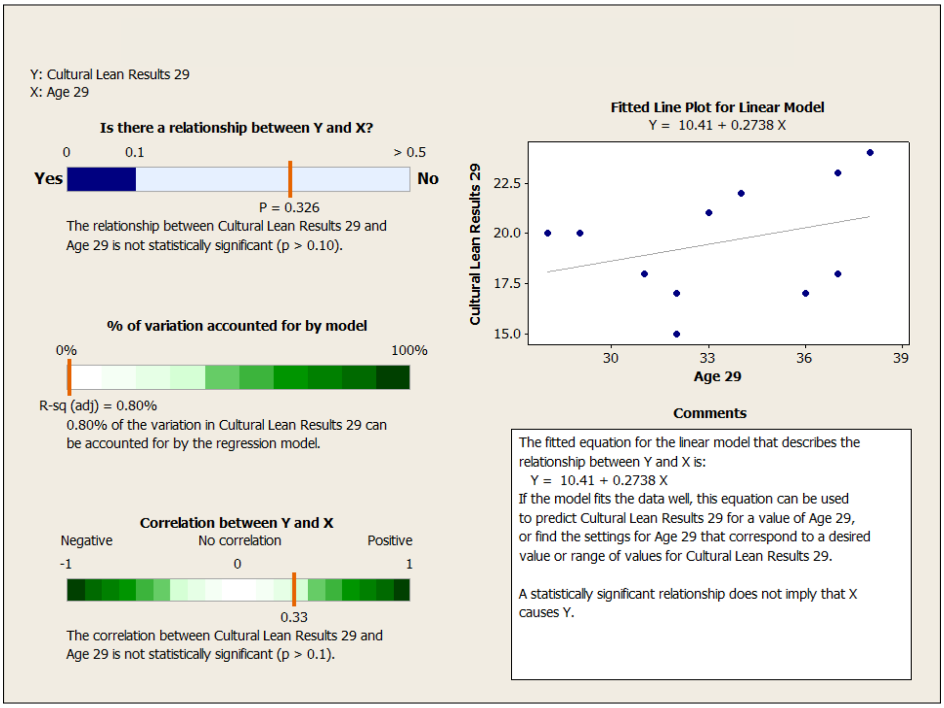

Figure 4-9. Summary of Cultural Lean v Age for 29(R) Squadron

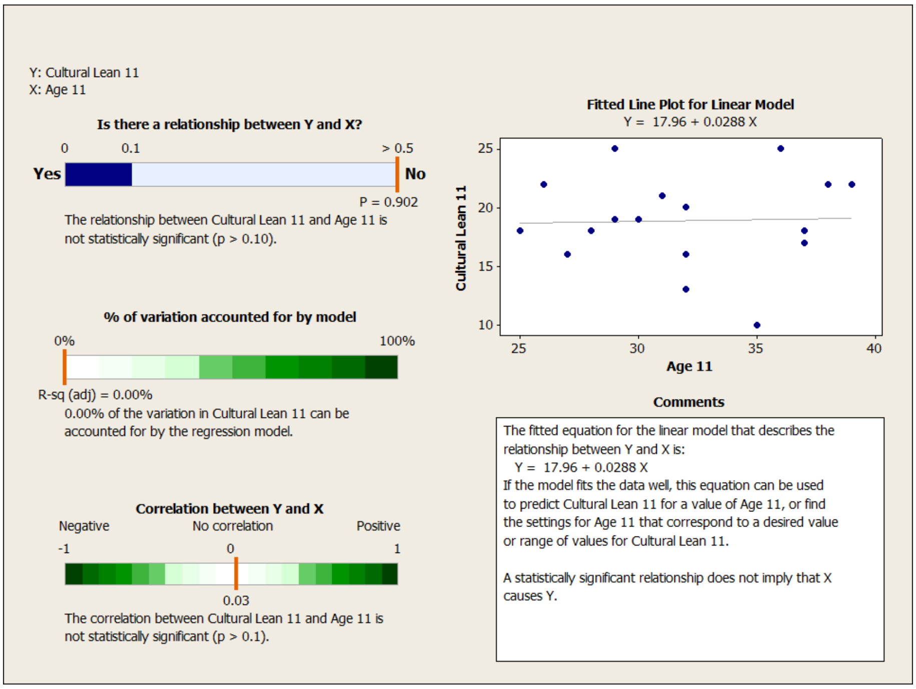

Figure 4-10. Summary of Cultural Lean v Age for XI Squadron

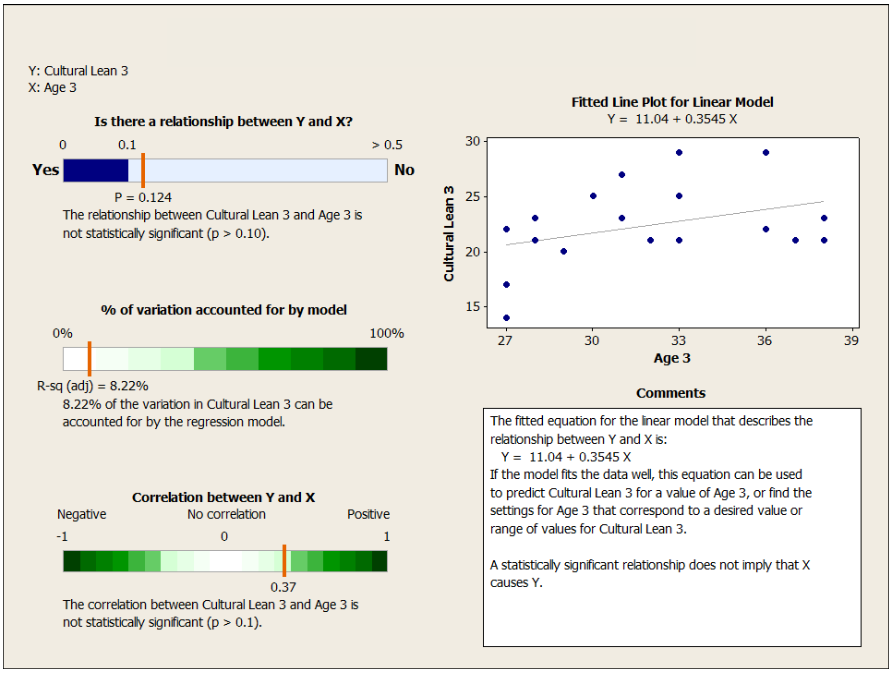

Figure 4-11. Summary of Cultural Lean v Age for 3(F) Squadron

The results for the squadrons can be seen in Figure 4-9, Figure 4-10 and Figure 4-11 and show positive values of correlation for 29(R) and 3(F) squadron of 0.33 and 0.37 respectively, although these values were not shown to be statistically significant. XI Squadron did have a single data point with a high residual but no reason could be found to remove it from the analysis, additionally a re-run of the data with this point missing did not yield a significant result. None of the squadrons returned a statistically significant result, thus the hypothesis upheld is **H_0**. However, it can be seen that the correlations for these particular samples is not the expected negative one if the belief that younger pilots were more accepting of simulation (Objective 3i) was to be supported.

4.16.2 TOTAL TYPHOON HOURS AND 'CULTURAL LEAN' AT RAF CONINGSBY

Both 29(R) and XI Squadron's scatterplots each contained a single outlier that had flown a total of 1600 and 1100 hours respectively on the Typhoon (Figure 4-12 and Figure 4-13). These data points were removed as there was not a sufficient density of points around these numbers of hours to allow realistic regression to these values. This logic mirrors that used for the age variable regression on 29(R) Squadron.

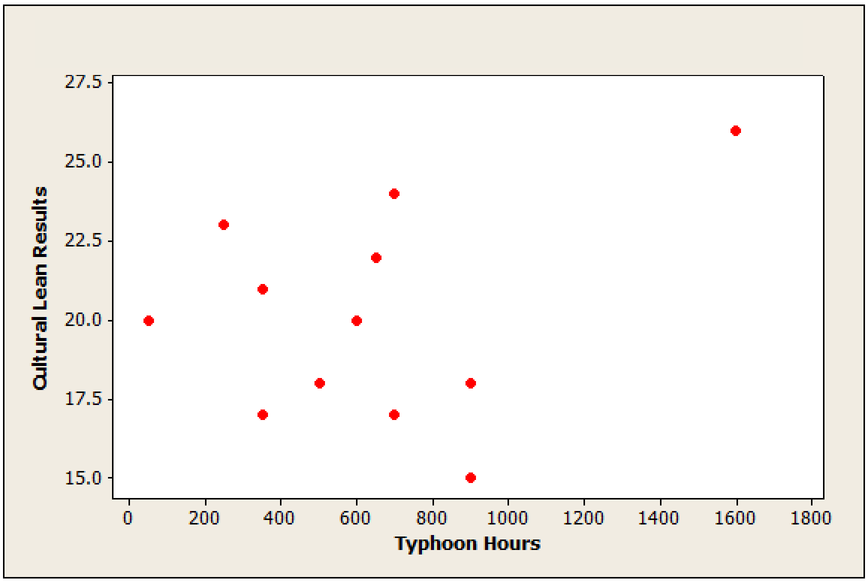

Figure 4-12. Scatterplot of Cultural Lean vs Typhoon Hours for 29(R) Squadron

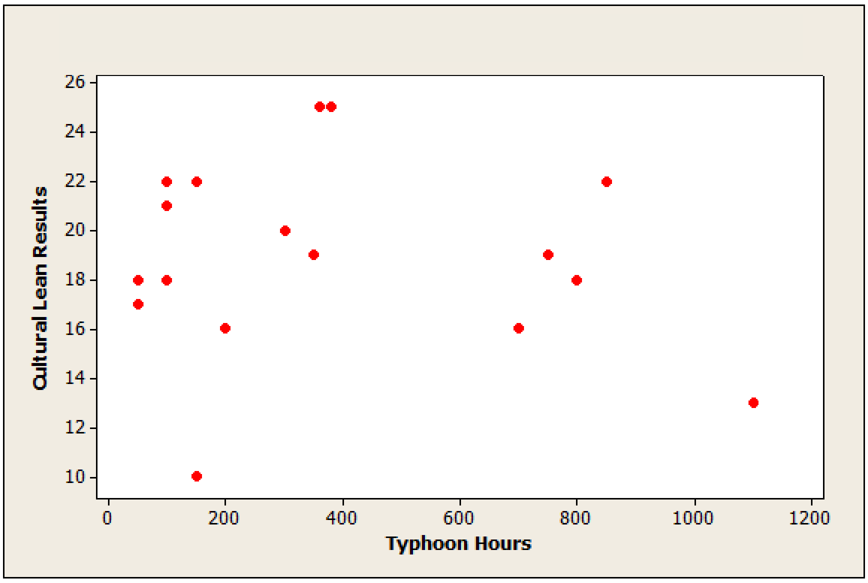

Figure 4-13. Scatterplot of Cultural Lean vs Typhoon Hours for XI Squadron

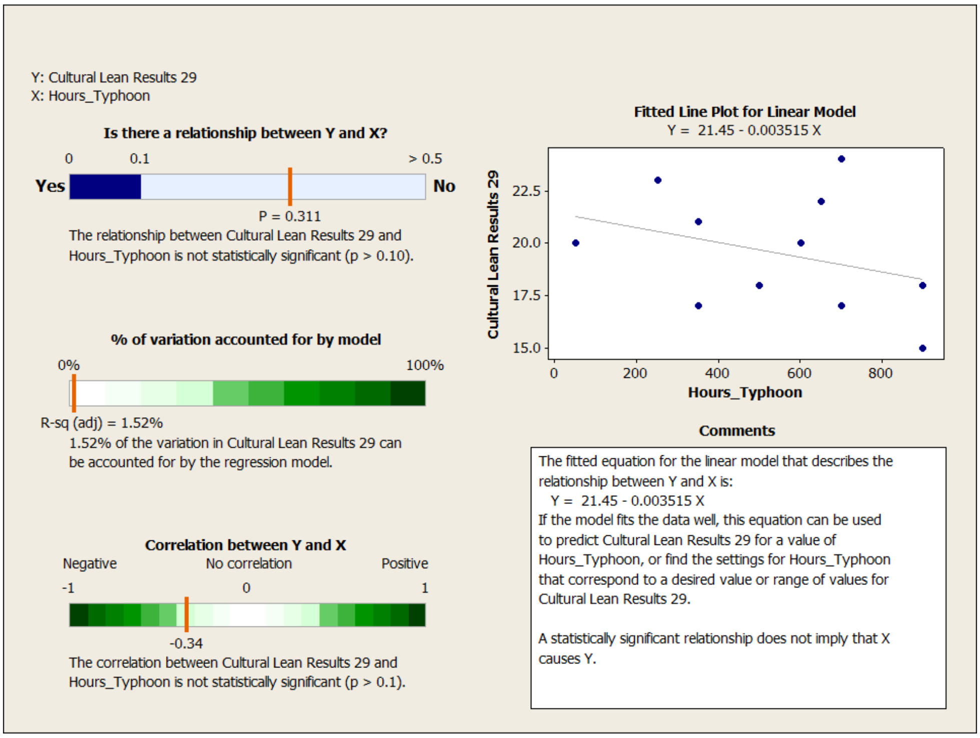

Figure 4-14. Summary Report for Regression of Cultural Lean vs Typhoon Hours for 29(R) Squadron

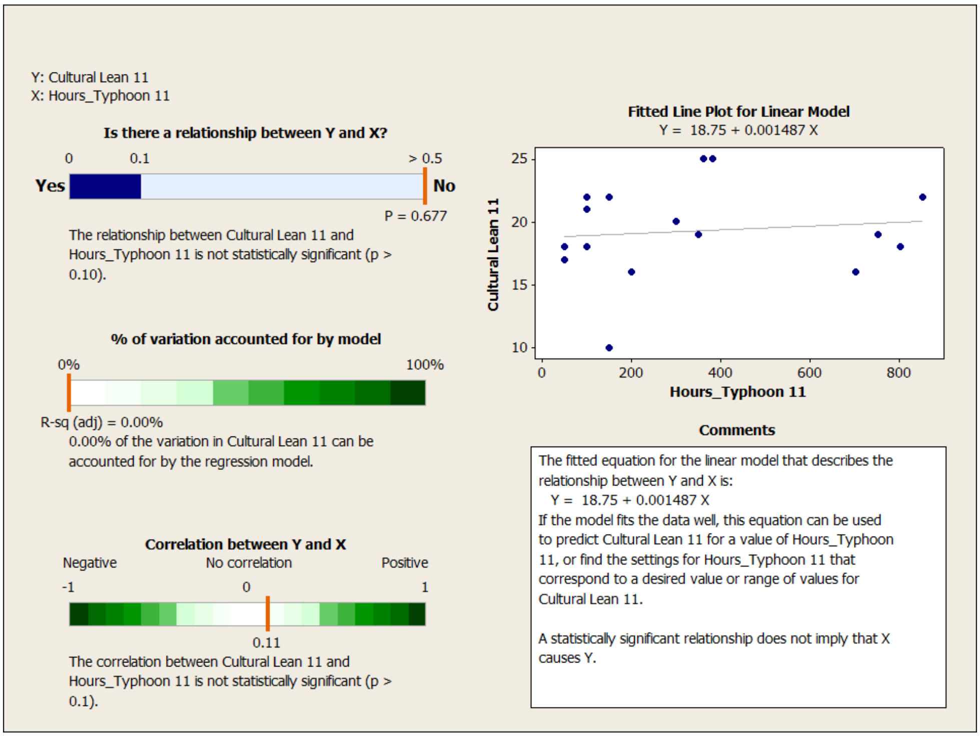

Figure 4-15. Summary Report for Regression of Cultural Lean vs Typhoon Hours for XI Squadron

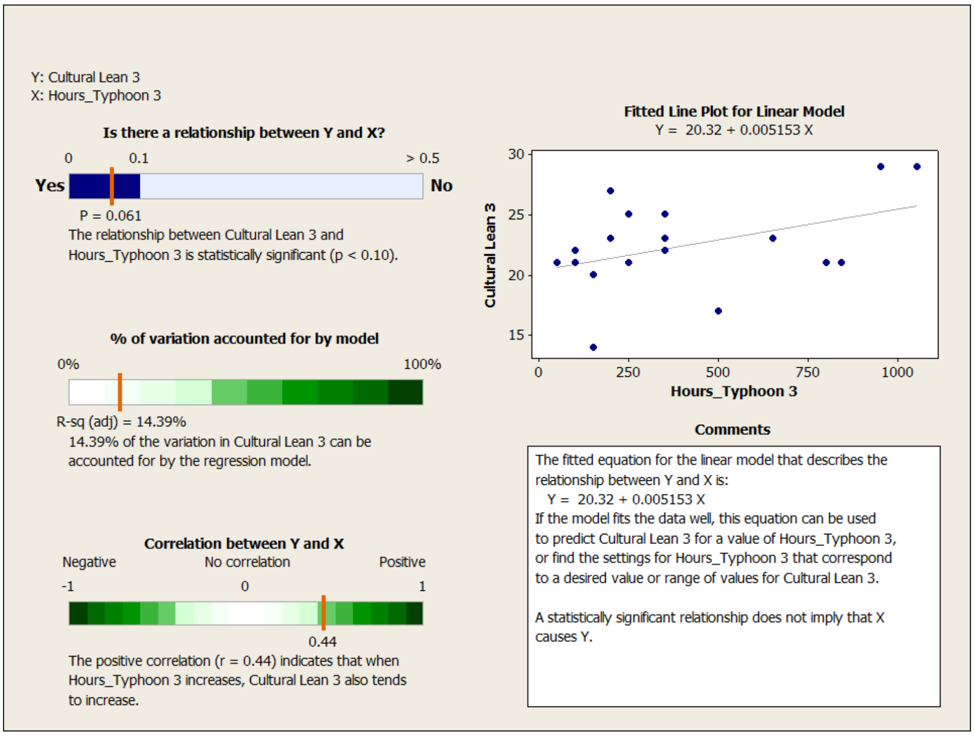

Figure 4-16. Summary Report for Regression of Cultural Lean vs Typhoon Hours for 3(F) Squadron

An inconsistent picture of the correlation coefficient was presented across the squadrons, both front-line squadrons returning a positive correlation coefficient whilst the training unit, 29(R), was negative. Only 3(F) produced a statistically-significant result allowing the alternative hypothesis **H_A** to be upheld for this squadron alone.

4.16.3 TOTAL FLYING HOURS AND 'CULTURAL LEAN' AT RAF CONINGSBY

A single outlier was found on 29(R) Squadron (see Figure 4-17). This individual with 6000 hours was a particular anomaly in the fleet and as such had no peers. His result was removed in line with the outlier argument advanced previously.

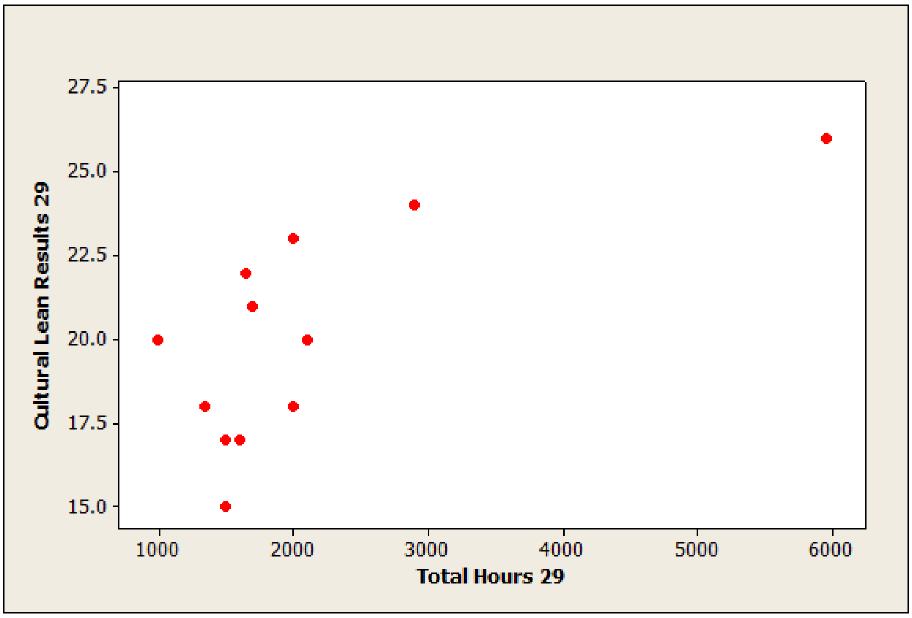

Figure 4-17. Scatterplot for Cultural Lean vs Total Hours for 29(R) Squadron

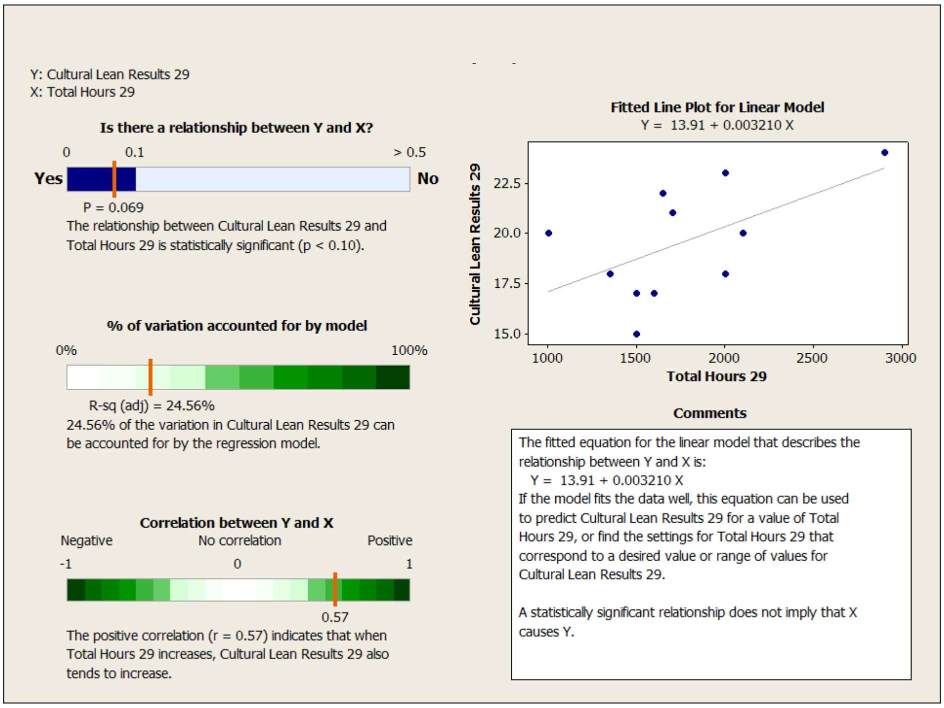

Figure 4-18. Summary Report for Regression of Cultural Lean vs Total Hours for 29(R) Squadron

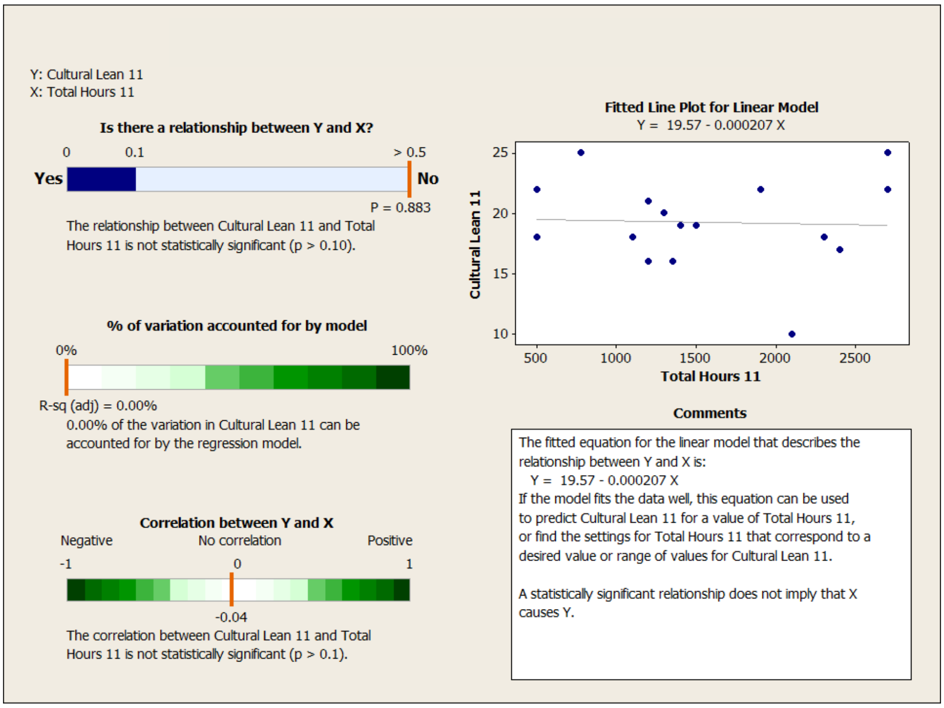

Figure 4-19. Summary Report for Regression of Cultural Lean vs Total Hours for XI Squadron

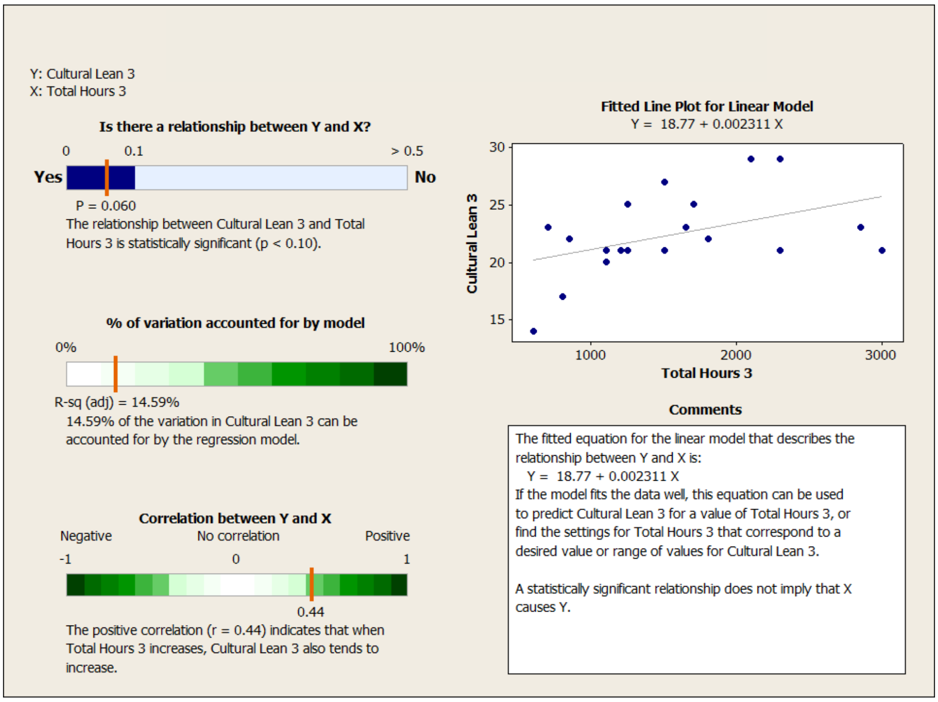

Figure 4-20. Summary Report for Regression of Cultural Lean vs Total Hours for 3(F) Squadron

Both 29(R) and 3(F) squadrons gave statistically-significant results with correlations of medium strength, thus supporting the alternate hypothesis. In contrast XI squadron showed virtually no correlation at all.

4.16.4 NUMBER OF 2-SHIP SIMULATORS FLOWN AND 'CULTURAL LEAN' AT RAF CONINGSBY

Outliers were discovered in both 29(R) and 3(F)'s data and were linked to a scarcity of individuals who had conducted a large number of 2-ship simulators. Thus in order not to over-extrapolate, the single data point at 10 simulator sorties for 29(R) (Figure 4-21) and the data points at 3, 4 and 6 simulator sorties for 3(F) (Figure 4-22) were removed, consistent with the previous outlier methodology.

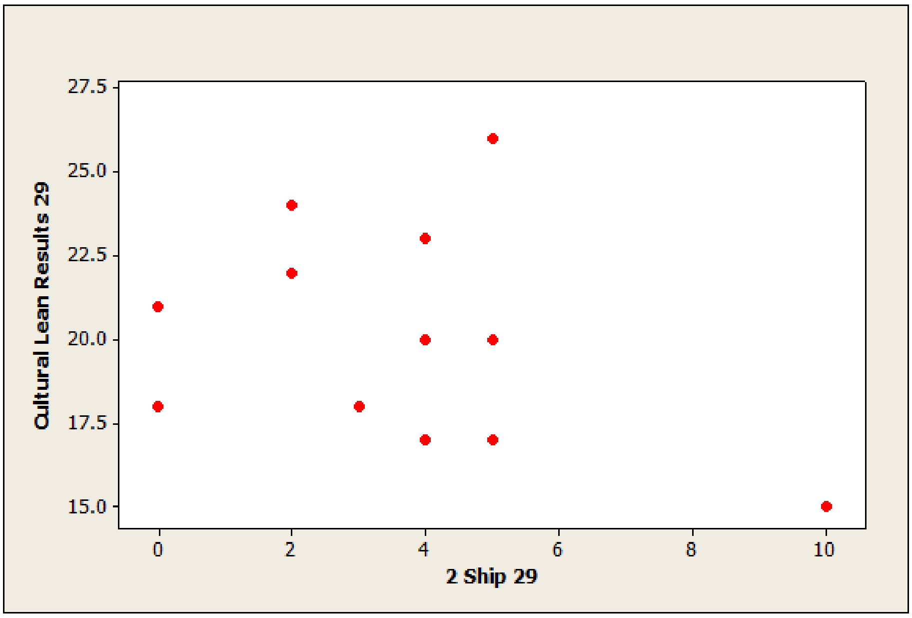

Figure 4-21. Scatterplot for Cultural Lean vs 2 Ship Sims for 29(R) Squadron

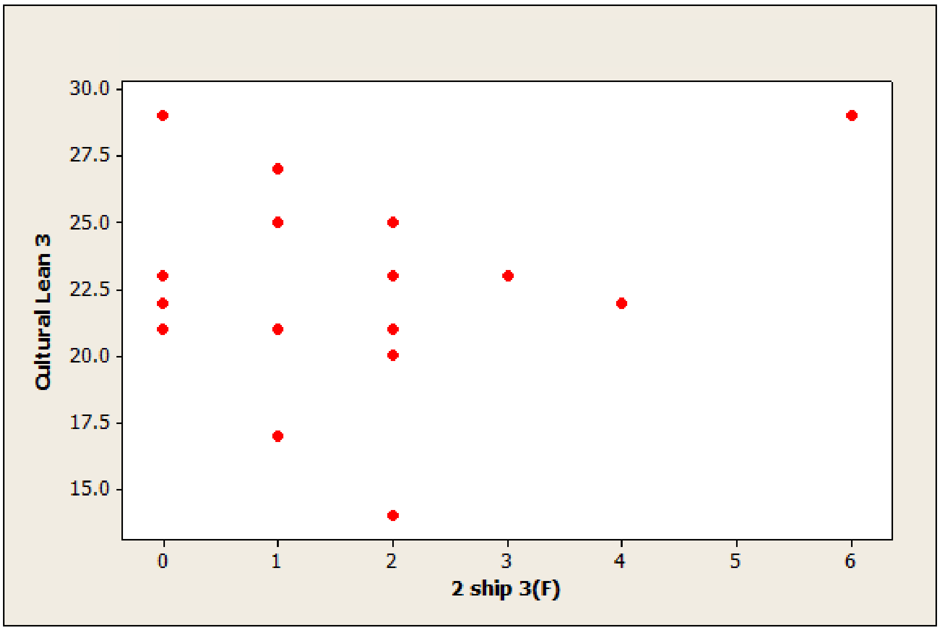

Figure 4-22. Scatterplot of Cultural Lean vs 2 Ship Sims for 3(F) Squadron

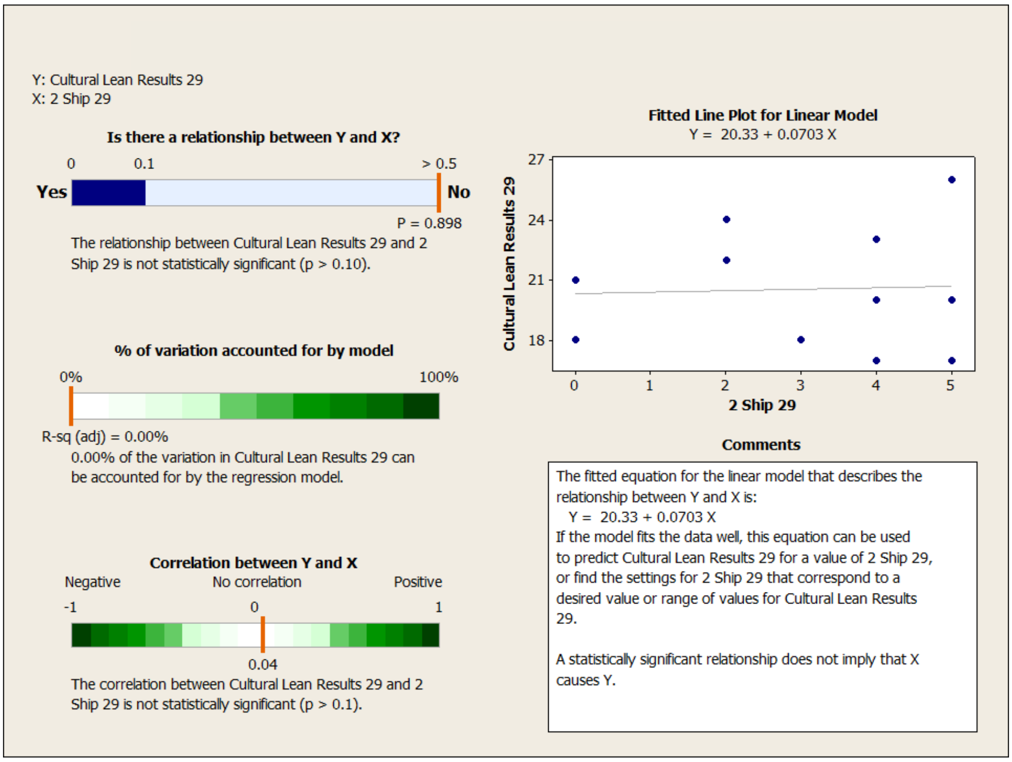

Figure 4-23. Summary Report for Cultural Lean vs 2 Ship Sims for 29(R) Squadron

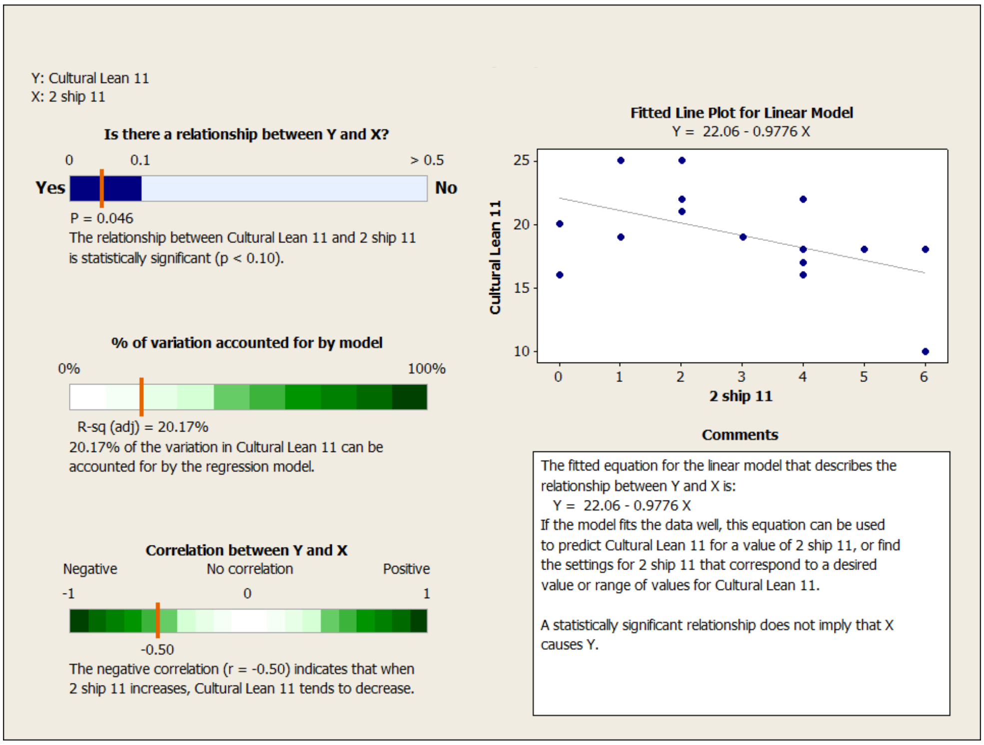

Figure 4-24. Summary Report for Cultural Lean vs 2 Ship Sims for XI Squadron

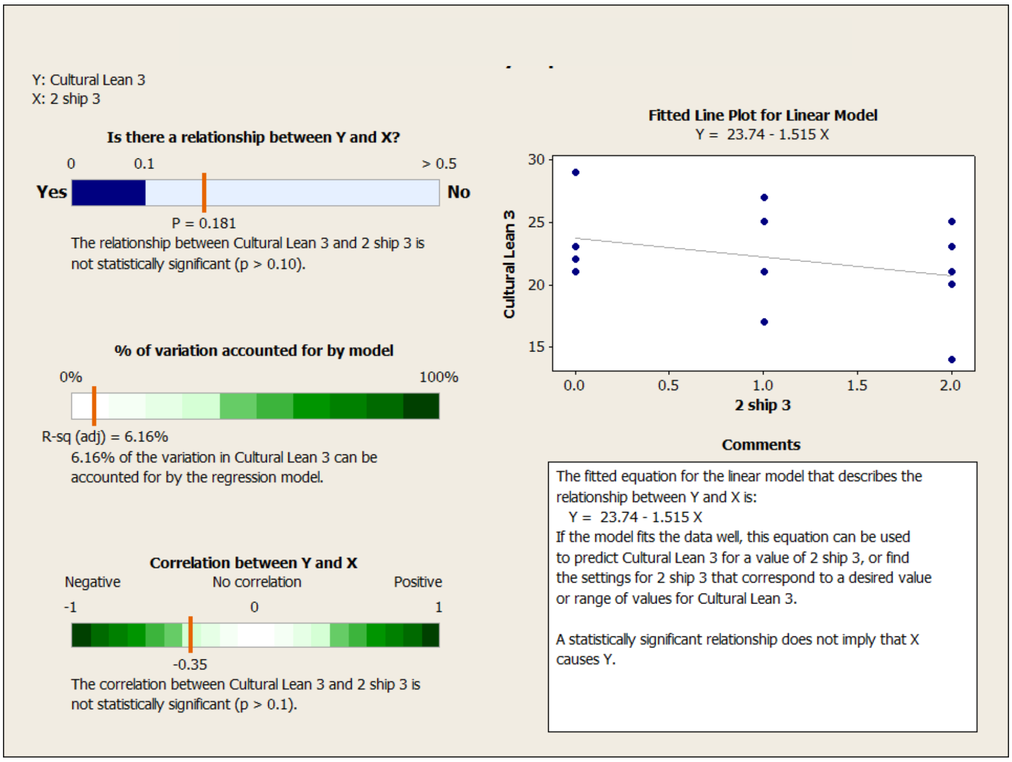

Figure 4-25. Summary of Cultural Lean vs 2 Ship Sims for 3(F) Squadron

Analysis of both front-line squadrons found that increasing 2-ship simulator sorties had a negative effect on the theme of ‘Cultural Lean’ with XI returning a result significant enough to reject the null hypothesis. Of more interest is that the XI Squadron results would be better fitted by a square law; up to two 2-ship simulators per 2-month period having a positive effect on ‘Cultural Lean’, in contrast to their sister squadron 3(F) Squadron. This is particularly important as it may hint to why XI Squadron’s overall ‘Cultural Lean’ figures are below those of their peers. Also highlighted was that XI Squadron would appear to conduct more work in simulation than the other squadrons at RAF Coningsby, their 17 pilots believing they had conducted 46 simulator sorties in the last 2 months, an average of 2.7 per pilot. 3(F)’s 19 pilots by comparison had conducted 32 sorties, an average of 1.7 each.

4.16.5 NUMBER OF 4-SHIP SIMULATORS AND ‘CULTURAL LEAN’ AT RAF CONINGSBY

Inspection of the scatterplots found 3 points for removal on the XI Squadron data. The data points of 2, 4 and 6 four-ship simulator sorties per 2 month period (Figure 4-26) were removed to ensure a better data density. Correspondingly the data for XI Squadron will only be valid up to a total of 1 four-ship simulator per 2 months.

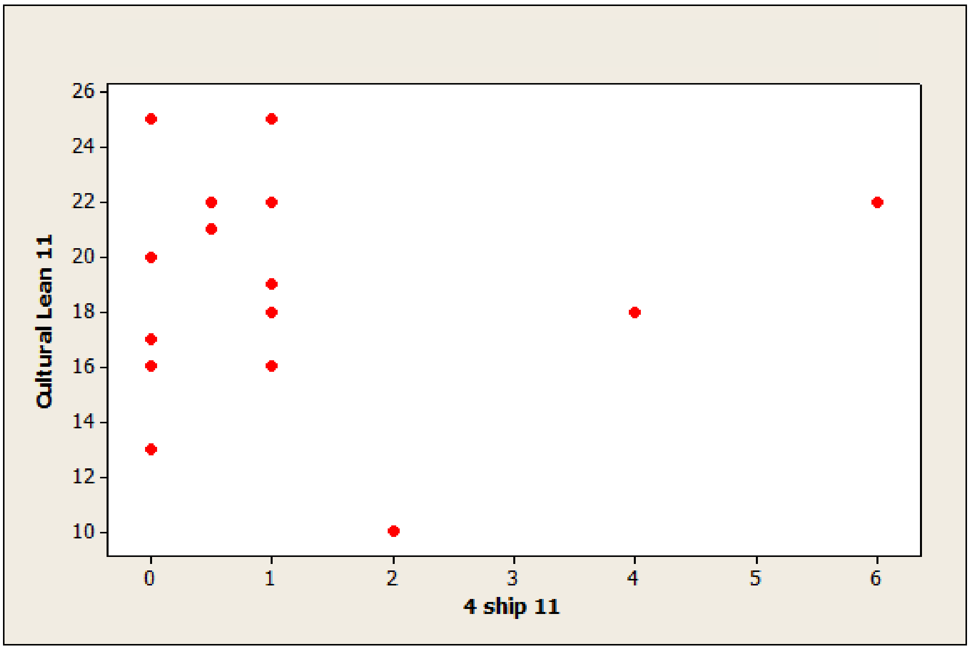

Figure 4-26. Scatterplot of Cultural Lean vs 4 Ship Sims for XI Squadron

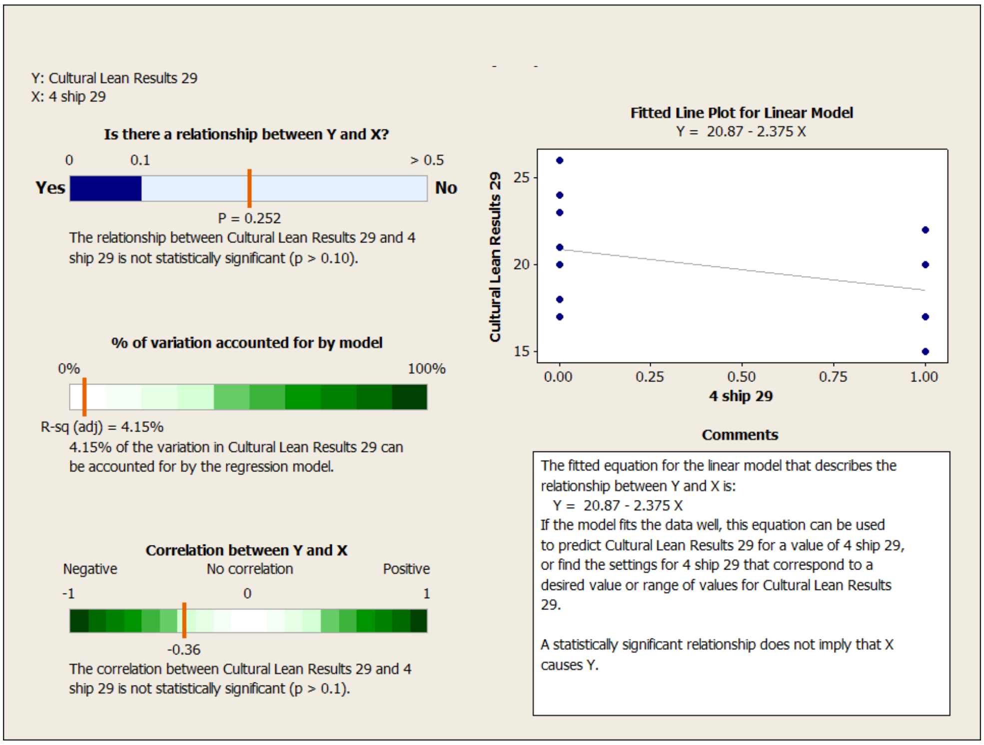

Figure 4-27. Summary Report for Cultural Lean vs 4 Ship Sims for 29(R) Squadron

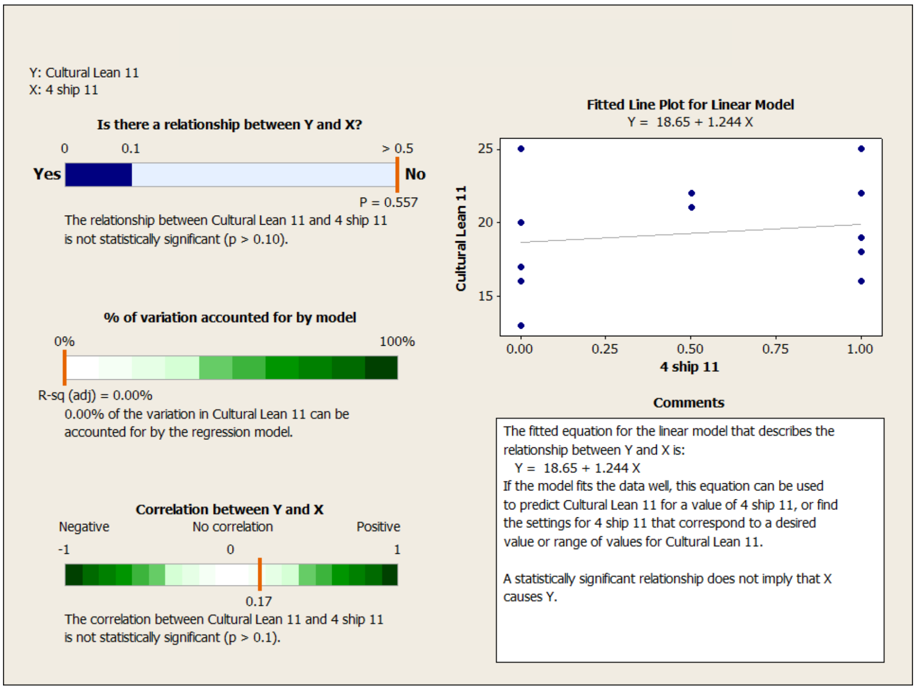

Figure 4-28. Summary Report for Cultural Lean vs 4 Ship Sims for XI Squadron

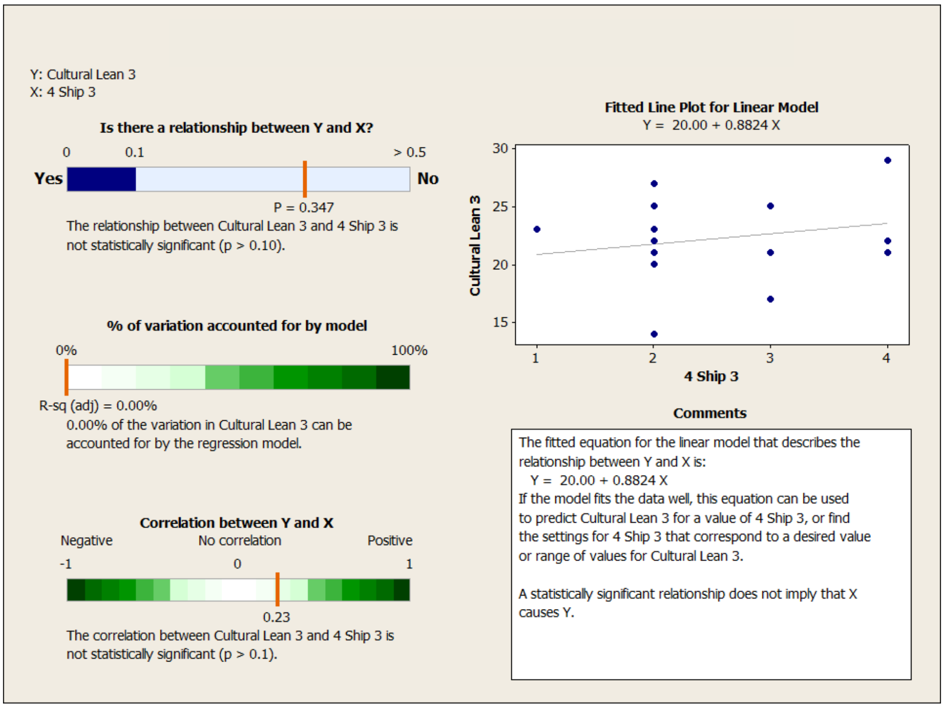

Figure 4-29. Summary Report for Cultural Lean vs 4 Ship Sims for 3(F) Squadron

Again both front-line squadrons agreed with each other, demonstrating a very weak positive correlation, albeit over differing data ranges, although neither was statistically significant enough to reject the null hypothesis. In common with the two-ship data it appeared that the squadrons conducted remarkably differing numbers of 4-ship sorties. XI squadron a total of 20; an average of 1.2 per pilot whilst 3(F) conducted a total of 51, equating to 2.7 per pilot.

Thus, it would appear that both front-line squadrons agree broadly with each other in terms of the correlation for 2 and 4-ship simulator sorties flown. It can also be seen that the squadrons are using the simulator in distinctly differing ways in order to get their mandated hours. 3(F) Squadron preferring to place their pilots in four-ship simulators whilst XI Sqn electing for more 2-ship simulator missions.

4.16.6 RAF CONINGSBY AND THE THEME OF 'NEAR FUTURE'

The same variables were used to regress the theme of 'Near Future' for the RAF Coningsby responses, the results being held in Table 4-2. All variables conformed to the assumption of normality and where outliers were found they were removed in accordance with the reasons articulated above.

Table 4-2. RAF Coningsby, 'Near Future' Pearson's Coefficient

Variable	Degrees of Freedom (dF)	Pearson's Correlation (r)	Statistical Probability (%)	r-sq (%)	Hypothesis upheld
Age	45 (1 x Outlier)	0.03	0.866	0.0	H_0
Typhoon Hours	46	0.09	0.553	0.0	H_0
Total Hours	45 (1 x Outlier)	0.08	0.611	0.0	H_0
2 Ship Simulators	45 (1 x Outlier)	0.02	0.888	0.0	H_0
4 Ship Simulators	45 (1 x Outlier)	0.24	0.099	3.8	H_A

4.16.7 RAF LOSSIEMOUTH AND THE THEME OF 'CULTURAL LEAN'

The data gathered by the RAF Lossiemouth returns were also treated in the same manner against 'Cultural Lean', correlations against the above variables being sought. The results are given in Table 4-3.

Table 4-3. RAF Lossiemouth. 'Cultural Lean' Pearson's Coefficient

Variable	Degrees of Freedom (dF)	Pearson's Correlation (r)	Statistical Probability (%)	r-sq (%)	Hypothesis Upheld
Age	27	0.37	0.046	10.8	H_A
Typhoon Hours	27	0.41	0.026	13.8	H_A
Total Hours	27	0.48	0.008	20.2	H_A
2 Ship Simulators	26 (1 x outlier)	-0.18	0.361	0.0	H_0

It should be restated that the 'Cultural Lean' variable at RAF Lossiemouth differed from that used at RAF Coningsby in that it was altered by the removal of a single question such that the theme was found to be consistent. The results for the variables are presented in Table 4-3; unless stated otherwise there were no outliers and all variables were normally distributed as tested by a Ryan-Joiner test. As an error check the tests on the variables above were re-run with the errant question being included (the Cronbach's alpha dropping to 0.68, 0.02 beneath the stated acceptable consistency level). The results were similar to those in Table 4-3, retaining both the statistical significance and the level of correlation.

4.16.8 MULTIPLE REGRESSION

Running a ‘best subsets’ regression for all the variables in Table 4-2 found that the themes of ‘Cultural Lean’ and ‘Near Future’ at RAF Coningsby were best explained using the following variables, shown in Table 4-4.

Table 4-4. Best Subset Regression, RAF Coningsby

Theme	R-Sq (adj)	Mallows Cp	S	Variables
Cultural Lean	9.3	2.0	3.7	Total Hours 4-Ship Sims
Near Future	3.3	2.2	3.4	Age Total Hours 4-Ship Sims

These results demonstrated an extremely low R-Sq (adj) (see Appendix I) showing that the model has virtually no predictive power; it did, however, provide a general indicator of trend when these variables are applied. In summary there appears to be no silver bullet to increase the acceptance of simulation or to increase favourable opinion on ‘Near Future’ capabilities.

4.16.9 SUMMARY OF WORK ON THEMES

This exploration has searched for evidence to support the traditional and emerging beliefs of the Typhoon Force, however the Cronbach Alpha results for the Entry Questionnaire’s internal validity makes it clear that the two main operating bases differ subtly in their views of what makes up cultural acceptance and also how simulation is viewed as a whole. There is, however, a notable similarity between both bases that cultural acceptance *is* related to age irrespective of their differing understanding of the theme. Whilst only the RAF Lossiemouth results are statistically significant the broadly positive correlation across both bases demonstrates no supporting evidence to suggest that the statement ‘younger pilots are more accepting of simulation’ is indeed true.

With respect to the effect of the number of Typhoon hours each pilot had flown Lossiemouth returned a statistically-significant positive correlation for their understanding of the theme. At Coningsby both front-line squadrons had positive correlations but with only 3(F)'s being statistically significant. Thus, of those questioned at both bases it can be stated that as a pilot's hours on the Typhoon increases he or she is likely to have a more positive attitude towards simulation (causality not inferred). The number of total hours a pilot has flown also returned a statistically-significant positive correlation at both bases with the exception of XI Squadron (non-significant neutral correlation). Thus, it can be said that in general the variable of Total Hours is positively correlated to a positive acceptance of simulation.

Across both bases there was a negative correlation between the number of 2-ship simulators flown and cultural lean for those asked, although this was only statistically significant at XI Squadron at Coningsby. In contrast the favourable opinion of those questioned at the front-line squadrons at Coningsby increased with the number of 4-ship simulators flown. This latter variable was the only one also to return a positive correlation in the theme of 'Near Future'.

Whilst similarities have been highlighted above it remains that there are observed differences between the two bases in the homogeneity of their viewpoint. Coningsby returned a negligible r-sq for the majority of the variables tested (0-5%), Lossiemouth by way of comparison was generally higher (10-20%). This is supported by the result of the Lossiemouth pilots believing that their perception of the Force's opinion matched their own, whilst at Coningsby they believed themselves to be individually more positive than the Force as a whole.

4.17 DETERMINING THE LIVE: SYNTHETIC BALANCES

Having examined the LSB for other air forces in the literature review (Section 2.3) and investigated the limits of the balance for the Operational Conversion Unit in Trial PANDORA'S BUZZARD, this section used the opinion of the current pilots to search for an answer to the question 'What is the correct

LSB for the Typhoon Force', with reference to day-to-day training. The intent was to determine if this is a broadly similar figure for all disciplines or if it was affected by the type of training and complexity the simulator was trying to replicate. The specific objectives to be examined in the section were:

a. **Objective 4**. Determine the subjectively-assessed LSB for each of the required tasks.

b. **Objective 5**. Investigate the effect of threat complexity on the LSB.

Each complexity level was examined in turn to source a LSB for individual skill sets, and for the complexity level as a whole. Each level had a number of tasks that were relevant to that level, each task being required to be trained a certain number of times per year as laid down by the Typhoon Force HQ, these were known as events. The questionnaire asked the pilots to give their opinion of the *minimum* number of events for each task that should be done in the Live and the Synthetic environments. In doing so there is a number of events left over that by definition could be done in either environment. This will be termed '*The Option'* as it will provide the commanders with the option to vary the synthetic blend from heavily synthetic to heavily live flying without exceeding the perceived minimums. This would be used in times of conflict, such as Operation ELLAMY in Libya, when many of the Force's aircraft were deployed overseas yet pilot training had to continue back in the UK. Thus, rather than the LSB terminology that has become commonplace in the RAF this paper utilises the term Live – Option – Synthetic Blend (LOSB) as a method of providing greater fidelity to the issue. The term LSB will be reserved for a specific target within the LOSB, such as the 1:1 set by the RAF.

4.17.1 <u>FINDING THE LIVE OPTION SYNTHETIC BLEND FOR LEVEL 1</u>

Level 1 represented the lowest threat level and as such all pilots within the Trial had had experience operating and training at this level (n=48). The Force HQ requirements for level 1 were broken down into sections: major tasks, some of which had specific sub-tasks associated with them, and common skills that were required for all major tasks. The pilot questionnaire asked for the minimum events for each of the major and sub-tasks and common skills.

Table 4-5. Task Breakdown by Group for Level 1

GROUP	MAJOR TASK	SPECIFIC SUB-TASK	COMMON SKILLS
QRA	LRA Intervene	High level Intercepts Low Fast Intercepts Helo Intercepts Low Slow Intercepts Scramble VID P1 VID P2 Intervene Shadow	Pairs Take off Pairs Landing DA ECM DA Flares DA DASS DA Chaff NME Chaff NME Flare NME Jam VID P4 CAP Low level Datalink
Counter Air (CA)	Engage Day	ACT A2A Gunnery DACT	
	CA Day		
	HVAD Day		
	Engage Night	Night Tactical Formation Night Landing	
	CA Night		
	HVAD Night		

To determine the LOSB the average number of minimum events for each discipline was found, see , however when testing for normality, even after outlier modification, very few of the tasks or skills were found to display a normal distribution, as recorded in the Ryan-Joiner test p-value column. Thus, the median was used for the majority of the remaining tasks. Additional complications were caused by the number of events per year for each task and skill, as laid down by the Force HQ, varying from 1 to 112; the HVAD Day and

Night tasking having only 4 and 1 events required per year respectively. For these tasks the mode was believed to provide a better indicator of the average for these tasks and skills. In sum, determination of the particular type of average was made after outlier modification (using the methodology previously detailed), normality testing and histogram inspection before recording the type of average in the 'Type' column of Table 4-6. It can be seen that tasks and skills with few events required per year, such as HVAD Day and Night provide little fidelity of pilot opinion, as such the final column of Table 4-6 provides an indication of the sensitivity of each event - a task requiring 10 events per year returning a sensitivity value of 10%. This column is colour coded to provide (subjective levels) of indication quality: red 100-25%, yellow 24-11% and green =<10%, the lower number the better as this indicates greater sensitivity e.g. HVAD night requires only 1 event a year thus, a single event alters the LOSB by 100%, not a sensitive measure.

Table 4-6. LOS Blend for Level 1

Major Task	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
		Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synth	
CA Day	112	53.5	46.6, 60.3	23.2	24.7	20.8, 28.56	13.3	>0.1	>0.1	Mean	48	30	22	0.9
CA Night	4	3	2, 3	1	1	1, 2	1.2	N/A	N/A	Median	75	0	25	25.0
Engage Day	48	25.9	22.5, 29.3	11.7	11.1	9.1, 13.1	6.9	>0.1	>0.1	Mean	54	23	23	2.1
Engage Night	24	13.5	11.7, 15.4	6.2	5.8	4.6, 6.9	3.9	>0.1	>0.1	Mean	56	20	24	4.2
LRA Intervene	10	1	0, 2	2.2	4	2, 5	3.2	0.04	>0.1	Median	10	50	40	10.0
HVAD Day	2	1			1			N/A	N/A	Mode	50	0	50	50.0
HVAD Night	1	1			0			N/A	N/A	Mode	100	0	0	100.0
Total LOSB	201	98.9			47.6						49	27	24	

Major Task	Sub Task	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
			Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synth	
LRA Intervene	High level Intercepts	2	1			1			N/A	N/A	Mode	50	0	50	50.0
	Low Fast Intercepts	1	1			0			N/A	N/A	Mode	100	0	0	100.0
	Helo Intercepts	4	3	2, 3	1.2	1	0.4, 1.0	1	N/A	N/A	Median	75	0	25	25.0
	Low Slow Intercepts	2	1			1			N/A	N/A	Mode	50	0	50	50.0
	Scramble	2	1			1			N/A	N/A	Mode	50	0	50	50.0
	VID P1	4	2	1, 2	1.2	2	1, 2	1.2	N/A	N/A	Median	50	0	50	25.0
	VID P2	4	2	1, 2	1.1	2	1, 2	1.2	N/A	N/A	Median	50	0	50	25.0
	Intervene	1	1			0			N/A	N/A	Mode	100	0	0	100.0
	Shadow	48	10	5.4, 15.7	10.8	10	5, 12	11.8	>0.1	<0.01	Median	21	58	21	2.1
Engage Day	ACT	12	12	11.4, 12	2.1	0	0, 0	1.2	<0.01	0.04	Median	100	0	0	8.3
	A2A Gunnery	1	1			0			N/A	N/A	Mode	100	0	0	100.0
	DACT	10	10	10, 10	2.2	0	0, 0	1.23	0.07	<0.01	Median	100	0	0	10.0
Engage Night	Night Tactical Formation	10	6.6	5.7, 7.4	2.8	1.7	1.2, 2.2	1.8	>0.1	>0.1	Mean	66	17	17	10.0
CA Night	Night Landing	10	6	5, 8	2.6	1	0.4, 2.0	1.9	>0.1	>0.1	Median	60	30	10	10.0
HVAD Night															

Common Skills	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
		Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synth	
Pairs Take off	10	10	6, 10	3.6	0	0, 0	0.23	>0.1	>0.1	Median	100	0	0	10.0
Pairs Landing	10	9	5.4, 10	3.7	0	0, 0	0.25	0.08	>0.1	Median	90	10	0	10.0
DA ECM	25	6.5	9.4, 13.2	6.5	8.1	6.6, 9.6	5.1	>0.1	>0.1	Mean	26	42	32	4.0
DA Flares	14	7	5.7, 8.2	4.2	4.7	3.6, 5.7	3.5	>0.1	>0.1	Mean	50	16	34	7.1
DA DASS	96	40	29.7, 48.0	24	24	12.0, 30.8	21.9	0.05	0.12	Median	42	33	25	1.0
DA Chaff	10	5	5, 5	3	4	2, 5	2.5	>0.1	>0.1	Median	50	10	40	10.0
NME Chaff	16	9.1	7.8, 10.4	4.5	3.8	2, 6	3	>0.1	>0.1	Mean	57	19	24	6.3
NME Flare	12	6.3	5.2, 7.5	3.9	3.3	2.5, 4.1	3.3	>0.1	>0.1	Mean	53	20	28	8.3
NME Jam	6	4.2	3.7, 4.6	1.5	1.6	1.2, 1.9	1.3	>0.1	>0.1	Mean	70	3	27	16.7
VID P4	4	2	2, 3	1	1	1, 1.1	0.8	>0.1	>0.1	Median	50	25	25	25.0
CAP	36	17.5	14.3, 20.7	10.9	7.5	5.7, 9.3	6.1	>0.1	>0.1	Mean	49	31	21	2.8
Low level	24	20	16.1, 24.0	6.4	0	0, 1	1.8	0.01	<0.01	Median	83	17	0	4.2
Datalink	48	20	16, 24	13.2	12	9.8, 18.2	11.1	>0.1	0.08	Median	42	33	25	2.1

Figure 4-30 provide a graphical breakdown of the LOSB for the major tasks at Level 1. The graph shows the average of the minimum events of each event for each realm; Live, Synthetic or the amount of 'The Option' that is available, the figures within the bars show the numerical quantity of corresponding events.

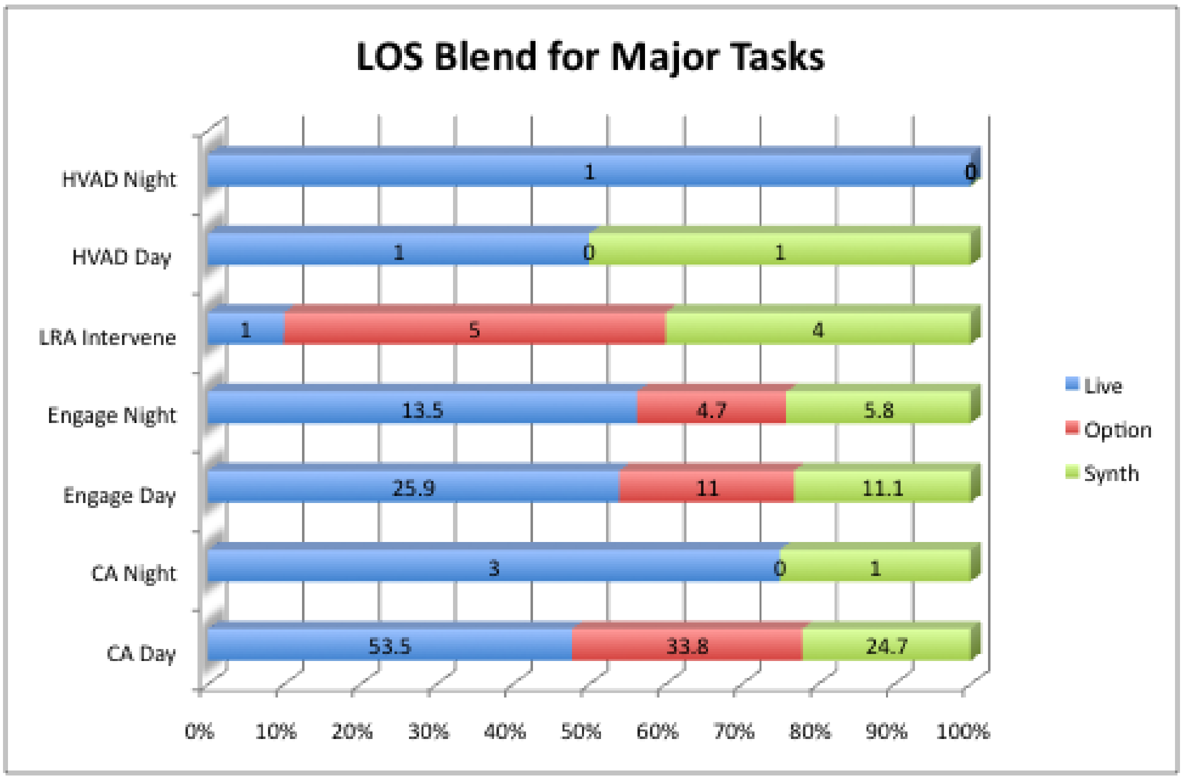

Figure 4-30. LOSB for Major Tasks – Level 1

Discounting the tasks that have poor sensitivity (HVAD Night, HVAD day and CA Night), it can be seen that the remaining Major Tasks have sufficient events per task per year to gain an understanding of the numbers the pilots would accept as a minimum in each environment.

Some of the Major Tasks have specific Sub-Tasks associated with them, see Table 4-5. Task Breakdown by Group for Level 1. These subtasks are not exclusive to a task *per se* but a pilot could be reasonably expected to utilise this skill when predominately operating or training this Major Task. The LOSB for these Sub-Tasks is presented at Figure 4-31. Finally the LOS Blend for the Common Skills used by all the Tasks is found at Figure 4-32.

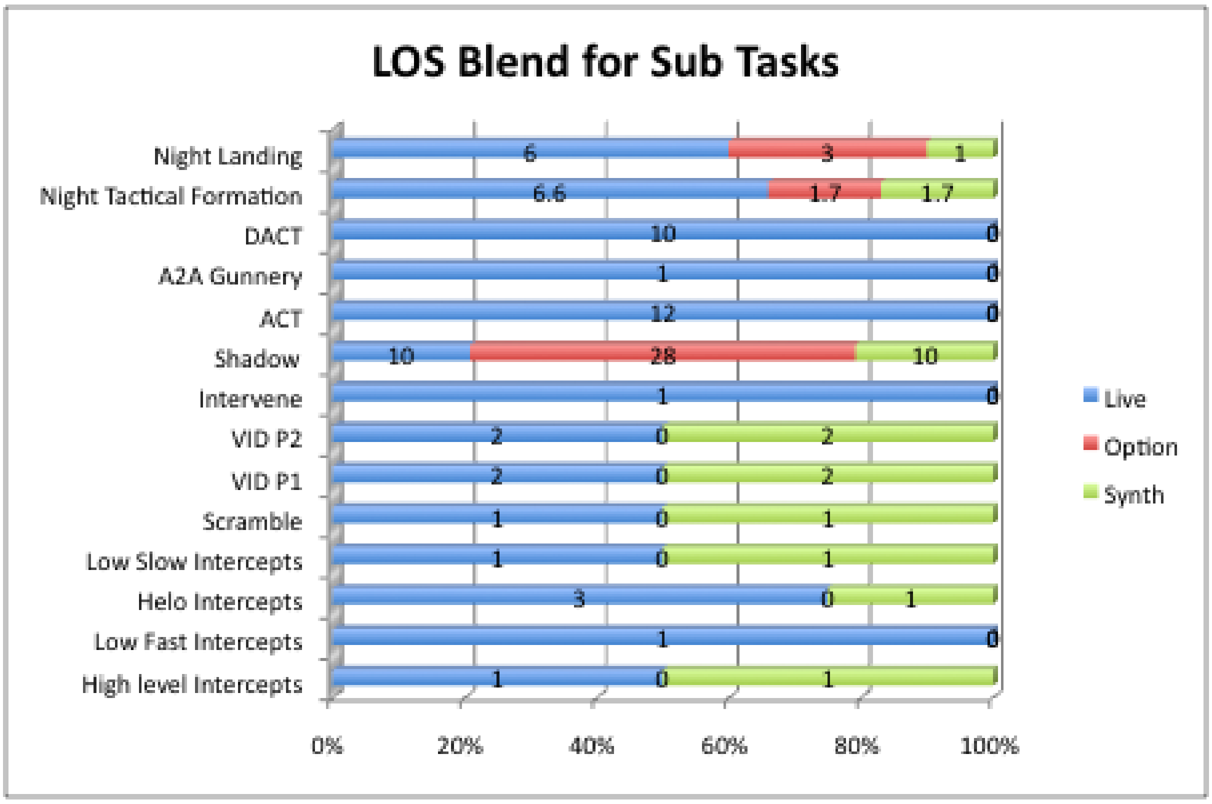

Figure 4-31. LOSB Sub-Task – Level 1

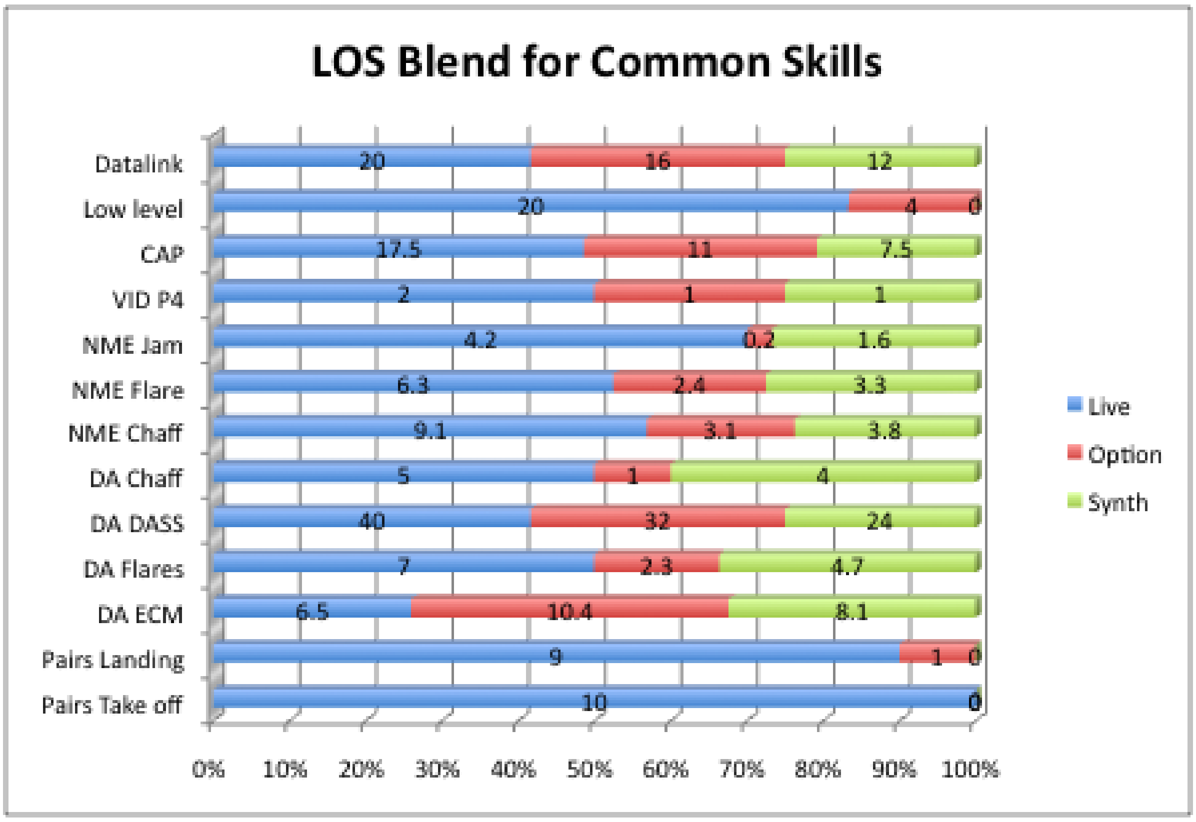

Figure 4-32. LOS Blend for Common Skills - Level 1

The results for the common skills showed, in broad terms, a minimum live proportion of approximately 50%. The exceptions to this figure were logical skills that required large amounts of hand-eye coordination with little reference to internal instrumentation; pairs take off and landing and low-level flying. DA ECM (Defensive Aids Counter Measures), the ability to jam the enemy's radar, also produces a logical result as once the switch is made at the beginning of the sortie there is no indication in flight of its success against the enemy. The remaining skills show a minimum synthetic proportion of approximately one third. Broadly the Common Skills LOSB is 50-20-30 with the exceptions mentioned above.

The Sub-Tasks were more difficult to interpret with a number of the tasks having poor sensitivity. However, tasks requiring large quantities of hand-eye coordination - ACT (dogfighting) and DACT (dogfighting with a dissimilar aircraft) - again have a high requirement for live flight. Shadow, a Sub-Task of LRA Intervene, had a large proportion of Option available indicating that this task is served equally well in both domains.

Of the major tasks with sufficient sensitivity CA Day, Engage Day, Engage Night all appear to favour a LOSB(%) of approximately 50-25-25. LRA Intervene is the practice of intercepting Russian Long Range Aviation in all weathers over the North Sea, such as regularly reported in the media. This skill, however, had a much larger Synthetic and Option proportion than the others which is commonsensical given its strong reliance on procedural working of the aircraft sensors, actions that occur entirely inside the cockpit with little reference to the outside world.

Finally the LOSB for level 1 as a whole was calculated by totalling the number of events and Live and Synthetic averages, this produced a LOSB(%) of 49-27-24. Given that 201 events per year were required at level 1 and the standard planning figure of 2 events per sortie could be achieved this equated to a total of 100 sorties per year for each level 1 pilot, of which 49 would be in the live environment and 24 in the simulator, the remaining 27 being placed in either domain.

4.17.2 LOS BLEND FOR LEVELS 2 AND 3

The same tables and figures were created for level 2 (see Appendix J) which created a similar picture. Level 2 having an overall LOSB(%) of 52-23-25 and for Level 3; 50-24-26. With the addition of the same treatment for Common Skills and Sub Tasks this represented the first time the culturally-accepted minimums for the combination of live and synthetic training had been mapped. If the same assumption of events achieved per sortie are made for levels 2 and 3 it can be shown that the Force HQ can calculate the Live and synthetic minimum requirement to be:

$100*P_1 + 68*P_2 + 65*P_3$ = Minimum No. of *Live* Sorties Required Per Year

$24*P_1 + *32P_2 + 35*P_3$ = Minimum No. of *Synthetic* Sorties Required Per Year

where $P_{1,2 \text{ and } 3}$ =No of pilots at Qualification Level 1,2 and 3

In addition to these minimum figures further sorties will have to be undertaken to ensure the correct number of events are achieved for each level. This is ‘The Option’ and can be undertaken in either synthetic or live flight. The formula is:

$27*P_1 + *31P_2 + 32*P_3$ = Additional Sorties Per Year to satisfy ‘The Option’.

These figures allow the RAF to forecast fuel costs and technical spares to a greater degree of accuracy in the long term. For the short term the use of ‘the option’ to increase the synthetic proportion will allow the planning forecast to be better met such that the Typhoon Force does not over run its budget yet still maintains valid tactical training. In addition any Combat or Contingency Operation would result in a reduction of aircraft at home and thus increasing the synthetic proportion of training into ‘the option’ will ensure pilots are trained for Operations without increased flying, and servicing of, the aircraft remaining in the UK.

Proportions. Figure 4-33 and Figure 4-34 show the (statistically-significant) relationship between the synthetics and live flight for the 2 major tasks of CA Day and Engage to be a quadratic. function Each curve peaks at an LSB of 67:33, showing this to be the maximum LSB presently acceptable to the sample at Level 1. This value is indicative only as the model is not strong enough to make it predictive in nature. Figure 4-35 is the only other statistically-significant relationship observed within the sample, but unlike the previous two this is linear in character. The equation for the relationship being Y = 6.3 + 0.4 X, where Y is ASTA and X is Live training events; as with the other relationships the model is not strong enough to be used as a predictive tool.

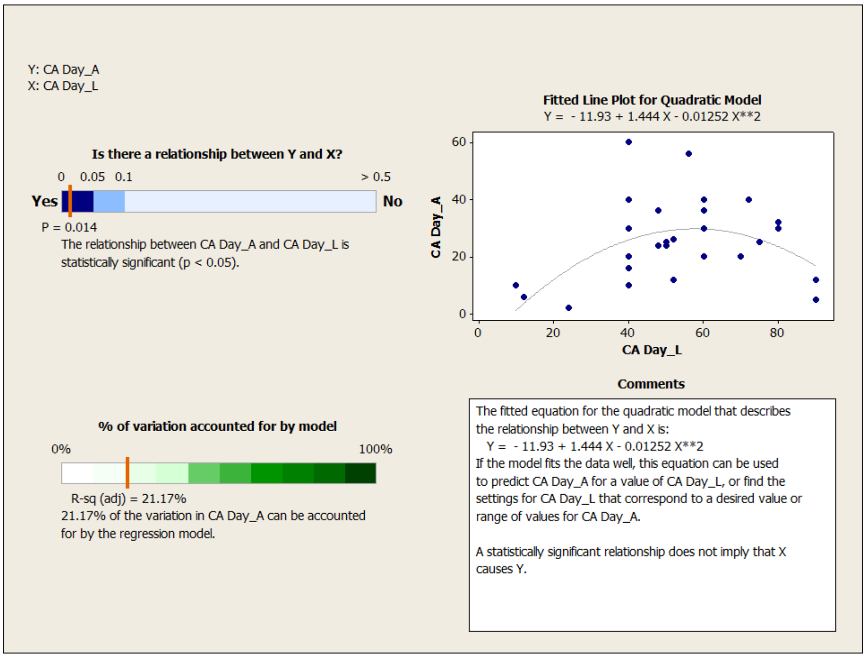

Figure 4-33. CA Day, ASTA v Live Proportion.

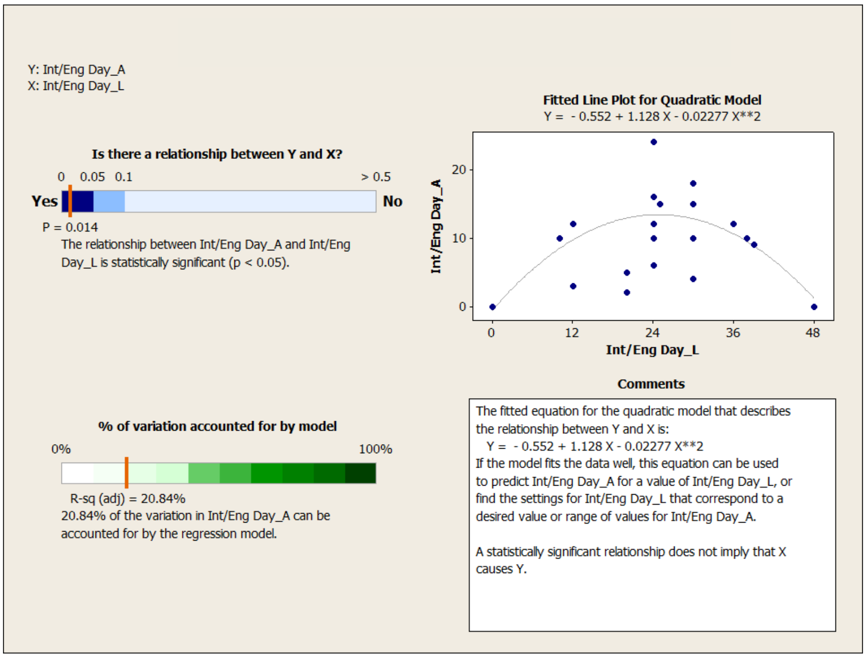

Figure 4-34. Engage, Day ASTA v Live Proportion.

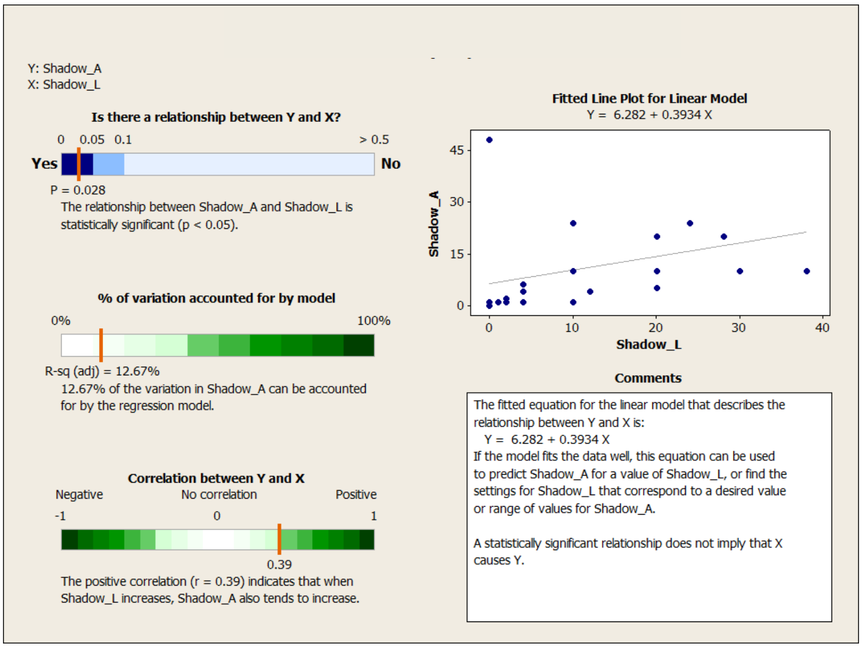

Figure 4-35. Shadow, ASTA v Live Proportion.

4.17.3 THE EFFECT OF CONCURRENCY

The lack of concurrency of the ASTA devices was often quoted as the reason why pilots do not want to train synthetically, this in turn affecting the

LSB. Concurrency primarily concerns aircraft software that drives elements of the displays and controls as well as the Defensive Aids Sub Suite (DASS) and Radar; poor concurrency affects motor programmes and display interpretation. However, achieving a fully-concurrent simulator with software that matches the aircraft within 30 days of its release has a significant cost, thus the intent of this section is to provide the procurement teams with information to assist their cost/benefit analysis calculations. A reduction in the number of events rto be conducted in the live environment would save both fuel (approximately £1050 per flying hour) and technical spares (~£1020 per hour). Should these savings, aggregated across the Force's pilots, be greater than the cost of securing the contract for concurrency then savings could be made with no loss of training.

Question 2 of the post-sortie questionnaires therefore sought to examine what effect, if any, fully-concurrent software in the simulator would have made on the LSB. The questions were centred on 3 topics: CA Day, Engage Day and Enemy (NME) Jamming techniques. The first 2 being the major tasks that form the bedrock of Typhoon training, the last being a subtask that had yet to be exploited fully in the synthetic environment and promised the largest mid-term change in capability.

Thus, to determine if the inclusion of concurrent software altered the LOSB averages for the 3 tasks the hypotheses below were used. As money would have to be found to fund the concurrent software this would have to be found by reducing live flight, thus the hypotheses look for both an increase in the synthetic proportion and a reduction of the live.

For the Live environment:

H_0: The amount of Live training remains the same or increases if the simulators were to have concurrent software.

H_A: The amount of Live training reduces if the simulators were to have concurrent software.

For the Synthetic environment:

H_0: The amount of Synthetic training stays the same or reduces if the simulators were to have concurrent software.

H_A: The amount of Synthetic training increases if the simulators were to have concurrent software.

To determine if there was a statistically-significant difference between the number of live and synthetic sorties recommended should the software be upgraded to a fully-concurrent one a paired-sample t-test was attempted. This found only the subject of 'NME Jam_Synthetic' to satisfy the assumption of no outliers and a normality of the differences. The other comparisons violated this assumption but did demonstrate symmetry in the distribution of the differences. Consequently the non-parametric Wilcoxon signed rank test (see Appendix I) was conducted for the variables of 'NME Jam_Live', 'Engage_Live' and 'CA_Live'. The final two variables of 'CA_Synthetic' and 'Engage_Synthetic' failed the requirement for symmetrical distribution of the differences and thus, despite its lower power, a paired sign test was conducted. Given the need to demonstrate a robust argument for change in the allocation of funding an alpha level of 0.05 was chosen. The results of these tests are given in Table 4-7. Effect of Concurrency on Recommended Sorties - Level 1

Table 4-7. Effect of Concurrency on Recommended Sorties - Level 1

Variable	Test Used	N	N for test	Statistic	p	Median delta	Hypothesis Upheld
CA_Live	Wilcoxon	40	20	W=68.5	0.089	-1.5	H_0
Engage_Live	Wilcoxon	40	13	W=28.5	0.124	0	H_0
NME Jam_Live	Wilcoxon	29	31	W=26.5	0.000	-1.85	H_A
CA_Synth	Sign	40	2 below 23 above	-	0.000	5.5	H_A
Engage_Synth	Sign	40	1 below 20 above	-	0.000	1.0	H_A
NME Jam_Synth	Paired - t	39	-	T=0.97	0.170	0.24 (mean)	H_0

Table 4-7. Effect of Concurrency on Recommended Sorties - Level 1 shows that for level 1 concurrent software would increase the number of sorties recommended in the synthetic environment for the tasks of Counter Air and Engage. In order to fund the software to achieve this recommended increase the only Live flying reduction could come from NME Jam. This financial saving would be equivalent to the fuel and spares costs associated with 1.85 sorties per pilot per year.

The process was repeated for Levels 2 and 3 which found multiple outliers in all the variables, which were unable to be modified or deleted without substantial changes to the overall results. As such Wilcoxon sign tests were carried out, see Table 4-8 and Table 4-9.

Table 4-8. Effect of Concurrency on Recommended Sorties - Level 2

Variable	Test Used	N	N for test	Statistic	p	Median delta	Hypothesis Upheld
CA_Live	Wilcoxon	36	13	W=40.0	0.363	0.0	**H_0**
Engage_Live	Wilcoxon	36	18	W=117.5	0.922	0.0	**H_0**
NME Jam_Live	Sign		22 below 1 above	-	0.000	-1.0	**H_A**
CA_Synth	Sign	36	0 below 17 above	-	0.000	0.0	**H_0**
Engage_Synth	Sign	36	3 below 19 above	-	0.000	4.0	**H_A**
NME Jam_Synth	Sign	36		-	0.000	0.8	**H_A**

Table 4-9. Effect of Concurrency on Recommended Sorties - Level 3

Variable	Test Used	N	N for test	Statistic	p	Median delta	Hypothesis Upheld
CA_Live	Wilcoxon	20	7	W=17	0.723	0.0	**H_0**
Engage_Live	Sign	20	4 below 4 above 5	-	0.637	0.0	**H_0**
NME Jam_Live	Sign	21	13 below 1 above	-	0.001	-1.0	**H_A**
CA_Synth	Sign	20	1 below 11 above	-	0.003	8.0	**H_A**
Engage_Synth	Sign	20	2 below 10 above	-	0.019	2.0	**H_A**
NME Jam_Synth	Sign	20	4 below 8 above	-	0.194	0.0	**H_0**

The results for Level 2 and 3 show a willingness on behalf of the pilots to reduce the minimum level of task 'NME Jam' that is conducted in the Live environment as well as increasing the proportion of the task 'Engage' in the synthetic. Level 3 also shows a significant rise in the number of minimum sorties recommended to be undertaken in the synthetic environment for ‘CA’.

4.17.4 SUMMARY

The intent of the investigation into the effect of concurrent software in the LOSB levels recommended by the pilots was to determine if, firstly, there was a desire to increase the amount of synthetic training if the software matched that of the most up-to-date aircraft. Secondly, and more pragmatically, if the funding for that concurrency could be found by the savings made through recommendations of a reduction in live flight.

The results found that there was a greater demand for synthetic training with the up to date software; an increase of 6.5, 4.8 and 10 sorties per pilot per year for levels 1,2 and 3 respectively. However, only the Live task of ‘NME Jam’ was offered to be traded to achieve this extra synthetic training. This was at the rate of 1.85, 1 and 1 sorties per pilot per year for each of the levels respectively. Thus in order to fund a contract for concurrency the contract should cost no more than:

Sortie saving x cost of saving (£2070) x number of pilots

Thus for levels 1, 2 and 3 respectively and assuming 20 pilots on each level:

(1.85*2070*20) + (1*2070*20) + (1*2070*20) = £159,390 per annum

In procurement terms this is an extremely low figure to generate a standing army of software coders to ensure the concurrency of the synthetic aircraft. Therefore whilst up-to-date software has been shown to have a positive effect on the number of synthetic sorties recommended there is no evidence to suggest that the Typhoon pilots are willing to trade live flying to fund that increase. It is recommended on a simple value for money basis therefore, that concurrent software is not procured.

4.17.5 COMPLEXITY

Sub Question 2, Objective 5 was to ' Investigate the effect of threat complexity on the LSB'. To this end it had been anticipated that a repeated measures ANOVA would be run to detect differences between each of the three levels of complexity and the number of events recommended for each Task, Sub-Task and Skill. In the event, however, there were 36 pilots who completed the simulator and questionnaires for levels 1 and 2, and only 20 for levels 2 and 3. Thus, whilst slower and more laborious, a paired t-test was elected to be used for comparison of the complexity levels, this was further complicated by outliers and a lack of symmetry on a number of the variables. Thus Wilcoxon and Signed Tests were used where appropriate to determine if there had been a change in the recommended events to be undertaken for each of the variables.

Selection of which tasks to compare was complicated by the issue that not all Tasks, Sub-tasks or Skills were used at each level and that in some cases the number of events required to be undertaken was not sensitive enough to be used for analysis. By discounting the Tasks with poor sensitivity (as stated in section 4.17.1 and shown in red in Table 4-6) and mapping for Tasks and Skills that could be compared across the levels a comparison table was constructed, see Appendix J. The comparison was then made between level 1 and 2, 2 and 3, and 1 and 3.

The hypothesis for the tests is set out below. Given the exploratory nature of the tests and the assumption that the outcome was to be informative only and would not be used for a re-write of the syllabus α was set at 0.1.

H_0 : the mean difference between the paired values was equal to, or greater than zero, or specifically there was no increase in the minimum number of events recommended as the complexity level increased.

H_A = the mean difference between the paired values was less than zero, or specifically there was an increase in the minimum number of events recommended as the complexity level increased.

The methodology described led to a total of 112 comparisons of the variables across the various levels, consequently the full results are given in Appendix J but a summary table is provided below containing only those results that gave the statistically- significant outcome for the alternate hypothesis.

Table 4-10. Summary of Recommended Sortie Increase Per Complexity Level

Levels Compared	Variable	Variable Type	Environment	p-value	Increase Recommended	% increase
1 v 2	DA ECM	Common Skill	Live	0.09	1.5	6
	DA Flares	Common Skill	Live	0.02	1	7
	DA Chaff	Common Skill	Live	0.09	0.5	3
	NME Flare	Common Skill	Live	0.08	0.5	4
	Datalink	Common Skill	Synthetic	0.09	0.5	1
2 v 3	Nil	Nil	Nil	Nil	Nil	-
1 v 3	CA Day	Major Task	Synthetic	0.01	9.2	8
	Engage Day	Major Task	Synthetic	0.02	3.9	8
	Engage Night	Major Task	Synthetic	0.03	2	8
	DA DASS	Common Skills	Synthetic	0.03	8.3	9
	DA Chaff	Common Skills	Live	0.05	1	6
	NME Flare		Live	0.09	1	8
	NME Jam		Synthetic	0.04	0.5	8

Table 4-10 shows that of the 112 tests conducted only 12 recommended an increase in the minimum number of events that was statistically significant. This low number is likely to be a function of the low power of the statistical tests used given the need to utilise non-parametric methods due to the sample sizes and distribution. Of note however is that the comparison between level 1 and 2 only returned results for the lowest level requirements – Common Skills, additionally these observations were made in the live environment for 4 of the 5 results. Despite small increases recommended for the number of events this did not filter up to the Major Task level. In contrast the comparison between the highest level of complexity and the lowest found an increase across all the Major Tasks and only within the synthetic environment.

4.17.6 SUMMARY

For the sample tested the minimum number of events recommended for all Major Tasks increased by approximately 8% of the annual requirement, for the synthetic environment only, when comparing training at the highest complexity level versus the lowest. In plain language this equates to 15 sorties per year moved into synthetics when training at the highest threat levels i.e. near-peer threats.

4.17.7 THE EFFECT OF COMPLEXITY ON THE MAJOR FACTORS

This section will initially determine if any of the 'major factors' are correlated to the minimum amount of simulation recommended by the front-line pilots. These results will be contrasted to those found at each level to search for similarities. The 'major factors' are defined as: age, flying hours (Typhoon), flying hours (total), average numbers of 2 and 4-ship simulator missions flown in a 2-month period. The intent is to determine if there is some experiential factor that is linked to the number of simulator sorties considered necessary at each level. This investigation is necessary as the number of simulator devices are limited, thus targeting the right pilots would increase efficiency of resource. As this section's major intent is to source efficiency of resource only the Major Task requiring the largest number of events per year will be considered – Counter Air (CA) Day. As this is intended to affect front-line pilots only, 29 Squadron (the OCU) will be discounted from the analysis.

A Pearson's product moment correlation was run to assess the relationship between Flying Hours (Typhoon) and the minimum number of simulator sorties recommended for the Counter Air Day Task. Initial analysis showed the relationship to be linear with both variables normally distributed (Ryan-Joiner test $P > 0.1$), and there were no outliers. At level 1 the correlation between amount of simulation recommended and Flying Hours (Typhoon) was found to be moderately negative $r(40) = -0.38$, $p < 0.01$with Flying Hours (Typhoon) explaining 12.5% of the variation.

The only other major factor that proved to be statistically-significant was the number of 2 Ship sims undertaken in a 2 month period. The correlation was weak - $r(40) = 0.37$ with $P=0.03$ and R-Sq (adj) =11%.

At level 2 the same factors were found to be significant:

a. Flying Hours (Typhoon) $r(33) = -0.47$, $P=0.01$, R-sq (adj) = 20%. The correlation being in the same sense and slightly stronger than the level 1 results.

b. Number of 2-Ship Sims $r(33) = 0.42$, $P = 0.01$, R-Sq (adj) = 14.7%. The correlation being of similar sense and strength to level 1.

At level 3 only the major factor of 2-Ship simulator missions was found to be statistically significant $r(31) = 0.52$, $P = 0.02$, R-Sq(adj) = 22.4%. Again the correlation coefficient increased when compared to the previous level.

Thus, it may be considered that the number of synthetic sorties recommended for each level has proved to have an experiential factor associated to it. Before this can be stated with confidence there is a need to determine if the results were being confused by a hidden correlation between the number of 2-ship simulator missions undertaken and Flying Hours (Typhoon) ie. the more experienced pilots were undertaking fewer simulator missions. This would determine if the factors of 2-Ship Sims and Flying Hours (Typhoon) were independent. Investigation into the association between these to variables found no statistically-significant correlation at any of the levels.

Thus, it can be said that the minimum number of Counter Air (Day) events recommended to be undertaken in the simulator was linked to the number of 2-Ship simulator missions undertaken by the pilots in a two-month period (for all three levels) and the number of Flying Hours (Typhoon) for the first two levels of complexity.

In order to contrast the tests above they were rerun to determine if any of the factors could be correlated to the minimum number of *live* flying events. However, no statistically-significant correlations were found at any of the levels.

4.18 ANALYSIS

The lack of effect, across all levels, of the major factors on the recommendations for events undertaken in *live* flight indicates a consensus of the front-line pilots on the value of live flight within their preparation for combat. The reduction below these levels (approximately 50%) is likely to face strong cultural challenges and should therefore be approached with caution or avoided unless necessary.

When examining the results for the recommendations for minimum numbers of events in *simulation* the correlations show that the experiential factor of flying hours is moderately negatively correlated. This shows that the newer members of the squadrons recommended a higher number of simulator missions than their more experienced colleagues. These figures may be influenced by the fact that these individuals have recently exited the Operational Conversion Unit, which has a higher proportion of simulation, and are thus accepting of the regime. This line of thought was supported by the positively correlated relationship between the number of simulator events recommended in the Counter Air task and the increasing number of 2-ship simulator missions flown in the last 2 months. Although causation is not determined from these results (there being longitudinal retesting over a period of years required to determine this) any increase in simulation required in the Typhoon Force front-line, due to events such as high-tempo operations or fiscal restrictions, should initially be targeted at the less experienced members of the squadrons in order to bring their LOSB into line with their recommendations.

Although the LOSB remained approximately the same for all levels of complexity there was a statistically-significant increase in the minimum number of synthetic Major Tasks recommended at the highest level of complexity when compared to the lowest. Thus, should increased simulation be required targeting this at the higher levels of complexity first is likely to meet less cultural resistance than those at the lowest.

4.19 CONCLUSIONS

Objective 1. *Determine the present level of cultural acceptance of synthetic use in front-line Typhoon training and investigate if this is common across both Typhoon bases.*

Trial PANDORA’S DIAMOND has shown that the two Main Operating Bases of Coningsby and Lossiemouth differ in terms of their character – the themed questions returning differing internal validity. The implication of this is that the questions had differing meanings to the two bases. It is postulated that these differing viewpoints on simulation may be related to the types of simulators available at the bases and the manner in which they are used. RAF Coningsby having four high quality and fidelity devices used for all levels of training, by comparison with Lossiemouth has two much lower fidelity devices that are used primarily as emergency trainers.

Because of the differing internal validity the two bases cannot be compared directly, however it can be stated that, unlike RAF Lossiemouth, Coningsby had a statistically significant difference between the individual’s cultural acceptance of simulation and their perception of the Force’s opinion, the individual being the more positive. The implication being that *questioning the Force’s pilots as a collective, such as has been done recently* (Holden 2015)*, may return a more negative view of simulation than interviewing the individual operators themselves.*

The overall level of cultural acceptance was similar at both bases. The Cultural Lean theme returned a mean slightly more positive than the average: RAF Coningsby 20.5 (neutral point being 18) and RAF Lossiemouth 18.4 (neutral point being 15) although the standard deviation for both was of the order of 3.5, indicating that approximately 15% of those questioned sat below the neutral point. The acceptance level of the proposed Near Future uses of simulation was only able to be tested at RAF Coningsby and was positive, 22.9 $\boldsymbol{\sigma}_x$=3.6 against a neutral point of 18.

Objective 2. *Identify any factors such as age or experience that correlate to the level of acceptance.*

The discovery that the Squadron a pilot was assigned to had an influence on the level of acceptance was unexpected and as such requires further investigation. Whilst it may be postulated that leadership or work environment may be the cause this research can only state that it differed and as such squadrons should be considered separately. *Examining the variable of age within this context found no supporting evidence to suggest that younger pilots were more accepting of simulation than older ones*. Indeed the sample itself suggested that the opposite was true i.e. an unexpected marginally positive sense, defying conventional understanding that the younger generations were more accepting of technology and gaming and thus would be more accepting of simulation within their training.

Consideration of the experience variables found that the number of Total Flying Hours was statistically significant on both 3(F) and 29(R) Squadrons which resonates with the sample's returns on age: the older pilots having attained more total hours than the younger. *Thus there is evidence to suggest pilots with higher total hours will be more accepting of simulation.* In the case of number of simulator sorties undertaken, both front-line squadrons showed a decline in cultural acceptance with increasing numbers of 2-ship simulator missions flown, however this trend was reversed when considering 4 ship simulators, a relationship further underlined in the theme of 'Near Future' with 4-ship simulator missions returning the only statistically significant correlation. As the relationship was *not* seen when considering 2-ship simulator missions it suggests that it is the larger-scale tactical missions that nurture a positive opinion of simulation. Whilst causality has not been proven there is little effort required to alter the manner in which simulation is programmed. *Thus it is recommended that where constraints allow simulator hours are programmed for 4-ship missions rather than 2.*

RAF Lossiemouth's results returned an altogether different picture. Age, Typhoon Hours and Total Hours all returned a statistically-significant, positively correlated result. These results were accompanied by r-sq values of circa 15% and whilst very low they were still a factor of 10 greater than those of RAF Coningsby. This would further support the suggestion of considering the two bases as stand-alone when collecting evidence in the future rather than collating the evidence in a single pool.

Objective 3. *Prove or disprove some commonly-held beliefs that had started to form since the initial trial (PANDORA'S BUZZARD) had increased synthetic use, namely:*

1. *Younger pilots are more accepting of simulation than older pilots.*

2. *Acceptance of simulation is based upon a pilots qualifications.*

The answer to Objective 3a has already been found above: that, at RAF Coningsby, there is no evidence to suggest age plays a factor in acceptance of simulation. Similarly there was no evidence to suggest that levels of qualification at either base were related to the acceptance of simulation. Therefore *it can be stated that at both RAF Coningsby and Lossiemouth there was no evidence to support either of the newly forming beliefs stated in Objective 3.*

Objective 4. *Determine the minimum levels of LSB for each of the required tasks, as subjectively assessed by the present Typhoon pilots.*

This work proposed the use of the term Live Option Synthetic Blend (LOSB) to articulate the limit of exploitation of the live and synthetic domains, LSB being now proposed as a term for the commander's guidance on the particular balance he or she wishes to achieve within those limits, depending on the stresses and strains the Force is experiencing in the short term. For the first time a graphical representation of the LOSB for all tasks and skillsets at all of the levels of training has been made (Figure 4-30-Figure 4-32 and Figure Apx J-1 - Figure Apx J-6) in order to provide the commander with a detailed and flexible method of using the two environments that will withstand the stresses of deployment and exercise cycles. As an overview the LOSB percentage average for the level 1 Major Tasks was 49-27-24, for level 2 52-23-25 and for level 3 50-24-26, thus the ratio of minimum live events to synthetic was a consistent 2:1. If this is allied to the results found at Objective 2 it would suggest that the best way to increase a pilot's opinion of the acceptable LSB would be to employ the Major Tasks within a linked 4-ship tactical scenario. Combination of these figures allowed an equation that sought to express the planning requirement for the forthcoming year in order that the Force HQ might be able to plan better their fuel and technical spares requirements.

Finally the effect on the LOSB of purchasing immediate concurrency was investigated. In the major tasks of CA Day and Engage live flight training was not reduced but the *minimum* number of events that should be done in the synthetic environment did increase. The data shows that only in the subtask of 'Jamming' did concurrency produce a reduction of live flight required for training. This shows a perception of the worth of simulation but does not provide a sufficient reduction in live flight with which to fund it. Thus simulation would provide a qualitative increase that must be funded externally.

Objective 5. *Investigate the effect of threat complexity on the LSB.*

As the complexity increased from level 1 to 2 there were small increases in the minimum numbers of live events recommended, however these were only in 4 of the 8 Common Skills, none were seen in the Major Tasks or Sub Tasks. In contrast comparisons between the most difficult threat level and the lowest found an increase of 8% in the number of sorties recommended for all 3 of the Major Tasks, the recommendation for live sorties remained unchanged. The indication being that the sample believed that training for the highest threat levels was served effectively by an increase in the synthetic proportions. Thus complexity does affect the LOSB.

4.20 TRIAL PANDORA'S DIAMOND SUMMARY

Trial PANDORA'S DIAMOND has sought to provide evidence for the immediate exploitation of synthetics within the training mix based on opinion from an informed audience. The correlation between variables and the themes could be used to target simulation towards pilots of particular experience (age or flying hours) or alternatively alter the manner in which the simulators are used (using 4-ship simulation sorties rather than 2-ship). The trial has gone further and determined the Force's subjective view on the Live-Option-Synthetic Balance that exists for the present simulation systems at each of the Combat Ready levels. Finally the trial has determined the effect of concurrency on the utilisation of the simulator in a number of key areas at different CR levels in order to provide evidence as to the cost effectiveness of purchasing concurrency contracts in these areas.

Chapter 5. Project JENX

5.1 OVERVIEW

The time taken to train a pilot to Multi Role Combat Ready (MR CR) in 2014 was of the order of 11 months. Given this starting point it can be shown (see section 5.4) that the *entirety* of a squadron's effort is focused on training their replacements. There cannot be any 'excursionary skillset' training (training to threats not included in MRCR) because the regular opportunity cannot mathematically exist. The associated effect is to depress CR pilot numbers. Project JENX was a feasibility study that sought to determine if an OCU pilot could be trained to MR CR using lessons from the PANDORA'S series of trials, with particular interest in understanding the resource savings and effect on training time for the Combat Ready syllabus should a targeted synthetic training syllabus be introduced. As a secondary objective the project was intended to understand any deficiencies in the ability to replicate Combat Ready Training. Unlike the trials, however, the focus was not on the individual being trained but the feasibility of the syllabus and the ability of the squadron to support this training regime. From the outset it was intended that JENX would utilise only a single pilot but by completion of the trial a total of 3 personnel had been assigned to, and completed, the syllabus.

5.2 KEY LITERATURE REVIEW DIRECTION

In 1991 Houck (Houck et al. 1991) provided the initial assessment of areas where simulation provided better training than live flying, this was further expanded in 2006 by Portrey's work (Portrey et al. 2006). Trials PANDORA'S BUZZARD and DIAMOND led on to demonstrate the application of this work in the context of Typhoon training, highlighting the sizable contribution simulation can make to training of the combat pilot. In contrast, however, there has been little work on the benefits of simulation with respect to resources employed. This is largely due to the nature of the experimentation carried out to date; part-task assessments, use of non-front-line aircraft and prohibitive costs. This work is particularly relevant given the amount of research focusing on financial implications (Schank *et al.*, 2002; Harper and Hillier, 2007; Kruzins, 2008; Wells *et al.*, 2009), the project therefore blends the tactical pilot and operation commander's interests on capability with the financial need to provide savings and, through a determining of resources employed, allows an understanding of capacity when employing this capability.

5.3 AIMS

The third theme of the primary research question was focused on the ability to pool resources in order to conduct simulation for CR training (the section of training conducted immediately after joining the front-line from the OCU). The objectives were:

a. **Objective 1**. Construct and demonstrate proof of concept of a CR work-up syllabus that minimises the resource input of the front-line squadrons. Any trainees used must pass their CR test sortie.

b. **Objective 2**. Determine the levels of resources required to conduct this syllabus and any resource savings made by its methodology.

c. **Objective 3**. Determine if the resources available are sufficient to employ this syllabus across all CR work-up pilots.

5.4 ISSUE

Given that tour lengths are 2.5 years and a front-line squadron has 18 pilots, an average of 7 pilots will be replaced per year. Assuming the replacement pilots will be delivered for training at equally spaced intervals there will be 5 of the 18 pilots on CR work-up at any one time. Of the 13 CR pilots that remain to train them 6 will be on diversions (1 in the Falklands Islands, 2 on leave, 1 course, 1 QRA, 1 QRA stand down). Thus the 5 CR work-up pilots have a total of 7 qualified pilots available to instruct them. The ratio of instructor to student is 1:1 leaving 2 qualified pilots to run the positions of Authorisers Desk and Duty Executive.

Accepting that the most complex sortie is the graduation Tactical Check, conducted once every 6-8 weeks, it follows that a squadron must be operating at or below this level whilst training their replacements. Thus a squadron's capability is limited by the time to achieve CR for its new pilots.

Investigation of the logbooks for all the Multi-Role Combat Ready pilots (n=24) at RAF Coningsby found the time taken to achieve the qualification after arrival on the front line ranged between 220 and 459 days (95% CI) with a median of 328 days.

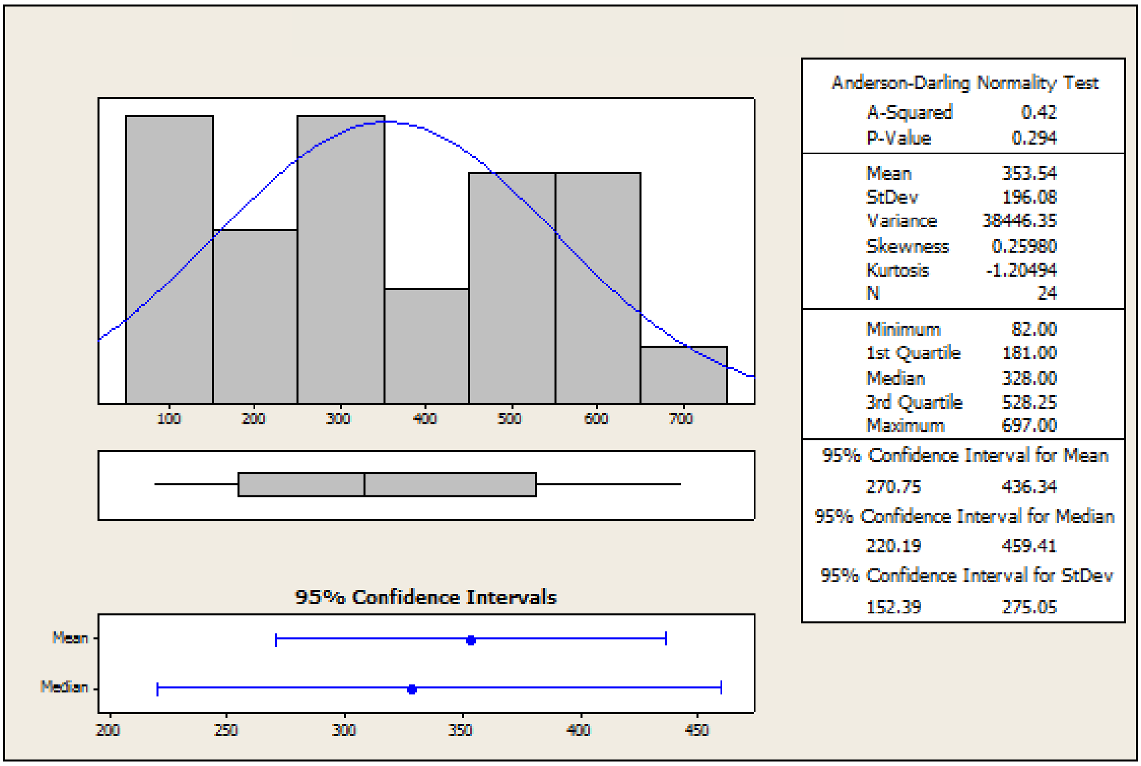

Figure 5-1. Time to Achieve Multi Role Combat Ready (Days).

Given this starting point it has been shown that the *entirety* of a squadron's effort is focused on training their replacements. There cannot be any 'excursionary skillset' (high level and complex threat) training because the regular opportunity cannot mathematically exist. The associated affect is to depress Combat Ready pilot numbers.

Project JENX intended to reduce the CR time to a maximum of 12 weeks, allowing the training time and assets returned to be used to increase Typhoon Force capability, capacity and efficiency.

5.5 PROJECT METHOD

Reduction in the JENX syllabus' CR timeline was to be achieved not through truncating the present syllabus but by porting those sorties that the simulator was capable of replicating into the synthetic domain[6]. The targeted use of simulation was intended to increase throughput by harnessing the following synthetic advantages:

a. **Configurable**. Numbers, replication and availability of radar equipped red air and Surface to Air Missiles (SAMs). Ensuring the trainees are always met with realistic enemy numbers of aircraft with correct numbers of Self Protection Jamming Pods and flight envelopes that are similar to the real aircraft. Additionally, environmental control of weather and light levels was possible in the simulator, with night sorties intended to be conducted without altering shift patterns or working days; sortie completion rates would also be unaffected by poor weather or weather could be artificially made worse to allow mandatory instrument approaches to be flown.

b. **Availability**. Consistent, forecast-able availability of devices. In contrast to the live flying that utilises 6 aircraft to generate 18 hours of training per day the simulators utilise 4 devices to achieve 35 hours. This allows reliable planning out to a realistic 4 week horizon which, allied to the present squadron manning tools, allows the correct allocation of trainee and student at least 2 weeks in advance.

[6] Identified through Trials PANDORA'S BUZZARD and PANDORA'S DIAMOND.

c. **Serviceability**. The simulator devices in use produce a consistent 10% failure rate, allowing mitigation to be built into the plan. The aircraft rate, however, was never consistent and was affected by issues of weather and spares availability. The effect of poor serviceability in the live environment is also greater with both the trainee and instructor's aircraft required to be serviceable as a minimum, with the other two formation members being a desirable requirement. In the simulator, however, it is possible to substitute a failed device with a computer generated entity, driven from the operating station. In such a manner the trainee is much less affected by any serviceability issues within their formation.

d. **Location.** Simulation gives the ability to determine fighting location and train tofor forthcoming exercises. Thus increasing the student's exposure to different areas of the UK and overseas. This allows the CR pilot to have a working knowledge, as they graduate, of Areas of Operation (AOs) the rest of their sqn will have experienced.

Compared to the standard syllabus LSB ratio of 66:34 the Project JENX Syllabus sought to utilise a LSB ratio of 30:70 which equates to approximately 21 live hours, although opportunity existed to fly outside the syllabus sorties. The syllabus was written to utilise no more than 4 of the 6 simulator slots allocated to each front-line squadron per day. For the initial trainee the supporting qualified pilot acting as instructor was intended to be the same for every sortie, so as to maximise feedback. They were programmed to achieve approximately 15 simulator sorties per month and these were intended to be aligned to the supporting pilot's annual task requirements (see Trial PANDORA'S DIAMOND) and the need to achieve 3 non-emergency based simulators per month.

The syllabus, included in Appendix K, provides the laydown of the sorties and briefs by day. Given the serviceability of the simulator this permitted programmers to bid for simulation assets 2 weeks or more in advance. It should be noted that provision was made for a catch-up simulator each Friday afternoon, which if not used was recycled for squadron emergency simulators. A fundamental difference between the two methods of training was that, within the day-to-day programming, Project JENX was deliberately prevented from the cultural norm of ‘sucking’ pilots out of the simulator into live flight in the case of sickness or manpower shortages. These short-notice changes had been witnessed in the past to lead to poor training value in the live environment and the loss of a completed syllabus sortie in the synthetic. In contrast Project JENX intended to maximise on the simulator’s availability and serviceability and by the inclusion of ‘catch-up ‘ sorties at the end of every week for the mitigation of any broken simulation devices the trainee would maintain the projected output date.

The test week at the end of the syllabus was intended to provide confidence to both the command chain and the subject that the CR MR standard was able to be met. As stated above, the sample size was too small to *guarantee* that all following subjects would pass, instead the project was intended to provide an understanding of resource usage with the result being indicative only.

5.6 RESULTS

The project commenced on 24 Oct 14 and used a single individual that exited the OCU during that same week. His training was conducted in the UK with 3(F) Squadron owing to his assigned squadron XI(F) participating in an exercise in Turkey. Access to large numbers of live assets was restricted by 3(F)’s training needs, nevertheless by week 4 he had completed the Air Defence Qualification, which from the logbook evidence took a median average of 202 days for his predecessors, and was commencing the Multi Role portion of instruction. The CR MR check ride was assessed completed with the Officer Commanding XI Squadron on 16 Jan 15 (Layden, 2015a), which was 12 calendar weeks after commencement of the trial. Allowing for the 2 week Christmas break this demonstrated that it was possible to complete the Project JENX syllabus in 10 weeks, to the Multi Role CR standard; in comparison the same standard was being achieved by the standard methodology in 328 days (σ=196).

The student pilot concerned passed his live MRCR sortie on his first attempt but it should be noted that despite all the Electronic Attack techniques being used by the red air only the techniques similar to the simulator made it to his side of the battlespace, thus not all techniques can be ratified as trained. Additionally the student had still to fly his live night test and night Air-to-Air Refuelling qualification (required for QRA) this is due to no live night phase being conducted during the time period and the simulator being unable to replicate air-to-air tanking of any sort.

5.7 ANALYSIS

Analysis of this project will take place in two parts firstly from the aspect of resource utilisation; and secondly the feedback from the single student will be considered in light of his performance and experience. The necessary data for these comparisons are held within Project JENX Supporting Information (Appendix K) that details the resource calculator for Project JENX, with the standard course for comparison. Additionally there is the comparison of resources saved.

5.7.1 LIVE SYNTHETIC BALANCE

Examination of Appendix K shows that Project JENX altered the Live Synthetic Balance (LSB) from 70:30 to 21:79. This is important with respect to the RAF's intended vision of an LSB of 50:50 as it demonstrates that the CR portion of Typhoon training can be completed 'synthetically heavy' and in doing so provides an opportunity to offset elements of training for which synthetics are not suitable or capacity in the simulator does not exist.

5.7.2 SAVINGS

The Project JENX syllabus saves a total of between 68 and 131 Typhoon flying hours per student pilot (the range of values being the difference between completing the standard syllabus using the stated resources or the minimum permitted). Using a nominal figure of 100 hours saved per pilot this can be represented in terms of marginal costs at £205,000 or the full cost capitalisation rate of £9,200,000 per pilot[7].

Financial representation has little immediate meaning at the tactical level, instead it may be more beneficial to represent this as engineering man hours saved. Present engineering man hours per flight hour stand at 13 and thus (assuming the nominal course saving of 100 hours) the Project JENX syllabus represents a reduction of 1300 engineering man hours. This saving can be reinvested back into either higher level training of front-line pilots or into 'softer' considerations such as leave, Adventurous Training or Force Development, all of which may go forward to tackling the increasing numbers of personnel seeking to leave the service prior to the end of their contract (MOD, 2016).

The number of DA20 and Hawk hours saved (up to 28 and 41 respectively) offers the opportunity for the savings to be repositioned to provide increased numbers of red air available to MRCR qualified pilots; allowing a numerically superior and more complex threat to be replicated.

5.7.3 SIMULATION ASSETS REQUIRED

An increase of 70 simulator hours was required to train under Project Jenx, bringing the total to 102 hours, equating to an extra 54 simulator slots that constructed an extra 24 missions. These slots were incorporated within the normal 4-6 slots allocated to front-line squadrons per day. No further simulator slots were used over and above the normal allocation and the Squadron still managed to maintain the 75:25 LSB directed, CR pilots being placed within

[7] Air Command HQ, *RAF Typhoon Capitation Rates* (High Wycombe: Royal Air Force 2012).

these missions as supporting crews. This demonstrates that capacity to train 2 pilots simultaneously, 1 on each of the front-line squadrons, exists at RAF Coningsby within existing resources. However, as the OCU currently trains 24 pilots a year, of which half are assigned to Coningsby squadrons, there would be a need to train 3 pilots per quarter. This is in excess of current capability and whilst it would be possible to make efficiency savings within the syllabus, e.g. training 2 trainee within the same mission, there would still be a required uptake in simulation capacity or a corresponding decrease in the demand from the OCU. Finally any increase above the recommended 2 pilots per quarter would affect the present 75:25 LSB on the front-line as more support pilots would be required to fly alongside the trainees in their synthetic sorties.

5.7.4 STUDENT FEEDBACK AND QUALITATIVE ASSESSMENT

Unedited feedback from the student pilot is held at Appendix K. Overall the student believed the course to be well planned and left him thoroughly prepared at the end of each phase. The tactical transition was easier than expected and the speed of the course had left him firmly ahead of his peers. Nevertheless there were areas of improvement:

- The long days at the beginning of the course where the subject was being asked to fly and simulated missions over and above the syllabus requirement. This detracted from the learning environment and made the student poorly prepared for the next event. These additional sorties were external to the syllabus.

- The reduction in priority during the Christmas run-up due to QRA currency considerations removed his available live-flight sorties, this was to affect g-tolerance when restarting post the Christmas break. However, this lack of priority reduced the sorties being conducted and indicates that in normal periods the training may be conducted more quickly.

- The Electronic Attack (EA) capabilities in the simulator are limited to a couple of techniques with the more advanced having to be intensively manipulated by the operators. EA exposure is therefore skewed towards certain techniques.

- The Laser Designation Pod differs from the aircraft, although the airborne re-familiarisation took only a couple of minutes. This may be longer for ab-initio pilots but is unlikely to be anything longer than a single sortie.

- An additional live sortie could have been included for ab-initio pilots during weeks 5 and 6 to ensure they remain current in the live environment.

5.8 CONCLUSIONS

Project JENX sought a new way of achieving CR in a shorter timescale and initial indications show the syllabus to be feasible although greater numbers are required to increase sample size. Nevertheless the savings attributable to the new methodology are considerable: ~100 flying hours per trainee, a consequent reduction in the engineering burden of 1300 hours per trainee, increase in front-line capability and capacity and allow CR work ups to continue during times of high fleet stress such as war or during deployment and recovery to exercises. Finally, the new syllabus will ensure a pilot spends approximately 9 months more of their tour being productive.

Objective 1. *Construct and demonstrate proof of concept of a CR work-up syllabus that minimises the resource input of the front-line squadrons. Any trainees used must pass their CR test sortie.*

Proof of concept completed and tested with trainee passing the relevant tests. Syllabus held in Appendix K.

Objective 2. *Determine the levels of resources required to conduct this syllabus and any resource savings made by its methodology.*

Resource requirements per syllabus trainee held at Appendix K. Savings of live flying hours are between 131 and 68 hours per trainee undergoing the course, flying hours saved from the DA20 (Red Air) contractor are between 3 and 28. The increase in simulation hours required is between 75 and 66.

Objective 3. *Determine if the resources available are sufficient to employ this syllabus across all CR work-up pilots*.

Resource exists to train 8 of the 12 pilots output by the OCU and assigned to RAF Coningsby. Increasing this number requires both an acceptance of a different LSB allocation to the front line, and an increase in synthetic capacity or a reduction in the OCU synthetic demand. Given that this present LSB (75:25) mirrors the minimum recommended by Trial PANDORA'S DIAMOND there is evidence to suggest that temporary synthetic increases would be accepted from a cultural viewpoint.

Chapter 6. DISCUSSION

This thesis began by breaking down the research question through a lens of 3 themes; culture, risk and performance, correspondingly those themes were broken down into objectives that underpinned each of the 3 research avenues. This approach was taken for two key reasons. Firstly, it was important that this research should be undertaken within a practical context, siding with Waag (1991, p4) rather than the quasi-transfer methodology preferred by others, this would allow access to the rare resources but more importantly ensure the standards on the trial were identical to those the RAF demanded of their current students rather than roughly translated approximations that were taken into a sterile experiment that would have been easy to dismiss by the user community. The method allowed peers a credible assessment of risk to life, risk to standards and risk to image (through failure). Secondly, the themed approach was taken to complement the strengths and weaknesses of each of the other trials. As an example, the small sample size of Trial PANDORA'S BUZZARD, whilst equal to the largest sample size of any contemporary trial and despite being more comprehensive in terms of tasks, skillsets and length, did still use only 8 trainee pilots. This, in turn, limited the fidelity of the answers that could be extracted from the results; the use of an all or nothing approach to recommendations in terms of trainee test ride results led to a lack of recommendation for the Combat phase. The inclusion of collective informed opinion from Trial PANDORAS DIAMOND allowed this finding to be reinforced together, the confidence in the recommendation is increased. This logic also reinforces the recommendation for the inclusion of the Counter Air phase as well as supporting the work on resources throughout all three trials.

The work has contributed to new knowledge in a number of areas. The trial to train OCU pilots in the simulator and test in live flight provided a European first: first solo from simulation for a front-line combat aircraft. Whilst this is supported in concept by the zero-flight-time airline training identified within the literature review it had not been thought transferable due to performance capabilities of modern fighter aircraft. This practical demonstration that knowledge acquired solely in the simulator could be translated into a valid performance in live flight challenges the RAF's standard methodology of using simulation simply to rehearse the next airborne sortie. This was reinforced in the successful Counter Air phase and to a lesser degree in the trial to train an OCU graduate to Multi Role Combat Ready and has already led to a wholesale rewrite of the OCU syllabus. An independent impression of the overall effect of the trial on the Typhoon Operational Conversion Unit is given by the Officer Commanding 29 Squadron in Appendix L.

The examination of synthetic balance within the front-line training based on collective opinion and the Force headquarters tasks requirement derived from the Mission Essential Competencies (Symons, France and Bell, 2006; Dstl, 2009) led to the LOSB (Live, Option, Synthetic Balance) method of representation and the first ever mapping of the pilot's annual task by training environment and complexity level. This offers the commander freedom to balance the Force's training, when resources are stressed, by altering the proportion undertaken in the simulator, all within culturally-accepted norms – that norm being approximately 50% conducted in a live environment, 25% in a synthetic and 25% available to be done in either. In this light it can be seen that the RAF's oft quoted 50:50 LSB ambition is achievable in the opinion of the Typhoon force at RAF Coningsby but it would be at the very limit of their recommendations and thus would be unlikely to be accepted on a permanent basis unless there was a cultural shift towards more acceptance of simulation.

The trials also produced the first attempt within the UK at understanding the fighter community's acceptance levels of simulation which, for both operating bases (Coningsby and Lossiemouth), was assessed as being only marginally positive. This piece of work is linked to the other trials through the theme of culture as it attempts to identify factors and areas of the pilot population that would accept more simulation in order that the synthetic assets may be targeted more efficiently. Producing the surprising observation that the squadron a pilot was assigned to had an effect was accompanied by a similarly unexpected outcome of a lack of evidence to support the view that younger pilots were more accepting of simulation. The first ever RAF investigation into the effect of threat complexity on training environment would also allow the correct targeting of synthetic resources; finding the most complex level of threat being recommended for more simulation over the lowest threat level.

The trial to reduce time to Multi-Role Combat Ready through heavy synthetic use, informed by the preceding trials, was the first of its type within the 5-nation Eurofighter community and demonstrated the effect of using synthetic strength to the maximum, namely CAST - Configurability, Availability and Serviceability to achieve speed or Throughput numbers. This methodology led to a reversal of the LSB, just as it had done in Trial PANDORA'S BUZZARD, and laid down a process-based syllabus that would output a trainee in a consistent time in between squadron deployments that would otherwise elongate this process. Whilst there are measureable cost savings associated with this work the intent was to generate equally-well trained pilots whilst creating spare capacity that could be reinvested into higher complexity training or alternatively reduce the burden on the Typhoon Force manpower by increasing leave, adventurous training or force development opportunities.

Not all elements of the trials have been successful but the observations from failure have led to a more informed position on the limitations of synthetics. The non-recommendation of the combat phase of Trial PANDORA'S BUZZARD owing to slow counters to the enemy manoeuvres, was identified as down to weaknesses of the visuals. Despite the failure, use of the simulator for conceptual training in the combat role was recognised, supporting the earliest of studies in the field (Pohlman and Reed, 1978; Payne, 1982). Even within the Counter Air phase that was recommended there were unexpected difficulties again associated with the visuals of the simulator, namely the effect of intrusive into-sun conditions and dissimilar representation of formation members between environments, both of which were to have a debilitating effect on the student's workcycle. This was further supported on a larger scale in the QRA phase that saw the inability of the visuals and radar software to represent real-world conditions found in the phase. These observations across a number of skillsets would have been unlikely to have been found in the increasingly popular (and cheaper) quasi-transfer trial and therefore they lend weight to the view of Bell & Waag (1998) that witnessing transfer to training within live flight is the only measure of establishing simulation effectiveness.

This research, therefore, has endeavoured to envelop the issue from three differing viewpoints and in doing so has led to new knowledge that has been employed by the peer community; re-issuing of the OCU and Combat Ready Syllabus (Layden, 2015b) and increased use of the synthetic assets up from 37% usage rate at the start of the research to 100% for FY 2015/16 (taken from Force Headquarters figures 23 Jul 16). Additionally the work has informed the future Typhoon simulation vision (Pemberton, 2014), leading to an investment in synthetics that will enhance collective training through the connection of the Air Battlespace Training Centre at RAF Waddington to new Typhoon simulators at the two bases. This will allow a distributed network that will promote training in the largest and most complicated environments with the UKs most important allies.

Chapter 7. CONCLUSIONS

The answer to the research question

" What is the maximum extent to which pilots of the Eurofighter Typhoon can incorporate synthetic training? "

has been demonstrated to be complex; bounded by cultural, resource and training location considerations. Broadly speaking increasing simulation from the lowest LSB of 75:25 on the front line will meet a cultural hurdle at 50:50 as well as a resource barrier for those training on the OCU. Correlations have been found between age, the manner in which simulation is used and the acceptance of simulation as a training environment that may offer insight as to where to target further increases. Nevertheless, increasing simulation past this point (50:50) is possible, with the maximum LSB achieved being 21:79 for the Multi-Role Combat Ready syllabus. However, heavy synthetic use requires targeting to particular simulators, requiring an understanding of their weaknesses and, strengths and equally importantly, knowledge of previous failings. Only in this manner can targeting simulation for heavy use within a syllabus be successful.

7.1 ACHIEVEMENTS

The aims of the research were to understand the limits of synthetics in terms of both training and culture as well as determining if it was resources that limited their use (SQ1,2 and 3). The work has achieved these aims and through the provision and delivery to Typhoon Force Headquarters of the first tested fully synthetic phases of Typhoon flight training, the first attempt to incorporate culture into longer term planning (LOSB equations of Section 4.17.2) and a conceptually tested training program, with resource assessment, to Combat Ready (Project JENX). In reaching these milestones the work has also achieved the first European first solo from simulation for a front line combat aircraft without the use of surrogate aircraft or use of the dedicated twin seat trainer. It is anticipated that when viewed as a whole these pieces of new knowledge will help inform the RAF with the creation of a ‘single seat OCU’ syllabus when the twin seat aircraft begin to be phased out in a few years time, as well as better planning and use of assets in the present day. Indeed the affect on the incorporation into the OCU of lessons already identified has led to the incorporation of the multi-role syllabus on to the OCU, whilst using extant resource, and a reduction of training time from 4 months to 4 weeks. The LSB has moved from 1:1 to 1:3 and the simulator usage has increased from less than 50% to over 95% (Statement of Officer Commanding 29 Sqn Typhoon OCU - Appendix L)

7.2 FURTHER WORK

Further work in this area should be centred around the issues of implementation; as simulation use increases resources in the live and synthetic domains will have to be managed to ensure all pilots receive training commensurate with their experience. As an example, extremely poor weather might preclude all but the most experienced going live flying whilst the less experience trainees stay on the ground without their valuable instructors available to take them into the simulator. Alternatively poor serviceability of aircraft might force the Squadron to choose between a pilot getting their live flying currency or participating in a simulator that will progress a trainee further down the syllabus. Further work examining the ability to inject simulation into live flight will allow these issues to be solved together, permitting formations split between the simulator and live flight to undertake simultaneous training no matter which domain they are in. The work should consider the synthetic replication of red air to those flying in the live environment, this would negate the need to choose between using valuable Typhoons or unrepresentative Hawk aircraft in day-to-day training. A final area to examine is the possibility of embedding synthetics further back in the training pipeline. Formations of trainees undergoing Advanced Fast Jet Training as part of larger mixed packages of Hawk and Typhoon aircraft would enhance training for all; giving

Typhoon pilots the escort aircraft they require for certain skillsets and the Hawk pilots an understanding of how to operate in larger formations.

The research has exposed some of the weaker areas of simulation, particularly the visuals, which could be used to progress an understanding of the fidelity required to ensure the skillsets do transfer into live flight. An area not covered during the synthetic training was the physiological effect of a heavy synthetic training regime. Might the performance of the students have been improved with physical coaching? Were the student's performances in the first live flights of each phase reduced due to he/she having to fight unanticipated physical demands? Does a highly synthetic syllabus increase the likelihood of neck injuries in live flight or can they be reduced through training and monitoring of head position in the simulator?

In the longer term this work can be incorporated within the developing realm of Live Virtual Constructive (LVC) synthetics: the injection of false (known as constructed) entities onto the displays and vision of the aircraft and pilot. These developments will allow simulation and live flight training to work together in a more symbiotic relationship to reduce resource use, improve output standards and increase the performance of students.

7.3 SUMMARY

In summary synthetic training is an underused resource that has the ability to reduce pressure on Typhoon resources and the potential to alter the manner in which air forces train, across the globe. Only through the continuous challenge of the status quo, intelligent implementation and a symbiotic relationship between industry and the military will this potential be realised.

REFERENCES

29 Squadron (2011a) 'Typhoon Operational Conversion Syllabus 2011'. RAF Coningsby: Typhoon Force HQ, (20111027–OPCON_2011_Syllabus_AL0–U).

29 Squadron (2011b) 'White Ticket Calculation - Hours and Instructors Required'. RAF Coningsby: Typhoon Force HQ, (20100403–FY1112 White ticket QPI plus Hrs).

Air Staff (2010) 'Air Staff Comment RE SDSR Sim Study'. RAF High Wycombe: Royal Air Force.

Alexander, A. L., Brunye, T., Sidman, J. and Weil, S. A. (2005) *From Gaming to Training: A review of Studies on Fidelity, Immersion, Presence, and Buy-in and Their Effects on Transfer in PC Based Simulations and Games*. Available at: http://www.aptima.com/publications/2005_Alexander_Brunye_Sidman_Weil.pdf (Accessed: 15 January 2014).

Anderson, B., Walls, K. and Read, P. (2011) *LiveWIRE - Post Trial Report*. Policy and Capability Studies, Portsdown West, Portsdown Hill Road, Fareham, PO17 6AD: DSTL.

Atkins, R. J., Lansdowne, A. T. G., Pfister, P. H. and Provost, S. C. (2002) 'Conversion between control mechanisms in simulated flight: an ab initio quasi-transfer study', *Australian Journal of Psychology*, 54(3), pp. 144–149.

Ausink, J. A., Taylor, W. W., Bigelow, J. H. and Brancato, K. (2011) *Investment Strategies for Improving Fifth-Generation Fighter Training*. 1776 Main Street, PO Box 2138, Santa Monica, CA, USA: RAND Corporation. Available at: http://www.rand.org.

Bacchetti, P. (2010) 'Current Sample Size Conventions: Flaws, Harms and Alternatives', *BMC Medicine*, 8(17), p. 17. doi: 10.1186/1741-7015-8-17.

Bacchetti, P., McCulloch, C. E. and Segal, M. R. (2008) 'Simple, Defensible Sample Sizes Based on Cost Efficiency', *Biometrics*, 64(2), pp. 577–594.

Baldwin, B. T. (2008) *Modified Flexible RAP: Linking Competencies and Experiences into Fighter Training and Readiness*. Maxwell Air Force Base, Alabama:

US Air Command and Staff College, Air University.

Bell, H. H. and Waag, W. L. (1998) ‘Evaluating the Effectiveness of Flight Simulators for Training Combat Skills: A Review’, *The International Journal of Aviation Psychology*, 8(3), pp. 223–242. doi: 10.1207/s15327108ijap0803_4.

Blyth, J. (2015) ‘Exploring the Balance of Live and Synthetic Training for Military Pilot Courses’, *MSc Thesis*. Cranfield University.

Bock, T. and Sergeant, J. (2002) ‘Small Sample Market Research’, *International Journal of Market Research*, 44(Collins 2000), p. 235.

Bryman, A. (2008) *Social Research Methods*. Oxford University Press.

CAA (2012) *FSTD User Approvals*. Available at: http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?type=simple (Accessed: 5 June 2013).

Chambers, M., Chong, M., Daniels, C., Davies-Knapp, G., Abraham-Araya, H. and Jackson, B. (2011) *Management of the Typhoon Project - Report by the National Audit Office*. London, UK: National Audit Office.

CINC-Air (2009) *Minimum Safe and Sustainable CFT (Live Crew Flying Task) for Fast Jet Operations*. MoD: RAF.

D’Alessandro, N. (2008) *Transference of PC-based simulation to mission training*, *Indian Aviation News*. InSite Solutions. Available at: net/allimages/AlessandroPaper. pdf (Accessed: 20 August 2011).

Dennis, K. a. and Harris, D. (1998) ‘Computer-Based Simulation as an Adjunct to Ab Initio Flight Training’, *The International Journal of Aviation Psychology*, 8(3), pp. 261–276. doi: 10.1207/s15327108ijap0803_6.

Dstl (2009) ‘Typhoon Mission Essential Competencies Summary Report’, *DSTL Report*. Farnborough, (DSTL No. 07-08-09).

Etherington, T. (2012) ‘RAF Typhoon Capitation Rates’, *FOI Response*. High Wycombe: Royal Air Force, (14-2-2012-155233–7).

FAA (2012) *FAA Regulations, Advanced Simulation Training*. Available at: http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?type=simple (Accessed: 12 December

2012).

Frazer-Nash (2010) ‘Simulation Assessment for the RAF’, *Frazer-Nash Report*. Stonebridge House, Dorking Business Park, Dorking, Surrey, RH4 1HJ: Fraser Nash Consultancy, (FNC 37508/36635R).

Gray, T. H. and Fuller, R. R. (1977) ‘Effects of Simulator Training and Platform Motion on Air-to-Surface Weapons Delivery Training’, *AFHRL Report*. Brooks Air Force Base: Air Force Human Resource Laboratory, (AFHRL-TR-77-29).

Hansard (2010) *Parliamentary Answers to Question by Peter Luff to Angus Robertson*. Available at: http://www.publications.parliament.uk/pa/cm201011/cmhansrd/cm100914/text/100914w0001.htm (Accessed: 29 August 2011).

Hansard (2011) *Evidence to Public Accounts Committee - MOD: Typhoon*. http://www.publications.parliament.uk/pa/cm201011/cmselect/cmpubacc/uc860-i/uc86001.htm. Available at: http://www.publications.parliament.uk/pa/cm201011/cmselect/cmpubacc/uc860-i/uc86001.htm (Accessed: 29 August 2011).

Harper, C. N. A. and Hillier, S. J. A. C. (2007) ‘RAF Air Synthetic Training Strategy’, *RAF Internal Paper*. Royal Air Force, (ID N/A).

Houck, M. R., Thomas, G. S. and Bell, H. H. (1991) ‘Training Evaluation of the F-15 Advanced Air Combat Simulation’, *ATRL Report*. Williams Air Force Base, Ariz.: USAF, Armstrong Laboratory, Aircrew Training Research Division, (AL-TP-1991).

JAA (2012) *Aeroplane Flight Simulation Devices (Testing)*. Available at: http://www.jaa.nl/publications/jars/JAR-FSTD-A_sec1_0508.pdf (Accessed: 1 May 2012).

Jacobs, J. W., Prince, C., Hays, R. T. and Salas, E. (1990) ‘A Meta-Analysis of the Flight Simulator Training Research’, *NTSC Report*. Orlando, (No. NTSC-89-006).

JTES (2009) *MoD Strategy for Simulation*. 20091214–MOD–Strategy–for Simulation–U. High Wycombe: Ministry of Defence.

Kline, P. (2000) *The Handbook of Psychological Testing*. London: Routledge.

Kruzins, E. (2008) *Defence Simulation Investment Reference Guide*. Department of Defence, Russell Offices [R1-3-B065], Canberra, ACT 2600: Australian Department of Defence, Simulation Office.

Laerd (2015a) *ANOVA Assumptions*. Available at: https://statistics.laerd.com/premium/spss/owa/one-way-anova-in-spss-7.php (Accessed: 30 December 2016).

Laerd (2015b) *Laerd Statistics*. Available at: https://statistics.laerd.com/premium/istt/independent-t-test-in-spss-9.php (Accessed: 5 December 2016).

Laerd (2015c) *One Way ANOVA Test, Laerd Statistics*. Available at: https://statistics.laerd.com/premium/stata/owa/one-way-anova-in-stata.php (Accessed: 30 December 2016).

Laerd (2015d) *Pearson Correlation*. Available at: https://statistics.laerd.com/premium/spss/pc/pearson-correlation-in-spss.php (Accessed: 30 December 2016).

Lard (2015) *T-Tests Leard Statistics*. Available at: https://statistics.laerd.com/premium/stata/pstt/paired-samples-t-test-in-stata-18.php (Accessed: 30 December 2016).

Layden, C. (2015a) ‘Project Jenx - OCXI Assessment’, *XI Sqn Report*. RAF Coningsby, (20150126–Project Jenx–OCXI).

Layden, C. (2015b) ‘XI(F) Squadron Recommendations Following Project Jenx’, *XI Sqn Report*. RAF Coningsby, (XI(F) Sqn/A7/Training-Air).

Lockwood, G. A. (2006) ‘The Assessment of Current and Future Synthetics to Augment and Enhance Conventional Flying Within the Typhoon Force’, *29 Squadron QWIC Report*. RAF Coningsby, (QWI_Cse1_Lockwood).

Macchiarella, N. D., Brady, T. and Arban, P. K. (2005) ‘High Fidelity Flight Training Devices in the Training of Ab Initio Flight Students’, in *The 24th Digital Avionics Systems Conference, 2005. DASC 2005.*, p. 5–B. doi: 10.1109/DASC.2005.1563375.

Magnusson, S. (2002) ‘Similarities and Differences in Psychophysiological Reactions Between Simulated and Real Air-to-Ground Missions’, *The International Journal of Aviation Psychology*, 12(1), pp. 49–61.

Marken, R. S., Taylor, W. W., Ausink, J. A., Hanser, L. M., Anderegg, C. R. and Wickman, L. (2007) *Absorbing and Developing Qualified Fighter Pilots (The Role of the Advanced Simulator)*. 1776 Main Street, P.O. Box 2138, Santa Monica, USA: RAND Corporation. Available at: http://www.rand.org.

McGrath, S. R. (2005) *Leveraging DMO’s Hi-Tech Simulation Against the F-16 Flying Training Gap*. Maxwell Air Force Base, Alabama: USAF Air Command and Staff College, Air University. Available at: https://www.afresearch.org/skins/RIMS/home.aspx.

Minitab (2012a) ‘Minitab Help: Mallows Cp’, p. Help page: R-squared. Available at: www.minitab.com.

Minitab (2012b) ‘Minitab Help: R-Squared’, p. Help page: R-squared. Available at: www.minitab.com.

Minitab (2012c) ‘Minitab Statistical Software’. Available at: www.minitab.com.

MoD (2006) *United Kingdom Glossary of Joint and Multinational Terms and Definitions*. Shrivenham, Swindon, Wiltshire, SN6 8RF: MOD - Development, Concepts and Doctrine Centre.

MoD (2010a) ‘JSP 550 Military Aviation Policy, Regulations and Directives’. Ministry of Defence, p. R375.105.2.

MoD (2010b) *SDSR Study 9.2 Simulation and Training*. London: Ministry of Defence.

MOD (2016) *UK armed forces monthly Service personnel statistics: May 2016*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/527838/Monthly_service_personnel_statistics-May_2016.pdf (Accessed: 16 June 2016).

Orlansky, J. and String, J. (1977) *Cost-Effectiveness of Flight Simulators for Military Training*. 400 Army-Navy Drive, Arlington, Virginia 22202: Institute for Defense

Analyses Science and Technology Division.

Payne, T. A. (1982) *Conducting Studies of Transfer of Learning: A Practical Guide*. Williams Air Force Base, Ariz.: Air Force Human Resources Laboratory, Operations Training Division.

Pemberton, G. (2014) *Typhoon Force Synthetic Vision*. RAF Coningsby: Typhoon Force HQ.

Pohlman, L. D. and Reed, J. C. (1978) *Air to Air combat skills: contribution of platform to initial testing*. Williams Air Force Base, Arizona: US Air Force Human Resources Laboratory.

Portrey, A. M., Keck, L. B. and Schreiber, B. T. (2006) *Challenges in Developing a Performance Measurement System for the Global Virtual Environment.* Human Effectiveness Directorate, Warfighting Readiness Research Division, 6030 South Kent Street, Mesa AZ, 85212-6061: Air Force Research Laboratory.

Roscoe, S. N. and Williges, B. H. (1980) 'Measurement of Transfer of Training', in *Aviation Psychology*. 2121 South State Avenue, Ames, IA, 50014-8300, USA: Iowa State University Press, p. 182.

Salas, E., Bowers, C. A. and Rhodenizer, L. (1998) 'It Is Not How Much You Have But How You Use It; Toward a Rational Use of Simulation to Support Aviation Training.', *International Journal of Aviation Psychology*, 8(3), pp. 197–208. doi: 10.1207/s15327108ijap0803_2.

Schank, J. F., Thie, H. J., Graff II, C. M., Beel, J., Sollinger, J., Graf, C. M., Beel, J. and Sollinger, J. (2002) 'Finding the Right Balance. Simulator and Live Training for Navy Units', *RAND/MR-1441*. Santa Monica, CA.: RAND Corporation.

Schank, J. F., Thie, H. J., Yardley, R. J., Taylor, W. W. and Grammich, C. A. (2009) 'RAAF F-35 Operational Pilot Training Live Training'.

Seaman, K. (1999) *Improving F15C Air Combat Training with Distributed Mission Training (DMT) Advanced Simulation*, *Training*. Maxwell Air Force Base, Alabama: USAF Air Command and Staff College, Air University. Available at: https://www.afresearch.org/skins/RIMS/home.aspx.

Symons, S., France, M. and Bell, J. (2006) *Linking Knowledge and Skills to Mission Essential Compency-Based Syllabus Development For Distributed Mission Operations*. Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ: Human Effectiveness Directorate.

Talleur, D. A., Taylor, H. L., Rantanen, E. and Bradshaw, G. L. (2003) 'Personal Computer Aviation Training Devices: Their Effectiveness for Maintaining Instrument Currency', *The International Journal of Aviation Psychology*, 13(4), p. 387.

Thompson, M. (2013) *Costly Flight Hours*, *Time.com*. Available at: http://nation.time.com/2013/04/02/costly-flight-hours/ (Accessed: 15 July 2016).

Vaden, E. a. and Hall, S. (2005) 'The Effect of Simulator Platform Motion on Pilot Training Transfer: A Meta Analysis', *The International Journal of Aviation Psychology*, 15(4), pp. 375–393. doi: 10.1207/s15327108ijap1504_5.

Waag, W. (1991) *The Value of Air Combat Simulation, Strong Opinions but Little Evidence*. Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061,USA: AFRL.

Wells, R., Franklin, A., Holden, R., Anderson, R., Jarvis, S., Kirby, B. and Simmonds, G. (2009) *The Optimal Blend for TornadoGR4 and Typhoon Post-OCU Aircrew Training in the Individual, Team and Collective Environments*. RAF High Wycombe: Headquarters Air Command.

Wikipedia (no date a) *Coefficent of Determination*, *Wikipedia, The Free Encyclopedia.* Available at: https://en.wikipedia.org/wiki/Coefficient_of_determination (Accessed: 30 December 2016).

Wikipedia (no date b) *Cronbach's Alpha*. Available at: http://en.wikipedia.org/w/index.php?title=Cronbach%2527s_alpha&oldid=653002619 (Accessed: 3 May 2015).

de Winter, J. C. F., Dodou, D. and Mulder, M. (2012) 'Training Effectiveness of Whole Body Flight Simulator Motion: A Comprehensive Meta-Analysis', *The International Journal of Aviation Psychology*, 22(2), pp. 164–183. doi: 10.1080/10508414.2012.663247.

Wong, K. (2010) 'Study Report for Phase 1 of ASTA TTF and ABTC Networked

Trial - Trial DevASTAtor', *QINETIQ/10/01413*. Farnborough: QinetiQ.

Appendix A SRF Supporting Documents

Table Apx A-1. Example of Grading Criteria for Overall Mark

	Score	Grade	Description	Action
Fail	0	*Unsafe (Re-train)*	The performance was unsafe or potentially dangerous. Any *element* marked "unsafe" will lead to an overall score of 0.	Remedial training **package** including at least a 're-fly' is required – student must be placed on Review
	1	*Below Average (Re-fly)*	Performance indicated a lack of ability or knowledge. At least one objective not achieved.	Remedial action required – student must be placed on Review
Pass	2	*Low Average*	The student's work was safe and ***satisfactory*** but with limited proficiency. Objectives achieved but with QPI prompting/assistance	Student will require remedial counseling and may benefit from Review action ***dependant on previous performance***
	3	*Average*	**Satisfactory**. Performance is almost correct. Student makes errors but ***recognizes and corrects them***	**This standard must be achieved in any dedicated End of Phase Check sortie to facilitate progress to the next stage of training**
	4	*High Average*	The student's work was correct, efficient and skillful. Only ***occasional non-critical errors*** were made.	
	5	*Above Average*	The student's work was of an unusually high degree of ability for this stage of training.	

Attachment 1: Level 3 Grading Criteria

1. **Mission Preparation**
Coordinated with QPI prior to briefing;
Demonstrated sufficient knowledge of TUG, TOPS, mission elements & SOPs prior to briefing;
SRF, PIFs and Reds & Greens reviewed and signed off;
Reviewed pertinent factors for flight: weather, NOTAMS, TOLD, fuel requirements, etc.

1a **TyAMPA Mission Planning**
Mission planning via applicable TyAMPA systems to ensure that all pertinent data transferred to data transfer device.

2. **SUTTO**

2a. **QRA Scramble Procedures**
Safely and efficiently performed scramble procedures to effect a timely departure IAW the brief.

3. **A/A Weapons Systems Checks**
Performed all necessary checks prior to combat IAW SOPs and TOPS; detected degraded systems and compensated appropriately.

4. **In Flight Checks**
Performed all necessary checks IAW applicable RTS, GASOs and local guidance to ensure safe flight operations.
4a. **G-Warm**
...and executed G-warm IAW current guidance and formation deconfliction.
4b. **AEA Check**
...and executed AEA check IAW current guidance and formation deconfliction.

5. **Fence Checks**
Properly configured the aircraft for combat IAW SOPs and TOPS.

Air Combat Fundamentals

6. **2 Ship Mutual Support**
Maintained element integrity throughout the mission in accordance with the brief, TUG and TOPS...
6a. **Skate**
...in a SKATE scenario.
6b. **Short Skate**
...in a Short Skate scenario.
6c. **Banzai**
...in a Banzai scenario.
6d. **Phase 4 VID**
...in a Banzai Phase 4 scenario.

7. **2 Ship Formation**
Maintained or regained formation position IAW the flight lead directive and Typhoon SOPs; demonstrated timely and smooth corrections.
7a. **CAP**
...while executing CAP
7b. **4 Ship Formation**
Maintained or regained formation position IAW the flight lead directive and Typhoon SOPs.

8. **2 Ship Night Operations**

8a. **Night Close Formation**
Maintained or regained formation position IAW the flight lead directive and Typhoon night SOPs; demonstrated timely and smooth corrections.

Figure Apx A-1. Example of Task Grading Criteria

Appendix B Descriptive Statistics for the Population

B.1 DESCRIPTIVE STATISTICS: IRT STD CSE

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Median	Maximum	Skewness
IRT Std Cse	56	1	3.732	0.100	0.751	1.000	4.000	5.000	-0.85

Variable	Kurtosis
IRT Std Cse	2.26

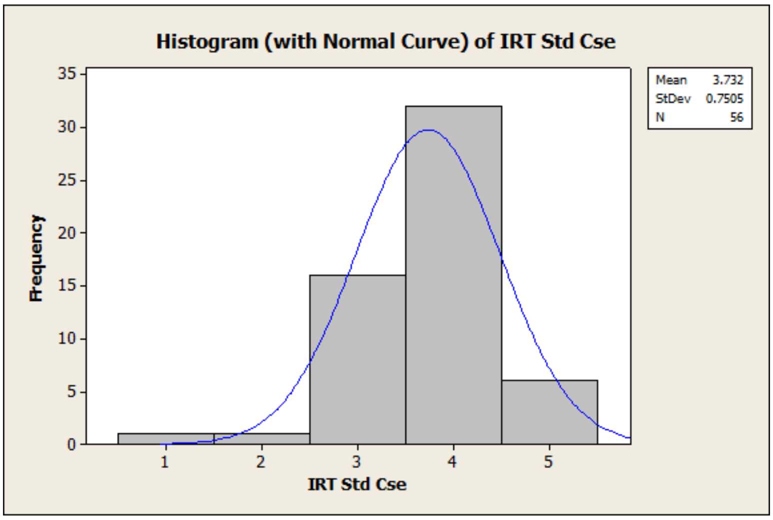

Figure Apx B-1. Histogram of IRT Std Cse Results

B.2 DESCRIPTIVE STATISTICS: BFM STD CSE

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Median	Maximum	Skewness
BFM Std Cse	57	0	3.404	0.106	0.799	1.000	4.000	4.000	-1.31

Variable	Kurtosis
BFM Std Cse	1.26

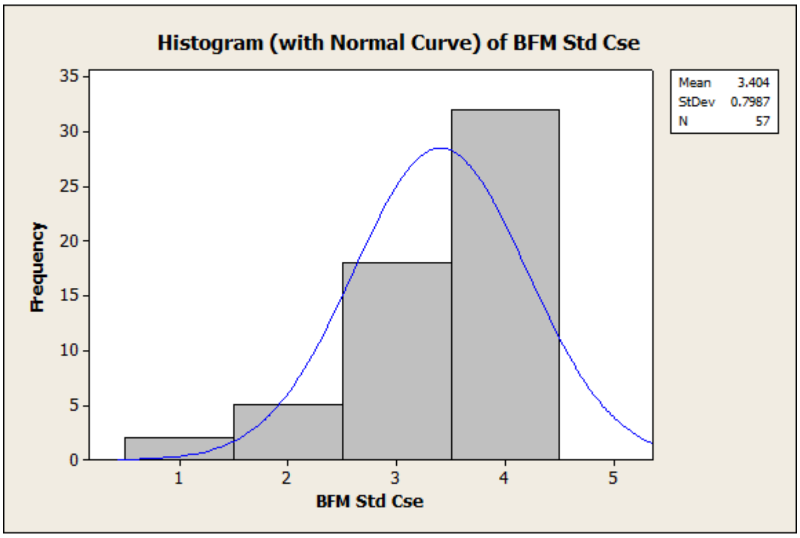

Figure Apx B-2. Histogram of BFM Cse Results

B.3 DESCRIPTIVE STATISTICS: CA STD CSE

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Median	Maximum	Skewness
CA Std Cse	57	0	3.842	0.119	0.902	0.000	4.000	5.000	-1.64

Variable	Kurtosis
CA Std Cse	5.08

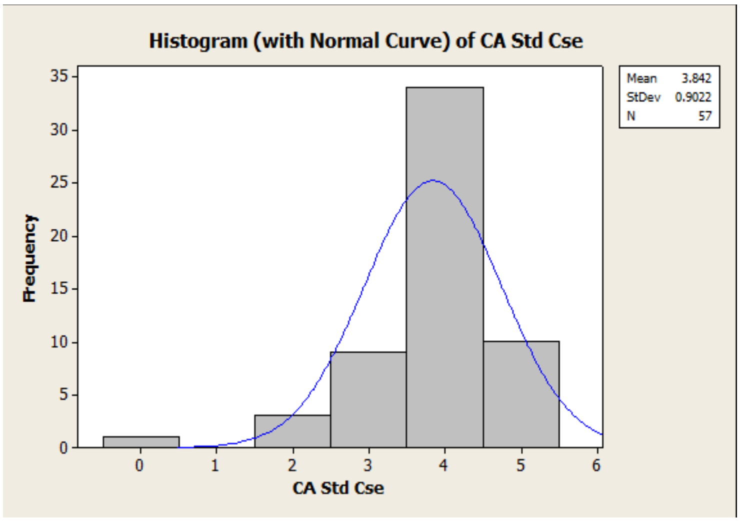

Figure Apx B-3. Histogram of Counter Air Std Cse Results

B.4 DESCRIPTIVE STATISTICS: LOW SLOW STD CSE

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Median	Maximum
Low Slow Std Cse	30	27	3.733	0.117	0.640	2.000	4.000	5.000

Variable	Skewness	Kurtosis
Low Slow Std Cse	-0.56	0.86

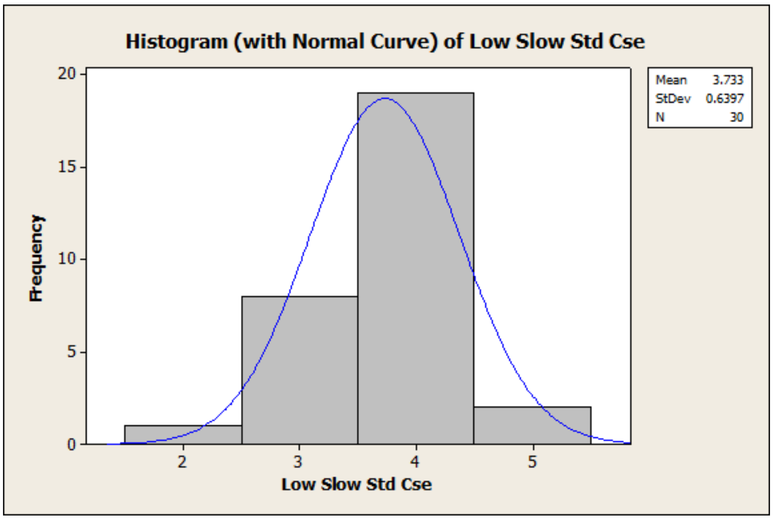

Figure Apx B-4. Histogram of QRA Low Slow Std Cse

B.5 DESCRIPTIVE STATISTICS: 3 PHASE SUM

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Median	Maximum	Skewness
3 Phase Sum	57	0	10.912	0.223	1.683	5.000	11.000	13.000	-1.26

Variable	Kurtosis
3 Phase Sum	2.22

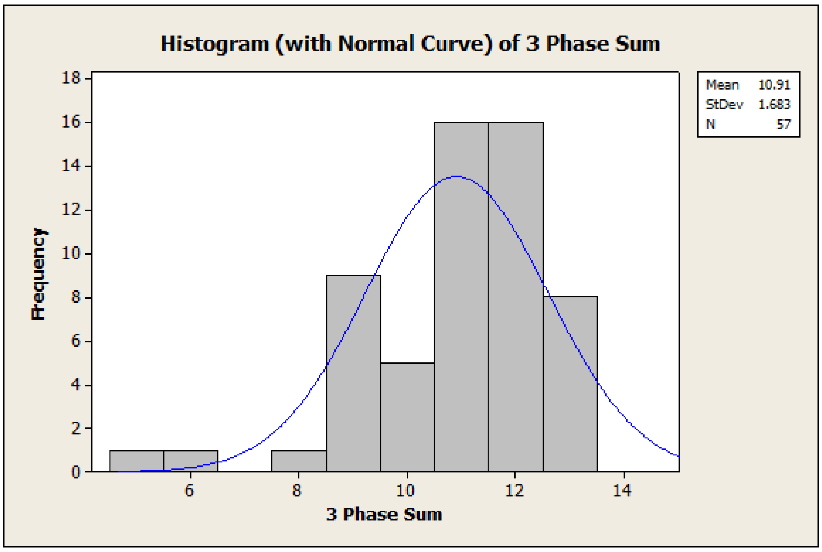

Figure Apx B-5. Histogram of 3 Phase Sum

B.6 DESCRIPTIVE STATISTICS: 4 PHASE SUM

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Median	Maximum	Skewness
4 Phase Sum	30	0	14.867	0.361	1.978	8.000	15.000	17.000	-1.43

Variable	Kurtosis
4 Phase Sum	3.58

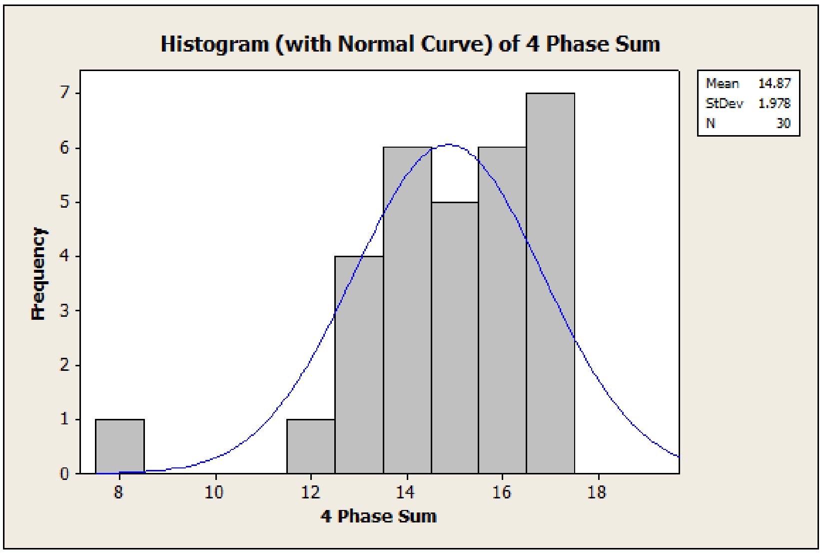

Figure Apx B-6. Histogram of 4 Phase Sum

Appendix C Risk Register Example

Table Apx C-1. Example Risk Register

ID	Date Entered	Risk	Mitigation	Likeness	Seriousness
1	**06/01/2011**	Malfunction during take off roll	Student completes emergency CTs and pre solo emerg sim. Level of knowledge is same as present student.	**Low**	**Medium**
2	**06/01/2011**	Heavy landing due to poor alpha Control	Technique will underline trend of alpha at 100'	**Medium**	**Low**
3	**08/01/2011**	SP G-LOCs during g- awareness	SP will be instructed to carry out g experience - a controlled pull up to but not exceeding 5g. This will specifically look for pressure in mask, LSJ and FAGTS.	**Low**	**High**
4	**08/01/2011**	Heavy landing caused by lack of simulator ability to replicate ground rush.	Technique taught does not use sense of ground rush. Landing technique is flown using 'numbers'.	**Low**	**Low**
5	**11/01/2011**	Aircraft damaged in crosswind landing.	Solo cross wind minima limited to 10 kts across. SP will conduct synthetic landings of up to 15 kts.	**Low**	**Medium**
6	**11/01/2011**	Student unable to generate a safe finals approach from finals turn.	If a safe touch and go has not been flown by div fuel + 400. SP will conduct short pattern GCA to land.	**Low**	**Medium**
7	**11/01/2011**	Incident due to error in SP checks.	SP is monitored by IP throughout synthetic phase. Sorties are end to end ie start up to shutdown. Risk is same as present methodology.	**Low**	**Medium**
8	**06/01/2011**	Student becomes lost during nav Route	ATC/ div airfields / Lon mil will be briefed on nature of solo. SP receives nav gndschl and practice exercises. SP briefed on use of guard / Buzzard. Chase ac available for first solo.	Low	Low
9	**06/01/2011**	Poor alpha control due to overly stable flight model and poor gust modelling leading to heavy landing.	IP will manually vary wind by +_ 3kts to achieve more realistic atmospheric variation.	**Medium**	**Low**
10	**07/01/2011**	Malfunction requiring need to Divert	Synthetic sorties include emergency handling, use of buzzard, frequency entry and practice diversions Chase ac available to lead to diversion.	**Low**	**Low**
11	**11/01/2011**	Loss of RADALT forcing ldg technique to use BARO - results in heavy landing.	SP will report loss of RADALT. Chase pilot briefed to find and report accurate QFE to SP.	**Low**	**Low**

Appendix D <u>Briefing Slide Example</u>

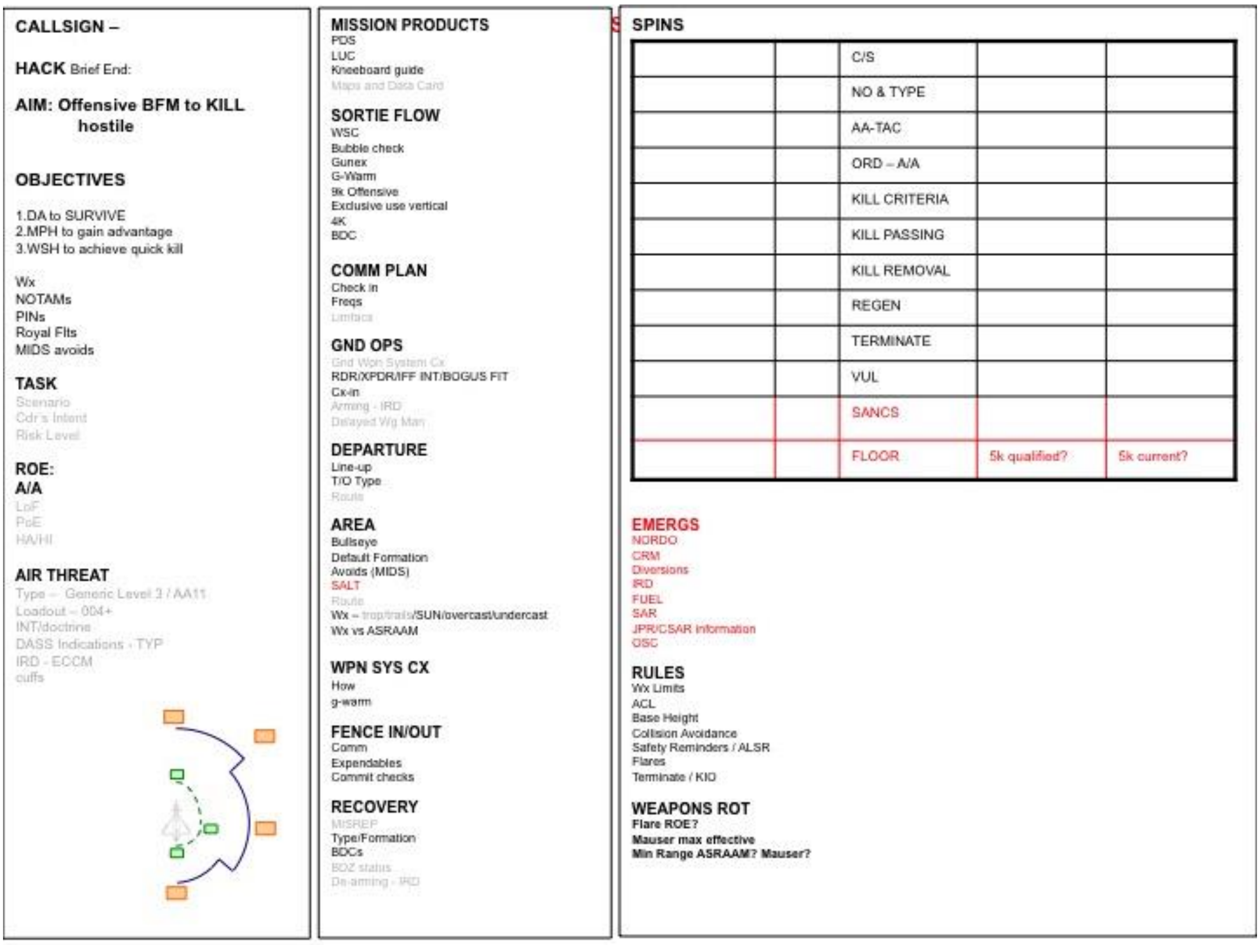

Figure Apx D-1. Example of Briefing Slide Used for Standardisation

Appendix E Approvals and Requests

MOD Research Ethics Committee (General)

Corporate Secretariat
Bldg 5, G01-614
Dstl Porton Down
Salisbury, Wiltshire
SP4 0JQ

Secretary: Marie Jones
telephone: 01980 658155
e-mail: mnjones@dstl.gov.uk

Sqn Ldr Alistair Allsop
SO2 TTF
RAF Coningsby
Tattershall
Lincs
LN4 4 SY

Ref: 1108/364

11th March 2011

Dear Sqn Ldr Allsop,

Re: Trial PANDORA'S BUZZARD Stages 1-5 version 1.1

Thank you for submitting this interesting protocol for ethical review and making minor amendments.

I am happy to give ethical approval for this research on behalf of the MOD Research Ethics Committee (General) and should be grateful if you would send me a copy of your final report on completion of the study. Please would you also send me a brief interim report in one year's time if the study is still ongoing.

This approval is conditional upon adherence to the protocol – please let me know if any amendment becomes necessary.

Yours sincerely,

Robert Linton

Dr Robert Linton
Chairman MOD Research Ethics Committee (General)

telephone: 020 8877 9329
e-mail: robert@foxlinton.org
mobile: 07764616756

Figure Apx E-1. MoD Ethic Committee Approval

18 Apr 11

Gp Capt Typhoon

Copy to:

Stn Cdr RAF Coningsby
SO2 Typhoon Trg

REQUEST TO PROGRESS TO TRIAL PANDORA'S BUZZARD STAGE 3

Reference: 20100526-Trial PANDORAS BUZZARD Tasking Letter-Gp Capt Typh.

1. It is formally requested that authority be granted to progress to Stage 3 (Combat) of Trial PANDORA'S BUZZARD in accordance with the ref above. Although Stage 1 is not yet complete Stage 3 will utilise the same students, thus early approval will allow unhindered progression through the syllabus. Stage 2 (Night solo) is intended to be postponed awaiting results and analysis of Stage 1 results.

2. The following individuals will be used in the trial, all are ab-initios / Hawk creamies.

a. Bewick, Flt Lt, allocated to OPCON AD 4 commencing 28 Feb 11.

b. Crickmore, Flt Lt allocated to OPCON AD 4 commencing 28 Feb 11.

c. Pucill, Flt Lt allocated to OPCON AD 4 commencing 28 Feb 11.

d. Saunders, Fg Off allocated to OPCON AD 4 commencing 28 Feb 11.

3. Approval is respectfully requested for the commencement of Parts 2 and 3 of the trial using the individuals and timeline above.

<original signed>

SEYMOUR A
Wg Cdr
OC 29(R) Sqn
95721 x 6332

Figure Apx E-2. Typhoon Force Approval Example

Appendix F Entry Questionnaire

QUESTIONNAIRE - BACKGROUND DATA Date: ____________

Name: ____________
Last 3: ____________
Squadron: ____________

Used to collate your responses only, will not be released.

Rank: ____________
Age: ____________

1. Hours Total (Nearest 50) ____________

2. Hours Typhoon (Nearest 50) ____________

3. Qualifications (Please circle one): Line Pilot QPI QWI

4. Are you an EWI-Typhoon (Please circle): YES NO

5. What is your CR level (Please Circle): LCR Q ADX Level 1 Level 2 ADX Level 2 MR Level 3

6. On average how many tactical 2 ship ASTA sorties do you do in a 2 month period? ______ *Insert number*

7. On average how many tactical 4 ship ASTA sorties do you do in a 2 month period? ______ *Insert number*

8. On balance when being tested on an IRT which environment is harder, ASTA or Aircraft? (Please Circle):

ASTA Aircraft

Entry Questions - Instructions

Please answer the questions on the next page.

There are 2 parts to each question.

For each question please indicate with a CROSS the degree to which YOU agree/disagree with the statement given.

For the second part please imagine the same question asked to *all* RAF Typhoon pilots in the Fleet and provide a cross in the box that represents the answer you belive the majority in the Fleet would have given.

Example:
Q1. All Typhoon pilots are incredibly good looking.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	X	☐
Majority of Ty Pilots	X				

Please turn over...

Questions:

1. In general the amount of day to day training in the ASTA simulator COULD be increased.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

2. In general the amount of day to day training in the ASTA simulator SHOULD be increased.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

3. ASTA Simulation in general provides battle winning training

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

4. Connected to a similar setup for land and sea the ASTA would provide battle winning training

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

5. The use of simulation in day to day training is too much

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

Please turn over…

6. Simulation can only be used for highly process driven and canned tasks.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

7. Simulation cannot replace live flight in any areas

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

8. Simulation can only be used to prepare for an airborne sortie of exactly the same profile.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

9. Simulation will never provide battle winning training.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

10. Simulation provides training I cannot get during day to day live training in the UK.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

11. My operational capability is enhanced by the inclusion of tactical sorties in the simulator.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

12. A reduction in live training in favour of simulation will have an overall negative effect on operational capability.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

13. I would be happy to swap one live event per month for a Joint Warfare simulator with the Army & Navy.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

14. A 4 ship Simulation sortie fighting a STANEVAL approved threat, using artificial intelligence to mimic known tactics and shots, would be a valid way of testing my abilities.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

15. Testing of my tactical abilities can only take place in the air.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐
Majority of Ty Pilots					

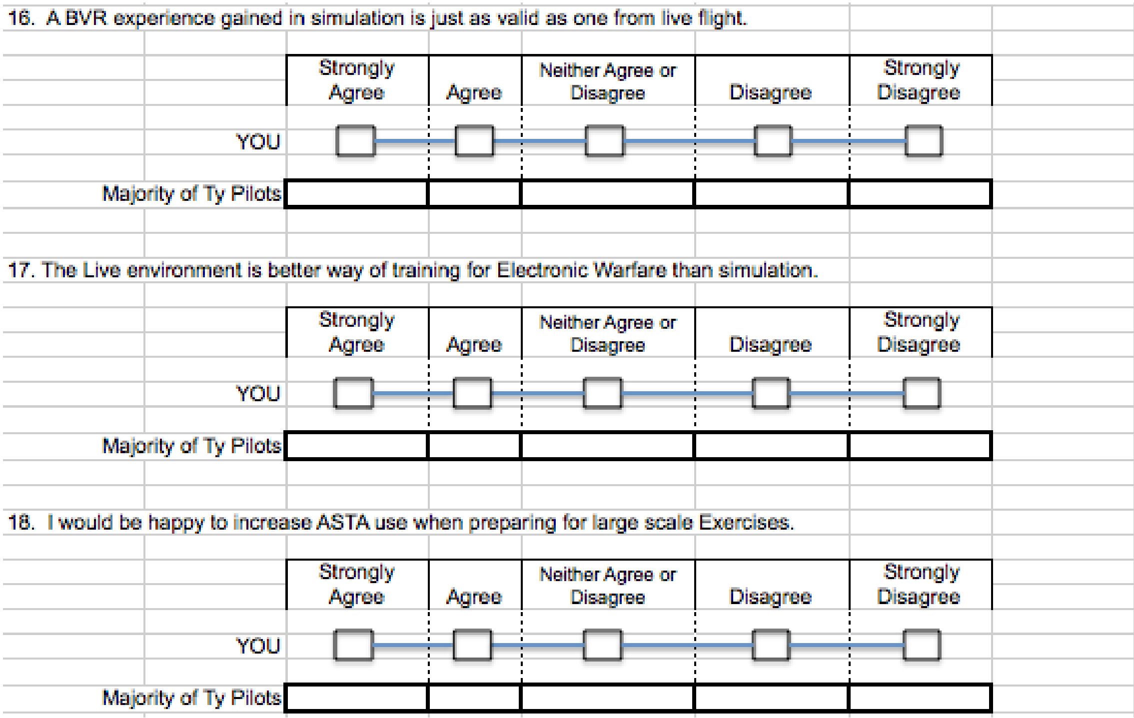

Figure Apx F-1. Entry Questionnaire - Trial PANDORAS DIAMOND

The table below provides the themes that each of the questions sought to address.

Table Apx F-1. Entry Questionnaire. Questions v Themes

Question No	Question	Theme
1	In general the amount of day to day training in the ASTA simulator COULD be increased.	Cultural lean
2	In general the amount of day to day training in the ASTA simulator SHOULD be increased.	Cultural lean
3	ASTA Simulation in general provides battle winning training	Near future
4	Connected to a similar setup for land and sea the ASTA would provide battle winning training	Near future
5	The use of simulation in day to day training is too much	Cultural lean
6	Simulation can only be used for highly process driven and canned tasks.	Simulation can provide

		experience
7	Simulation cannot replace live flight in any areas	Cultural lean
8	Simulation can only be used to prepare for an airborne sortie of exactly the same profile.	Simulation can provide experience
9	Simulation will never provide battle winning training.	Near future
10	Simulation provides training I cannot get during day to day live training in the UK.	Simulation can provide experience
11	My operational capability is enhanced by the inclusion of tactical sorties in the simulator.	Cultural lean
12	A reduction in live training in favour of simulation will have an overall negative effect on operational capability.	Cultural lean
13	I would be happy to swap one live event per month for a Joint Warfare simulator with the Army & Navy.	Near future
14	A 4 ship Simulation sortie fighting a STANEVAL approved threat, using artificial intelligence to mimic known tactics and shots, would be a valid way of testing my abilities.	Near future
15	Testing of my tactical abilities can only take place in the air.	Near future
16	1. A BVR experience gained in simulation is just as valid as one from live flight.	Simulation can provide experience
17	The Live environment is better way of training for Electronic Warfare than simulation.	Simulation can provide experience
18	I would be happy to increase ASTA use when preparing for large scale Exercises.	Simulation can provide experience

Appendix G Post Sortie Questionnaire

Name:				Date:	
Last 3:					
Which Device were you in ? (Please circle)			FMS	CT	

QUESTION 1. Below are a list of **EVENTS** along with the **NUMBER of times** that event must be conducted in a year according to TACTS.

You are challenged to provide the best blend of training.

Please provide, in your opinion, the minimum number that should be conducted in ASTA and the minimum number the should be conducted in LIVE flight for each of the **EVENTS**, for a Level 1 threat environment. The total does not have to equal the **NUMBER OF TIMES** - it is assumed that the remaining can be conducted in eithier ASTA or LIVE.

EXAMPLE

EVENTS	NUMBER of times	Minimum ASTA	MinimumLive
FQMD	10	2	2

EVENTS	NUMBER of times	Minimum ASTA	Minimum Live
Counter Air Day	112		
Counter Air Night	4		
DCA CAP Profile	36		
Intercept / Engage Day	48		
Intercept / Engage Night	24		
Defensive Aids Trg (ESM) - DASS	96		
Defensive Aids Trg (IR)- Flares	14		
Defensive Aids Trg (ECM)	25		
Defensive Aids Trg (RF) - Chaff*	10		
Enemy EW-Radar Jamming	6		
Enemy EW-IR Flares	12		
Enemy EW-Chaff	16		
DACT*	10		
DATALINK	48		
Helo Intercepts	4		
High Level Intercepts (HFF)	2		
Intervene	1		
Long Range Aviation Intercept*	10		
Low Fast Intercepts	1		

EVENTS		NUMBER of times		Minimum ASTA	Minimum Live
Low Slow Intercepts		2			
Phase 1 VID		4			
Phase 2 VID		4			
Phase 4 Offensive VID		4			
Phase 5 Passing VID LL		2			
Scramble Procedures		2			
Shadow		48			
Low Level Flying		24			
Night Landing*		10			
Night Tactical Formation*		10			
Pairs Landing*		10			
Pairs T/O*		10			
ACT		12			
Air-Air Gunnery		1			
DCA HVAD profile Day		2			
DCA HVAD profile Night		1			

* A TACTIS Requirement but with no stated number of times. The substituted number, usually 10, is to provide a picture of perceived ASTA capability in this area.

QUESTION 2. Finally please answer the same question for the subjects below but this time ASSUME THE SIMULATOR IS 100% CONCURRENT WITH THE FE@R AIRCRAFT.

EVENTS		NUMBER of times		Minimum ASTA	Minimum Live
Counter Air Day		112			
Intercept / Engage Day		48			
Enemy EW - Radar Jamming*		60			

Figure Apx G-1. Post Sortie Questionnaire - Trial PANDORA'S DIAMOND

Appendix H Exit Questionnaire

EXIT QUESTIONNAIRE Date: ____

Name: ____
Last 3: ____
Squadron: ____

Used to collate your responses only, will not be released.

Exit Questions - Instructions

Having completed the synthetic serials please answer the following questions.

For each question please indicate with a CROSS the degree to which YOU agree/disagree with the statement

Example:
Q1. All Typhoon pilots are incredibly good looking.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☒	☐

Questions:

1. In general the amount of day to day training in the ASTA simulator COULD be increased.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

2. In general the amount of day to day training in the ASTA simulator SHOULD be increased.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

3. ASTA Simulation in general provides battle winning training

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

Please turn over.

4. Connected to a similar setup for land and sea the ASTA would provide battle winning training

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

5. The use of simulation in day to day training is too much

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

6. Simulation can only be used for highly process driven and canned tasks.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

7. Simulation cannot replace live flight in any areas

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

8. Simulation can only be used to prepare for an airborne sortie of exactly the same profile.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

9. Simulation will never provide battle winning training.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

10. Simulation provides training I cannot get during day to day live training in the UK.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

Please turn over…

11. My operational capability is enhanced by the inclusion of tactical sorties in the simulator.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

12. A reduction in live training in favour of simulation will have an overall negative effect on operational capability.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

13. I would be happy to swap one live event per month for a Joint Warfare simulator with the Army & Navy.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

14. A 4 ship Simulation sortie fighting a STANEVAL approved threat, using artificial intelligence to mimic known tactics and shots, would be a valid way of testing my abilities.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

15. Testing of my tactical abilities can only take place in the air.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

16. A BVR experience gained in simulation is just as valid as one from live flight.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

Please turn over…

17. The Live environment is better way of training for Electronic Warfare than simulation.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

18. I would be happy to increase ASTA use when preparing for large scale Exercises.

	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
YOU	☐	☐	☐	☐	☐

Figure Apx H-1. Exit Questionnaire - Trial PANDORA'S DIAMOND

Appendix I <u>Statistical Tests and Results Explained</u>

This Appendix expands on some of the presentation of the results and tests used. Firstly the results referred to in Section 4.13:

‘ Thus, a statistically significant difference of 2.125 was demonstrated (95% CI, 1.237 to 3.013), t(47) = 4.82, p<0.001, d = 0.70 ’

1. 95% CI, 1.237 to 3.013 - the difference that was stated as 2.125 for this test can, for the population, be assured (to a confidence of 95%) to be between 1.237 to 3.013.

2. t(47) = 4.82 – the ‘t’ shows that the comparison for this test is against a t distribution whilst the ‘(47)’ denotes the degrees of freedom of the test, which is one less than the sample used. ‘4.82’ is the t-value obtained from the statistical look up tables.

3. p<0.001 – is the probability of returning the t-value stated if the null hypothesis was correct.

4. d = 0.70 is the *effect size* and is becoming increasingly requested for journals (Lard, 2015) hence its inclusion here. Known as *Cohen’s d* it is calculated, for a pairs t test, by dividing the mean difference between the two groups by the standard deviation of that difference. Nominally the values returned can be thought of as 0.2 – low, 0.5 – medium and 0.8 – high strength.

<u>Types of Tests</u>

The **One-way ANOVA** (ANalysis Of VAriance) test referred to in Section 4.14 is a method of determining if there are any statistical differences between two or more independent groups (Laerd, 2015c).

A **boxplot** as referred to in section 4.14.1 is a method of comparing sample distributions the components of which are shown in **Error! Reference**

source not found.. In particular boxplots were used to determine if there were any outliers associated with the data collected.

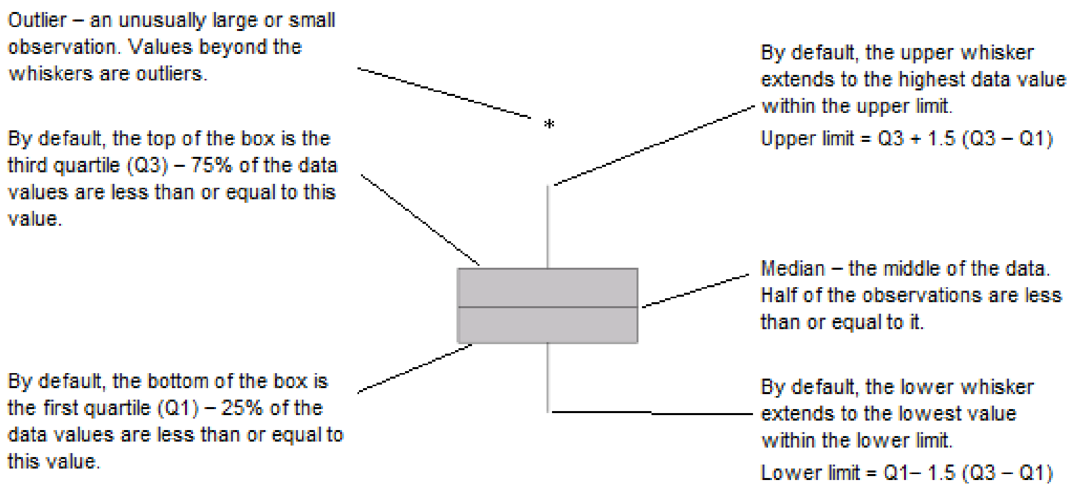

Figure Apx I-1. Components of a Boxplot. Reproduced from the Help Section of Minitab 2012.

Levene's test (section 4.14.1) is a method of checking that the variances of two or more groups are the same i.e. that there is homogeneity of variances. Tests such as the one way ANOVA are sensitive to unequal variances, such as those that occur within groups that significantly differ in size, and as such it is an assumption that the tests for homogeneity of variance and normality along with a boxplot examination have been run prior to an ANOVA being commenced (Laerd, 2015a).

Pearson's product moment correlation measures the direction and strength of the relationship between two variables that are continuous in nature the relationship is assumed to be linear. The correlation value (denoted as *r*) returns a value between the two extremes of +1, a perfect positive relationship, and -1, a perfect negative correlation (Laerd, 2015d).

R-Sq (adj) from section 4.16.8 and is a truncated form of 'R-Squared'. The figure provided indicates the amount of variance in the dependant variable (in the given section this is either the 'Cultural Lean' or 'Near future') that is predicted from the independent variable. An alternative view is that it attempts to indicate how well a model might predict a response for a given observation. R-Squared (adjusted) is the attempt to allow for the increase in the R-Sq value when multiple variables are added to the model (Wikipedia, no date a; Minitab, 2012b).

Mallows Cp used in section 4.16.8 assists with determining the best model that provides the best balance of bias and precision against the number of predictors used. Too many predictors may result in a model that in imprecise whilst too few may be biased. In practice a model Mallows Cp number that is approximately equal to the number of predictors is optimum choice (Minitab, 2012a).

The **non-parametric Wilcoxon signed rank test** used in section 4.17.3 is the non-parametric version of the paired samples t-test. As it is a non-parametric test the data used is not required to pass a normality test correspondingly, it returns the median difference between paired observations.

Appendix J PANDORAS DIAMOND Supporting Tables and Figures

Table Apx J-1. LOS Blend for Level 2

Major Tasks	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
		Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synthetic	
AI	4	2			2			N/A	N/A	Mode	50	0	50	25.0
ML CAS	12	8	6.0, 10.0	2.7	2	2.0, 4.0	2.3	0.03	<0.01	Median	67	17	17	8.3
LL CAS	4	4			0			N/A	N/A	Mode	100	0	0	25.0
CAS Night	4	3			1			N/A	N/A	Mode	75	0	25	25.0
ECAS	4	2			0			N/A	N/A	Mode	50	50	0	25.0
SCAR	10	5.5	4.0, 6.0	2.9	2	2.0, 2.0	2.6	0.1	<0.01	Median	55	25	20	10.0
SOF	4	2			0			N/A	<0.1	Mode	50	50	0	25.0
TST	2	1			1			N/A	N/A	Mode	50	0	50	50.0
OCA Self Escort	12	6	6.0, 8.0	2.4	4	3.7, 5.0	2	<0.01	<0.01	Median	50	17	33	8.3
Engage Day	48	24	20.0, 30.0	10.9	10	10.0, 15.0	7.2	0.08	0.12	Median	50	29	21	2.1
CA Day	112	53.5	50, 60	21.9	30	20.0, 30.0	13.9	0.06	0.19	Median	48	25	27	0.9
HVAD Day	2	1			1			N/A	N/A	Mode	50	0	50	50.0
Engage Night	24	12	12, 14.3	5.8	6	4.0, 10.0	4.1	0.06	0.03	Median	50	25	25	4.2
CA Night	4	4			0			N/A	N/A	Mode	100	0	0	25.0
HVAD Night	1	1			0			N/A	N/A	Mode	100	0	0	100.0
OCA Slow Mover	2	1			1			N/A	N/A	Mode	50	0	50	50.0
OCA Sweep	2	1			1			N/A	N/A	Mode	50	0	50	50.0
OCA Screen	12	6	6.0, 8.0	3.1	4	2.0, 4.0	2	0.12	<0.01	Median	50	17	33	8.3
Total LOSB	263	137			65						52	23	25	

Major Task	Sub Task	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
			Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synthetic	
AI	LDP	4	4			0			N/A	N/A	Mode	100	0	0	25.0
	EPW	12	6	6.0, 6.0	2.9	4.5	3.0, 6.0	2.8	<0.01	<0.01	Median	50	13	38	8.3
	PWII	12	6	6.0, 6.0	3	4	2.7, 6.0	2.5	<0.01	<0.01	Median	50	17	33	8.3
Counter Air	ACT	12	12	11.7, 12.0	1.3	0	0.0, 0.0	0.8	<0.01	<0.01	Median	100	0	0	8.3
	DACT	10	10	8.0, 10.0	1.8	0	0.0, 0.0	0.8	<0.01	<0.01	Median	100	0	0	10.0

Common Skills	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
		Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synthetic	
DA ECM	25	12	10.0, 15.0	6.6	5	4.7, 10.0	4.8	0.01	<0.01	Median	48	32	20	4.0
DA Flares	14	8	5.7, 10.0	4.3	4	1.0, 5.0	3.1	0.06	<0.01	Median	57	14	29	7.1
DA DASS	96	40	28.7, 48.0	19.5	20	120, 30.0	17.9	<0.01	0.11	Mean	42	38	21	1.0
DA Chaff	10	6	5.0, 7.0	3.3	2	1.0, 5.0	3	0.03	<0.01	Mean	60	20	20	10.0
NME Chaff	16	8	6.0, 12.0	4.7	2.5	21.0, 4.3	2.9	0.02	<0.01	Median	50	34	16	6.3
NME Flare	12	6	5.0, 10.0	4	2	1.0, 3.3	2.4	>0.1	0.02	Median	50	33	17	8.3
NME Jam	6	4	3.0, 5.0	1.5	2	0.74, 2.0	1.4	<0.01	<0.01	Median	67	0	33	16.7
NME Comms Jam	12	6	6.0, 8.0	3.4	2	1.7, 4.0	2.7	<0.01	<0.01	Median	50	33	17	8.3
CAP	36	12	10.0, 18.5	8.6	10	4.0, 10.0	6.54	0.03	<0.01	Median	33	39	28	2.8
Datalink	48	24	17.8, 24.0	12.9	11	9.5, 21.1	11.1	<0.01	<0.01	Median	50	27	23	2.1
Dragon High	10	5	5.0, 6.3	2.6	4	2.0, 5.0	2.3	<0.01	<0.01	Median	50	10	40	10.0
Dragon Low	12	6	6.0, 8.0	2.8	4	2.0, 5.0	2.8	<0.01	<0.01	Median	50	17	33	8.3
Air C2	24	12	10.0, 12.5	6.5	4.5	1.5, 6.5	6	<0.01	0.04	Median	50	31	19	4.2

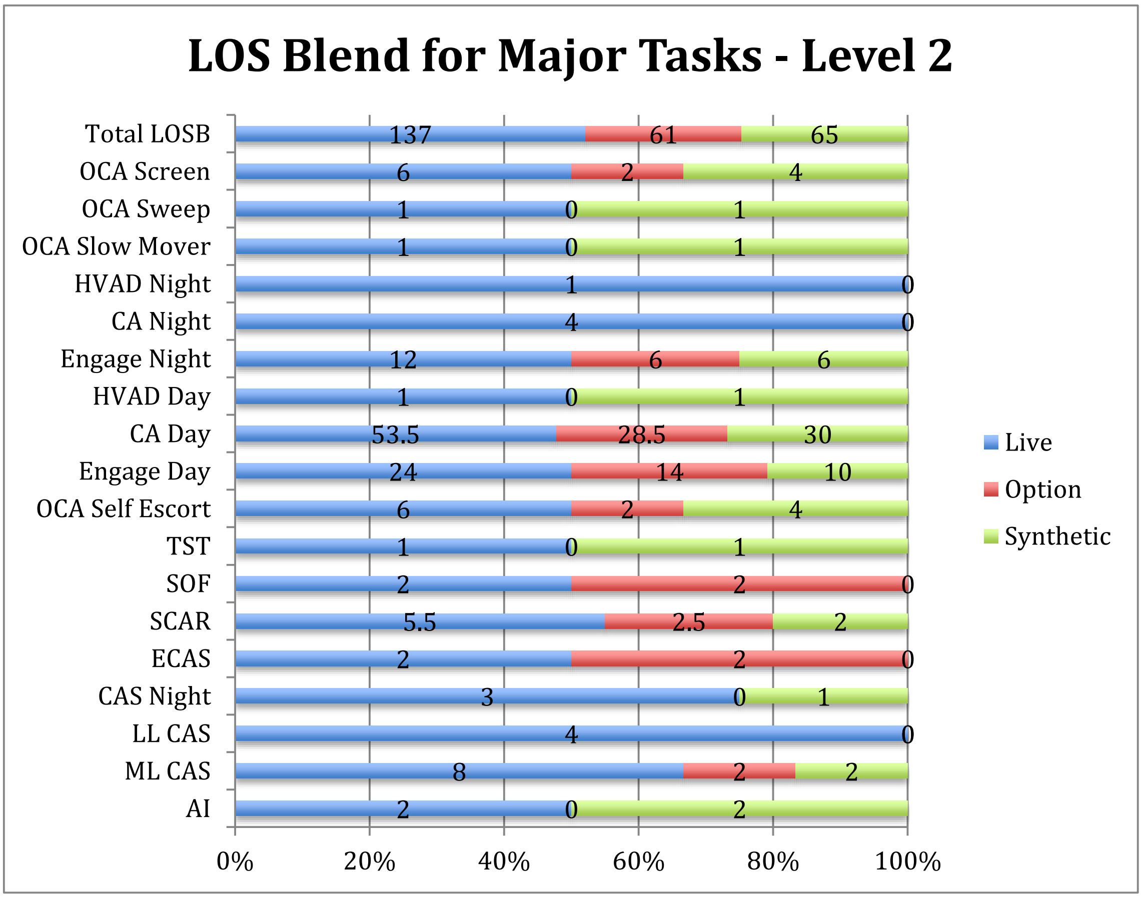

Figure Apx J-1. LOS Blend for Major Tasks - Level 2

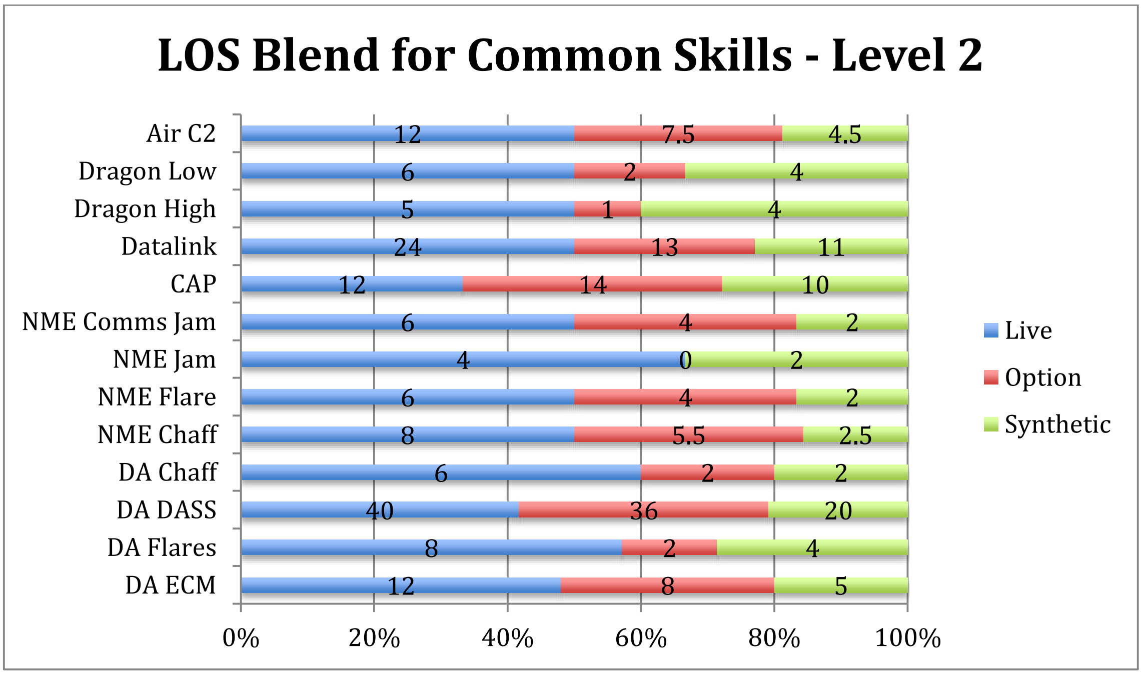

Figure Apx J-2. LOS Blend for Common Skills - Level 2

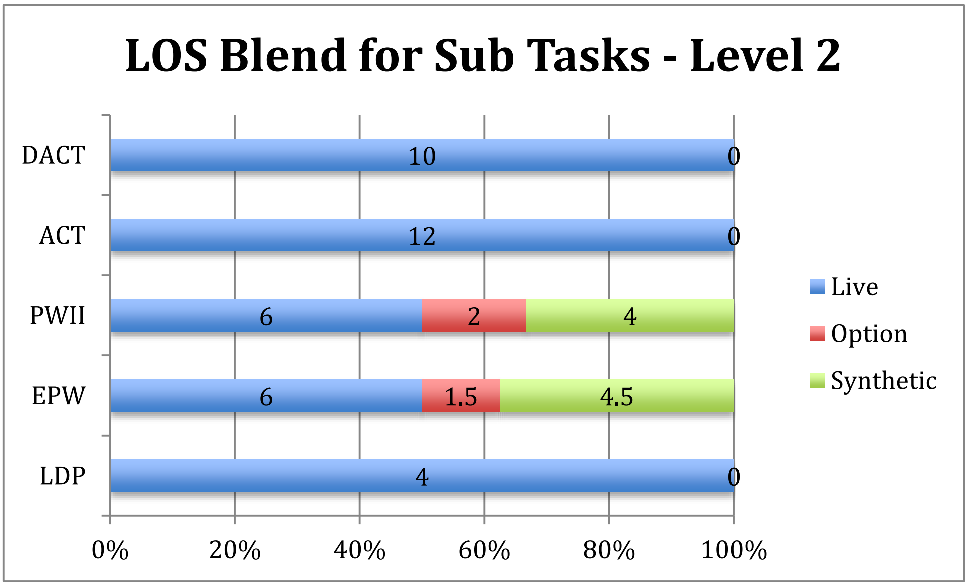

Figure Apx J-3. LOS Blend for Sub Tasks - Level 2

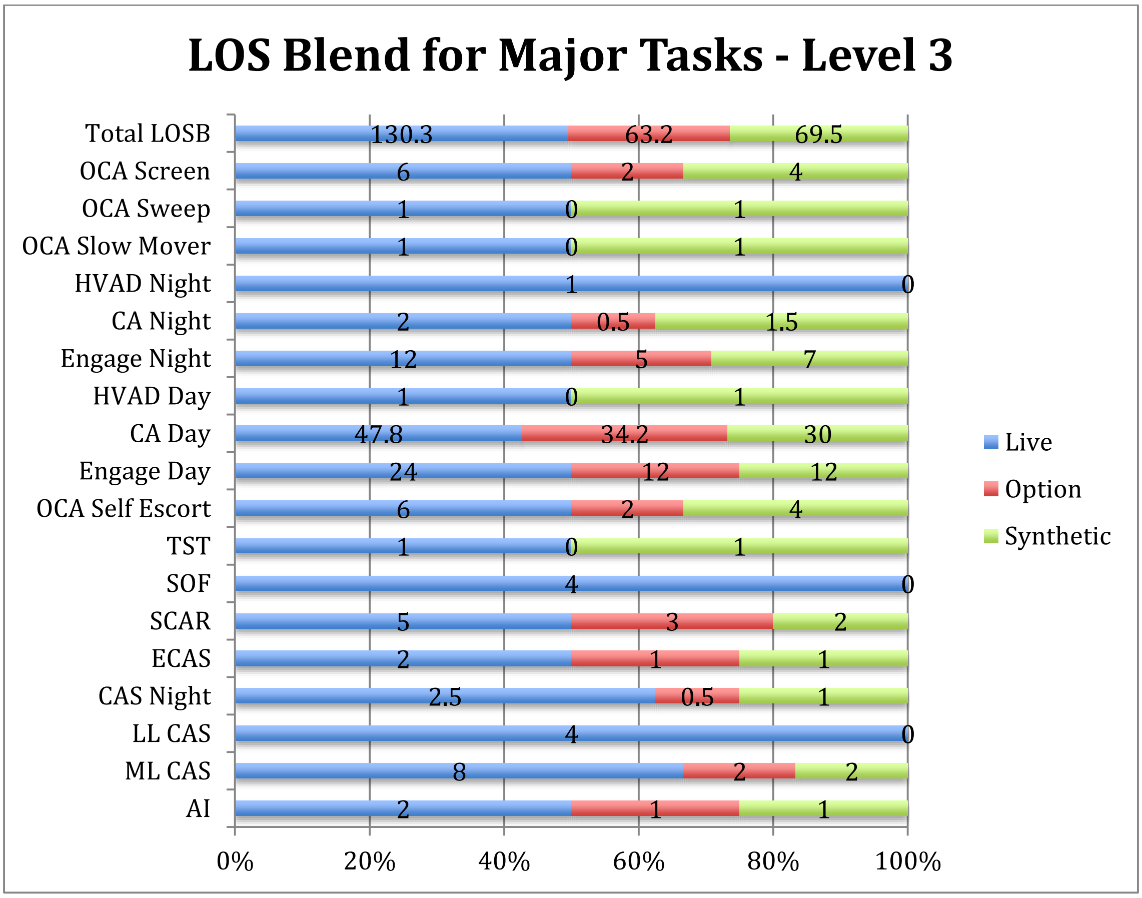

Figure Apx J-4. LOS Blend for Major Tasks - Level 3

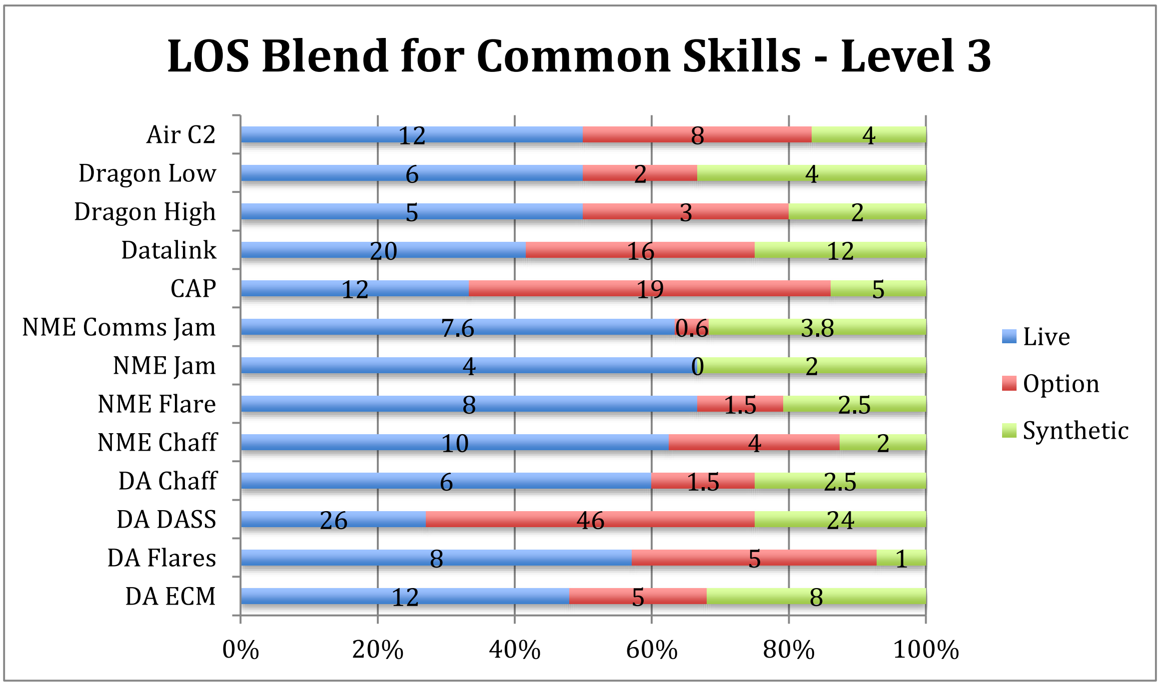

Figure Apx J-5. LOS Blend Common Skills - Level 3

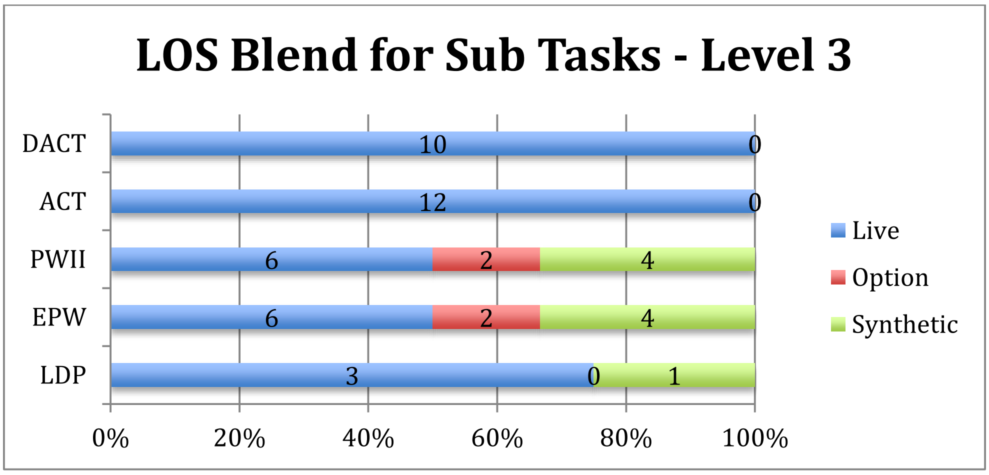

Figure Apx J-6. LOS Blend for Sub Tasks - Level 3

Table Apx J-2. LOS Blend for Level 3

Major Tasks	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
		Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synthetic	
AI	4	2			1			N/A	N/A	Mode	50	25	25	25.0
ML CAS	12	8	6.0, 10.0	3.4	2	2.0, 4.0	2.1	0.12	0.01	Median	67	17	17	8.3
LL CAS	4	4			0			N/A	N/A	Mode	100	0	0	25.0
CAS Night	4	2.5			1			N/A	N/A	Mode	63	13	25	25.0
ECAS	4	2			1			N/A	N/A	Mode	50	25	25	25.0
SCAR	10	5	3.3, 6.7	2.9	2	2.0, 5.0	2.3	0.15	0.04	Median	50	30	20	10.0
SOF	4	4			0			N/A	N/A	Mode	100	0	0	25.0
TST	2	1			1			N/A	N/A	Mode	50	0	50	50.0
OCA Self Escort	12	6	6.0, 8.0	2.4	4	4.0, 5.0	1.9	0.08	<0.01	Median	50	17	33	8.3
Engage Day	48	24	12.0, 30.0	11	12	10.5, 18.0	6.8	0.08	0.05	Median	50	25	25	2.1
CA Day	112	47.8	39.2, 56.3	18.8	30	22.5, 37.4	15.9	0.6	0.84	Mean	43	31	27	0.9
HVAD Day	2	1			1			N/A	N/A	Mode	50	0	50	50.0
Engage Night	24	12	12.0, 15.0	4.5	7	5.0, 10.0	3.6	0.02	0.16	Median	50	21	29	4.2
CA Night	4	2			1.5			N/A	N/A	Mode	50	13	38	25.0
HVAD Night	1	1			0			N/A	N/A	Mode	100	0	0	100.0
OCA Slow Mover	2	1			1			N/A	N/A	Mode	50	0	50	50.0
OCA Sweep	2	1			1			N/A	N/A	Mode	50	0	50	50.0
OCA Screen	12	6	6.0, 8.0	2.5	4	4.0,6.0	2.2	0.06	0.04	Median	50	17	33	8.3
Total LOSB	263	130.3			69.5						50	24	26	

Major Task	Sub Task	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
			Average	95% CI	Std Deviation	Average	95% CI	Std Deviation	Live	Synth		Live	Option	Synthetic	
AI	LDP	4	3			1					Mode	75	0	25	25.0
	EPW	12	6	4.0, 8.4	3.5	4	1.3, 6.0	2.6	0.00	0.01	Median	50	17	33	8.3
	PWII	12	6	6.0, 10.0	3.5	4	1.3, 4.0	2.5	0.01	0.01	Median	50	17	33	8.3
Counter Air	ACT	12	12	12.0, 12.0	1.4	0	0.0, 0.0	1.1	<0.01	<0.01	Median	100	0	0	8.3
	DACT	10	10	9.7, 10.0	1.9	0	0.0, 0.0	0.8	<0.01	<0.01	Median	100	0	0	10.0

Common Skills	Total Events	Live			Synthetic			Ryan Joiner p		Type	LOSB (%)			Sensitivity (%)
		Average	95% CI	Std Deviation	Average	5 CI Std Deviation		Live	Synth		Live	Option	Synthetic	
DA ECM	25	12	10.0, 1530	9.4	8	5.0, 10.0	3.9	<0.01	0.05	Median	48	20	32	4.0
DA Flares	14	8	5.7, 12.0	4.6	1	0.0, 4.0	3	0.20	<0.01	Median	57	36	7	7.1
DA DASS	96	26	24, 46.0	15.5	24	10.5, 40.0	18.1	0.02	0.17	Median	27	48	25	1.0
DA Chaff	10	6	5.0, 9.3	2.9	2.5	0.2, 3.0	1.9	0.09	0.02	Median	60	15	25	10.0
NME Chaff	16	10	8.0, 12.0	5	2	0.0, 4.0	2.3	0.09	<0.01	Median	63	25	13	6.3
NME Flare	12	8	5.6, 10.3	3.8	2.5	1.0, 4.0	2.06	0.16	0.05	Median	67	13	21	8.3
NME Jam	6	4			2			N/A	N/A	Mode	67	0	33	16.7
NME Comms Jam	12	7.6	6.1, 9.1	3.3	3.8	2.0, 5.8	3.1	0.14	0.21	Mean	63	5	32	8.3
CAP	36	12	5.7, 18.7	8.3	5	1.7, 11.5	6.4	0.09	0.03	Median	33	53	14	2.8
Datalink	48	20	12.0, 24.0	11.5	12	5.7, 20.0	11.9	0.07	0.09	Median	42	33	25	2.1
Dragon High	10	5	5.0, 8.0	2.9	2	2.0, 5.0	2.6	0.05	0.02	Median	50	30	20	10.0
Dragon Low	12	6	5.7, 8.0	3.4	4	2.0, 6.0	2.9	0.07	0.08	Median	50	17	33	8.3
Air C2	24	12	8.7, 15.0	8	4	0.0, 5.3	4.5	0.10	0.00	Median	50	33	17	4.2

Table Apx J-3. Available Comparisons between the Complexity Levels

Major Task	**Level 1 v 2**	**Level 2 v 3**	**Level 1 v 3**
CA Day			
Engage Day			
Engage Night			
ML CAS			
SCAR			
OCA Self Escort			
OCA Screen			
Sub Task			
ACT			
DACT			
EPW			
PWII			
Common Skills			
DA ECM			
DA Flares			
DA DASS			
DA Chaff			
NME Chaff			
NME Flare			
NME Jam			
CAP			
NME Comms Jam			
Datalink			
Dragon High			
Dragon Low			
Air C2			

Table Apx J-4. Tasks and Skills Comparison Against Complexity-Level 1v2

Major Task	Environment	Test Used	N	N for test	Statistic	p	Delta of Average	Hypothesis Upheld
CA Day	Live	Paired-t	36	36	T = -0.02	0.491	-0.08	Ho
	Synthetic	Paired-t	36	36	T =-0.63	0.268	-1.14	Ho
Engage Day	Live	Wilcoxon	36	25	W= 164	0.527	0	Ho
	Synthetic	Wilcoxon	36	24	W = 134	0.334	0	Ho
Engage Night	Live	Wilcoxon	36	27	W = 204	0.645	0.5	Ho
	Synthetic	Wilcoxon	36	27	W = 167	0.307	-0.5	Ho
Sub Task								
ACT	Live	Wilcoxon	36	18	W = 69	0.243	0	Ho
	Synthetic	Wilcoxon	36	15	W = 79	0.866	0	Ho
DACT	Live	Wilcoxon	36	12	W = 38.5	0.5	0	Ho
	Synthetic	Wilcoxon	36	7	W = 12.5	0.433	0	Ho
Common Skills								
DA ECM	Live	Wilcoxon	36	26	W = 123	0.093	-1.5	Ha
	Synthetic	Wilcoxon	36	22	W = 173	0.938	1.5	Ho
DA Flares	Live	Wilcoxon	36	25	W = 88	0.023	-1	Ha
	Synthetic	Wilcoxon	36	25	W = 213	0.917	0.5	Ho
DA DASS	Live	Wilcoxon	36	27	W = 165	0.29	-3	Ho
	Synthetic	Wilcoxon	36	25	W = 171	0.596	0	Ho
DA Chaff	Live	Wilcoxon	35	26	W = 123	0.095	-0.5	Ha
	Synthetic	Wilcoxon	36	21	W = 119	0.555	0	Ho
NME Chaff	Live	Sign	35	15 below 13 above	-	0.425	0	Ho
	Synthetic	Wilcoxon	35	21	W = 145	0.855	0.5	Ho
NME Flare	Live	Wilcoxon	36	24	W = 99	0.077	-0.5	Ha
	Synthetic	Wilcoxon	36	20	W = 113	0.625	0	Ho
NME Jam	Live	Sign	35	8 below 8above	-	0.599	0	Ho
	Synthetic	Sign	35	7 below 7 above	-	0.605	0	Ho
CAP	Live	Sign	36	9 below 14 above	-	0.895	0	Ho
	Synthetic	Paired-t	36	36	T = -0.19	0.426	-0.22	Ho
Datalink	Live	Wilcoxon	35	23	W = 139	0.524	0	Ho
	Synthetic	Wilcoxon	35	18	W = 55	0.099	-0.5	Ha

Table Apx J-5. Tasks and Skills Comparison Against Complexity-Level 2v3

Major Task	Environment	Test Used	N	N for test	Statistic	p	Delta of Average	Hypothesis Upheld
CA Day	**Live**	Paired-t	36	36	T = -0.02	0.491	-0.08	Ho
	Synthetic	Paired-t	36	36	T =-0.63	0.268	-1.14	Ho
Engage Day	**Live**	Wilcoxon	36	25	W= 164	0.527	0	Ho
	Synthetic	Wilcoxon	36	24	W = 134	0.334	0	Ho
Engage Night	**Live**	Wilcoxon	36	27	W = 204	0.645	0.5	Ho
	Synthetic	Wilcoxon	36	27	W = 167	0.307	-0.5	Ho
Sub Task								
ACT	**Live**	Wilcoxon	36	18	W = 69	0.243	0	Ho
	Synthetic	Wilcoxon	36	15	W = 79	0.866	0	Ho
DACT	**Live**	Wilcoxon	36	12	W = 38.5	0.5	0	Ho
	Synthetic	Wilcoxon	36	7	W = 12.5	0.433	0	Ho
Common Skills								
DA ECM	**Live**	Wilcoxon	36	26	W = 123	0.093	-1.5	Ha
	Synthetic	Wilcoxon	36	22	W = 173	0.938	1.5	Ho
DA Flares	**Live**	Wilcoxon	36	25	W = 88	0.023	-1	Ha
	Synthetic	Wilcoxon	36	25	W = 213	0.917	0.5	Ho
DA DASS	**Live**	Wilcoxon	36	27	W = 165	0.29	-3	Ho
	Synthetic	Wilcoxon	36	25	W = 171	0.596	0	Ho
DA Chaff	**Live**	Wilcoxon	35	26	W = 123	0.095	-0.5	Ha
	Synthetic	Wilcoxon	36	21	W = 119	0.555	0	Ho
NME Chaff	**Live**	Sign	35	15 below 13 above	-	0.425	0	Ho
	Synthetic	Wilcoxon	35	21	W = 145	0.855	0.5	Ho
NME Flare	**Live**	Wilcoxon	36	24	W = 99	0.077	-0.5	Ha
	Synthetic	Wilcoxon	36	20	W = 113	0.625	0	Ho
NME Jam	**Live**	Sign	35	8 below 8above	-	0.599	0	Ho
	Synthetic	Sign	35	7 below 7 above	-	0.605	0	Ho
CAP	**Live**	Sign	36	9 below 14 above	-	0.895	0	Ho
	Synthetic	Paired-t	36	36	T = -0.19	0.426	-0.22	Ho
Datalink	**Live**	Wilcoxon	35	23	W = 139	0.524	0	Ho
	Synthetic	Wilcoxon	35	18	W = 55	0.099	-0.5	Ha

Common Skills								
DA ECM	Live	Wilcoxon	21	11	W = 20.5	0.143	0	Ho
	Synthetic	Sign	20	8 below 2 above	-	0.055	0	Ho
DA Flares	Live	Wilcoxon	21	11	W = 34	0.553	0	Ho
	Synthetic	Wilcoxon	20	9	W = 29	0.797	0	Ho
DA DASS	Live	Sign	21	6 below 9 above	-	0.849	0	Ho
	Synthetic	Wilcoxon	20	14	W = 33	0.123	-2.5	Ho
DA Chaff	Live	Wilcoxon	21	12	W = 35	0.407		Ho
	Synthetic	Sign	20	8 below 3 above	-	0.113	0	Ho
NME Chaff	Live	Paired-t	21	21	T = -0.41	0.342	-0.29	Ho
	Synthetic	Wilcoxon	20	11	W = 45	0.867	0	Ho
NME Flare	Live	Wilcoxon	21	9	W = 23	0.571	0	Ho
	Synthetic	Sign	20	6 below 2 above	-	0.144	0	Ho
NME Jam	Live	Wilcoxon	21	8	W=10.5	0.163	0	Ho
	Synthetic	Wilcoxon	20	8	W = 14.5	0.337	0	Ho
CAP	Live	Sign	21	3 belwo 7 above	-	0.945	0	Ho
	Synthetic	Wilcoxon	20	12	W = 28	0.216	0	Ho
NME Comms Jam	Live	Wilcoxon	21	10	W = 22	0.305	0	Ho
	Synthetic	Sign	20	-	5 below 2 ab	0.227	0	Ho
Datalink	Live	Sign	21	5 below 6 above	-	0.725	0	Ho
	Synthetic	Sign	21	8 below 5 above	-	0.291	0	Ho
Dragon High	Live	Wlcoxon	21	13	W = 43	0.444	0	Ho
	Synthetic	Wilcoxon	21	13	W = 37.5	0.3	0	Ho
Dragon Low	Live	Sign	21	6 below 6 above	-	0.613	0	Ho
	Synthetic	Paired-t	21	21	T = -0.90	0.189	0	Ho
Air C2	Live	Wilcoxon	21	10	W = 29.0	0.581	0	Ho
	Synthetic	Sign	21	5 below 1 above	-	0.109	0	Ho

Table Apx J-6. Tasks and Skills Comparison Against Complexity-Level 1v3

Major Task	Environment	Test Used	N	N for test	Statistic	p	Delta of Average	Hypothesis Upheld
CA Day	Live	Paired-t	21	21	T=0.56	0.708	2.29	Ho
	Synthetic	Paired-t	20	20	T=-2.89	0.005	-9.15	Ha
Engage Day	Live	Wilcoxon	21	14	W = 64.5	0.784	2.5	Ho
	Synthetic	Paired-t	20	20	T=-2.13	0.023	-3.9	Ha
Engage Night	Live	Paired-t	21	21	T=0.63	0.732	0.86	Ho
	Synthetic	Sign	20	11 below 3 above	-	0.029	-2	Ha
Sub Task								
ACT	Live	Sign	21	6below 0 above	-	0.016	0	Ho
	Synthetic	Wilcoxon	20	6	W = 13	0.735	0	Ho
DACT	Live	Wilcoxon	21	6	W = 6.0	0.201	0	Ho
	Synthetic	Sign	21	3 below 1 above	-	0.312	0	Ho
Common Skills								
DA ECM	Live	Wilcoxon	21	15	W = 37	0.101	-2.5	Ho
	Synthetic	Paired-t	20	20	T = 0.04	0.517	0.05	Ho
DA Flares	Live	Sign	21	9 below 6 above	-	0.304	0	Ho
	Synthetic	Paired-t	20	20	T = 1.52	0.928	1.35	Ho
DA DASS	Live	Sign	21	6 below 9 above	-	0.849	0	Ho
	Synthetic	Paired-t	20	20	T = -1.94	0.034	-8.3	Ha
DA Chaff	Live	Paired-t	21	21	T = -1.74	0.048	-1.6	Ha
	Synthetic	Wilcoxon	20	16	W = 78	0.715	0	Ho
NME Chaff	Live	Sign	20	9 below 5 above	-	0.212	0	Ho
	Synthetic	Paired-t	19	19	T = 1.45	0.918	0.9	Ho
NME Flare	Live	Wilcoxon	21	16	W = 42	0.094	-1	Ha
	Synthetic	Sign	20	9 below 4 above	-	0.133	0	Ho
NME Jam	Live	Paired-t	20	20	T = 0.76	0.772	0.35	Ho
	Synthetic	Paired-t	19	19	T = -1.92	0.035	-0.474	Ha
CAP	Live	Sign	21	3 below 11 above	-	0.994	1	Ho
	Synthetic	Paired-t	20	20	T = -0.03	0.486	-0.05	Ho
Datalink	Live	Paired-t	20	20	T = -0.02	0.493	-0.05	Ho
	Synthetic	Sign	20	8 below 4 above	-	0.194	0	Ho

Appendix K Project JENX Supporting Information

PROJECT JENX SYLLABUS

Table Apx K-1. Project JENX Syllabus

Day		Synth Sortie Title	Sims Req			Live Sortie Title
WEEK 1 - Skillset generation						
1	AM	DCA Sim 1	2	SIM		
1	PM	ASDM BRF		BRF		
2	AM	ASDM SIM	1	SIM		
2	PM	EPM Brf		BRF		
3	AM	EA 1	1	SIM		
3	PM	EPM Tactics Brf		BRF		
4	AM	EA 2	2	SIM		
4	PM	AA ROE BRF		BRF		
5	AM	EA 3	2	SIM		
5	PM	AA 10 Brf		BRF		
WEEK 2 - Develop skillsets						
6	AM	DCA Sim 2	2	SIM		
6	PM			LIVE	F7	Arrival Cx
7	AM	DCA 2	2	SIM		
7	PM			LIVE	F8	ACT 1
8	AM	DCA Sim 3	4	SIM		
8	PM			LIVE	F9	ACT 2
9	AM	DCA 3	4	SIM		
9	PM	AA 12 BRF		BRF		
10	AM	DCA Sim 4	4	SIM		
10	PM	Spare slot		SIM		
WEEK 3 - Contextual employment of skills 1						
11	AM	DCA 4	4	SIM		
11	PM	OCA BRF		BRF		
12	AM	OCA SIM 1	2	SIM		
12	PM			LIVE	F10	2v1 ACT
13	AM	OCA 1	4	SIM		

13	PM			LIVE	F11	2v2 ACT
14	AM	OCA 2	4	SIM		
14	PM			LIVE	F12	ASDM 1
15	AM	Night DCA sim 1	2	SIM		
15	PM	Spare slot		SIM		
WEEK 4 - Contextual employment of skills 2						
16	AM	Night DCA 1	4	SIM		
16	PM			LIVE	F13	ASDM 2
17	AM	Night DCA sim 2	4	SIM		
17	PM	Helo BRF		BRF		
18	AM	Night DCA 2	4	SIM		
18	PM	Attack Brf		BRF		
19	AM	Helo SIM	1	SIM		
19	PM			LIVE	F25	Helo Affil
20	AM	AS,WSH,HOTAS		BRF		
20	PM	EP2/PW4 Fuse+Planning Tools		BRF		
WEEK 5 - AS Skillset generation						
21	AM	Attack Sim 1	2	SIM		
21	PM	AS ROE BRF		BRF		
22	AM	Attack 1	1	SIM		
22	PM	Range BRF		BRF		
23	AM	Range Sim 1	1	SIM		
23	PM	DT/TST BRF		BRF		
24	AM	Attack Sim 2	1	SIM		
24	PM			LIVE	F31	Range 1
25	AM	Attack 2	2	SIM		
25	PM	Spare slot		SIM		
WEEK 6 - CAS						
26	AM	Night Attack	2	SIM		
26	PM	Strafe BRF		BRF		
27	AM	Strafe Sim 1	1	SIM		
27	PM	CAS Brf		BRF		
28	AM	Strafe Sim 2	1	SIM		
28	PM			LIVE	F34	Strafe 1

29	AM	CAS sim 1	2	SIM		
29	PM	Opposed AI Brf		BRF		
30	AM	AI Sim 1	4	SIM		
30	PM	Spare slot		SIM		
WEEK 7 - Test Week						
31	AM	AI 1	4	SIM		
31	PM			LIVE	F36	CAS Consol
32	AM	AI 2 (night)	4	SIM		
32	PM			LIVE	F35	Strafe 2
33	AM					
33	PM			LIVE	F37	CAS Qual (day)
34	AM					
34	PM			LIVE	F39	MR TAC Cx
35	AM					
35	PM			LIVE	N37	CAS Qual (night)

RESOURCE CALCULATOR

Table Apx K-2. Resource Calculator - Project JENX

Serial	Serial Name	Content	Planned Blue			Minimum Blue			Planned Red			Minimum Red			PROJECT Jenx				Hours
			TYP	STRIKER	SIM	TYP	STRIKER	SIM	TYP	DA20	Hawk	TYP	DA20	Hawk	TYP	DA20	Hawks	Sims	
1	Arrival Brief																		01:00
2	SERE Brief	ISOPREP, PRC112 and EPA Refresh with Sqn																	00:00
3	Secure Radios Brief																		01:00
4	HMSS																		02:00
5	Red Air Phase Brief																		02:00
6	RAIDS Brief																		01:00
7	Arrival Cx		2					1										1	01:30
8	Night Arrival Cx		2					1										1	01:30
9	ACT 1	HMSS Famil	1			1			1			1			1				01:15
10	ACT 2		1			1			1			1			1				01:15
11	2v1 (D)ACT		2			2			1					1	2		1	1	01:15
12	2v2 (D)ACT		2			2			2					2	2		2	2	01:15
13	ASDM Phase Brief	DASS tech, DASS Ds&Cs, ASDM Mech																	02:30
14	ASDM Sim	Dragon/Dragon+			1			1										1	01:15
15	ASDM 1	Spade/Spaces Dragon/Dragon Hi	2			1			1			1			2				01:30
16	ASDM 2	Spade/Spaces Dragon+/CCM	2			1			1			1			2				01:30
17	A-A ROE Brief																		01:30
18	DCA Sim 1	2vX DCA 10C Stagger Back FQMD																	01:15
19	DCA 1	2v2 DCA 10C Stagger Back FQMD	2			2			2				1	1				1	01:30
20	EPM Phase Brief	Dii Secret																	03:00
21	EA 1	2v1 Academic EA 10C	2			1				1			1					1	01:45
22	EPM Tactics Phase Brief	Dii Secret																	01:30
23	EA 2	2v1 EA Tactics 10C	2			2				1			1					2	01:30
24	EA 3	2v2 EA Tactics 10C	2			2				2			1	1				2	01:30
25	AA-10D Brief	Dii Secret																	01:00
26	DCA Sim 2	2vX DCA Point Def 10C/D EA Off VIDs			2			2										2	01:15
27	DCA 2	2vX DCA Point Def 10C/D EA Off VIDs	2			2			2	1				2				2	01:30
28	DCA Sim 3	4vX DCA Point Def 10C/D EA			4			4										4	01:15
29	DCA 3	4vX DCA Point Def 10C/D EA	4			3			2	2				3				4	01:30
30	AA-12 Brief	Dii Secret																	01:00
31	DCA Sim 4	4vX DCA Area Def 12/10C EA			4			4										4	01:00
32	DCA 4	4vX DCA Area Def 12/10C/D EA	4			3			2	2				3				4	01:15
33	OCA Phase Brief																		
34	OCA Sim 1	2vX OCA Detached Escort 10C EA			2			2										2	01:15
35	OCA 1	4vX OCA Detached Escort 10C/D EA	3	2		3	1		2	2				3				4	01:30
36	OCA 2	4vX OCA Fighter Sweep 12/10C/D EA	3	4		3	1		2	2				3				4	01:30
37	Night DCA Sim 1	Night 2vX DCA 10C/D EA (IMC)			2			2										2	01:15
38	Night DCA 1	Night 4vX DCA 10C/D EA	4			3			2	2				3				2	01:30
39	Night DCA Sim 2	Night 4vX DCA 10C/D EA			4			4										4	01:15
40	Night DCA 2	Night 4vX DCA 10C/D EA	4			3			2	2				3				4	01:30

41	Helo Affil Phase Brief	Helo Affil and Engagement																	02:00
42	Helo Affil Sim	Helo Affil and LL Slow Gunnery																	01:15
43	Helo Affil																		01:30
44	Attack Phase Brief																		01:30
45	A-S WSH and HOTAS																		01:00
46	EP2/PW4 and Fusing	MFBF Refresh If Required																	
47	Planning Tools	Planning Tools and Wpns Effects PB																	
48	LIII Phase Brief																		01:00
49	Attack Sim 1	LDP EP2/PW4 and Co-op Des			2			2										2	01:15
50	Attack 1	LDP EP2/PW4 and Co-op Des	2			2												1	01:30
51	A-S ROE Brief																		01:00
52	Range Brief																		01:30
53	Range Sim 1	Academic Range Profiles			1			1										1	01:15
54	Range 1	Academic Range Profiles	2			1									1				01:30
55	DT/TST Brief	Dynamic Targeting and Time Sensitive Tasking (DT/TST)																	01:00
56	Attack Sim 2	DT/TST			1			1										1	01:15
57	Attack 2	DT/TST	2			2												2	01:30
58	Night Attack		2			2												2	01:30
59	Strafe Phase Brief																		01:30
60	Strafe Sim 1	Strafe Intro			1			1										1	01:15
61	Strafe Sim 2	Strafe Consol and Solo Clearance			1			1										1	01:15
62	Strafe 1	Strafe	2			1									1				01:30
63	Strafe 2	Strafe Consol and Qualification	2			1									1				01:30
64	CAS Phase Brief																		02:00
65	CAS Sim 1	CAS Intro			1			1										1	01:15
66	CAS 1	CAS Consol	2			2												2	01:30
67	CAS 2	CAS Qualification (Day)	2			1												2	01:30
68	Night CAS 1	CAS Qualification (Night)	2			2												2	01:30
69	LIII Recce	AMT 8 -NTISR	1			1												1	01:30
70	Opposed AI Phase Brief																		01:30
71	AI Sim 1	2vX AI Intro 10C/D EA IADS			1			1										1	01:15
72	AI 1	4vX AI 10C/D EA IADS	4			3			2	1				2				4	01:30
73	AI 2 (Night)	4vX AI 10C/D EA IADS	4			3			2	1				2				4	01:30
74	MR Tac Cx	4vX D/OCA or AI vs 12/10C/D EA IADS	4	1		3			2	2	2			4	4	2	2		01:30

PROJECT JENX SAVINGS - OVERVIEW

Table Apx K-3. Project JENX Savings Overview

	Planned Blue			Minimum Blue			Planned Red			Minimum Red			PROJECT Jenx			
	TYP	STRIKER	SIM	TYP	STRIKER	SIM	TYP	DA20	Hawk	TYP	DA20	Hawk	TYP	DA20	Hawks	Sims
TOTALS	77	7	27	59	2	29	29	21	2	4	4	33	17	2	5	82
Number of events	32	3	14	30	2	16	17	13	1	4	4	14	10	1	3	38
	Planned Blue			**Minimum Blue**			**Planned Red**			**Minimum Red**			**PROJECT Jenx**			
Hours flown	**113.50**	**10.50**	**27.00**	**86.50**	**3.00**	**35.75**	**41.75**	**31.25**	**3.00**	**5.50**	**6.25**	**48.00**	**24.00**	**3.00**	**6.75**	**102.50**
HOURS: Proj Jenx v Std Syllabus	**-89.50**		**75.50**	**-62.50**		**66.75**	**-41.75**	**-28.25**	**3.75**	**-5.50**	**-3.25**	**-41.25**				
	Live	/Synth		Live	/Synth								Live	/Synth		
LSB=	70	30		65	35								21	79		
Proj Jenx v Std Syllabus	*Planned CR work up*			Min Assets Work up												
Typhoon (hours)	-131.25			-68.00												
DA20(hours)	-28.25			-3.25												
Hawk(hours)	3.75			-41.25												
Sim(hours)	75.50			66.75												
Total Typhoon Sorties (Missions/Waves required)	-39			-24												

PROJECT JENX FEEDBACK FROM STUDENT

Overall impressions of course

Overall I thought it was a well structured course and was pleased to have briefs scheduled into the program. Having completed a CR workup on the GR4 and spoken to my counterparts the 'traditional' way of doing things was to just fit it in when ever possible – sometimes this would be 10 minutes before a trip and the overriding priority would be to fly/Sim regardless of preparation. I also particularly liked having the initial phase being on a one-to-one basis (Flt Lt Skinner – effectively a QWI!). It meant he would know what standard I was at and could tailor the Sims as required – plus the briefs would not overlap and he knew what I had been taught.

How do I consider the training transferred to live flight?

Tactically the transition has been easier than I thought. When I started actually flying DCA/AI it was hard because I had flown twice in the previous 2 months and G tolerance was noticeably poor during combat. During the week commencing 19 Jan 15, I flew 5 times. Currency was therefore now not an issue and I found the tactical side relatively straight forward including EA, 4 ships and the occasional problematic jet! To feel fully comfortable I would like to live fly a night DCA (4 ship) as I would then feel more confident in my abilities to do this for real (not just in the Sim).

Do I feel disadvantaged compared to counterparts on the standard course? How does it compare to the progress from my OCU course?

With respect to knowledge and tactical exposure I would say (certainly compared to my course mates from the OCU) I have the upper hand. Just after Christmas I had spoken to 2 of them and both had completed 4 workup events. They had flown more than me but their workup was scattered and they were definitely bottom of the priority list. By this point I was almost complete. Due to Christmas and the rest of the squadron having to gain enough flying for Q currency, my priority dropped and so barely flew. This is where I may have been disadvantaged (as they were flying more regularly).

In summary I feel I am in a better situation as have completed all the CR workup trips (albeit in the Sim) and now can fly in 4 ships in any scenario and feel competent. The original plan was to fly 2/3 times per week so that would negate what I see as the main disadvantage of my individual workup. Christmas just got in the way! My OCU course may have flown more but have not achieved anywhere near the number of CR workup events, in fact the situation is not much better with people from 1 or even 2 courses ahead of me.

Course structure/ Level of training provided

The structure was well planned and I felt thoroughly confident at the end of each phase. Again noting one-to-one teaching and planned afternoons for briefs and study. I would like to mention that it has been good to fly again a bit more regularly for familiarity in the jet – plus I have been doing live DCA where as good as the Sim is – it just cannot replicate the entire sensation (G, environmental factors, jet not working properly etc). The level of EA in the Sim is a known limitation so has been good to see live.

Other comments

This workup has been from my point of view very good. I definitely think it is the way forward, perhaps with a few tweaks as it hopefully gets people CR quicker and therefore able to concentrate on improving rather than constantly trying to get a CR tick. Changes I would make would be to fly a couple of elements (DCA/AI/OCA) towards the end even if it was just for confidence of being able to do it live, but more so to get used to real life jet issues.

Negatives:

- Lack of flying. Not down to the workup plan but what actually happened. This made both my G tolerance and ability to fly markedly reduced.
- EA – not fully correct in the Sim.

- A-S – Pod not exactly the same as in the Sim – Only took a few minutes to get used to in the live environment – but I have had previous experience using it.
- It is difficult to see the standard at which I should be at as I am comparing myself to the QWI (during the first 3 weeks) thus always appearing poor! A potential low moral point (was for me!)

Positives:

a. Programmable phase briefs and time to consolidate.

b. Quick enough pace so everything remains fresh for the next event.

c. The Sim can be limited but it can replicate many more hostiles that we get in the 323's and they can all be going at parameters that we rarely can simulate live – this I think is great training.

d. You can really nail down 4 ship ops - I didn't have to think about it when I flew the first 4 ship. My arrow may have been a little slack but generally I felt it was good.

Finally I think a point to make that makes this whole work up worth while is sticking to the plan. The lack of flying (couldn't be helped I realise) detracted from the outcome. Additionally the first 4 weeks on 3(F) Sqn (who were kind enough to host us whilst 11 Sqn were on Exercise) were crucial to the initial learning. During this time there were occasions where the programmer was required to fill sims/flights and I was then tasked to do these on top of the workup. The upshot of this was very long days, increased fatigue and the inability to read/prepare as much for trips (due to crew duty). This had no benefit (apart from the every event being filled) and detracted from my ability to learn/perform well in the next CR trip.

Experience

As a second tourist I found some aspects (CAS theory, some threats and EA theory) easier to understand than an ab-inito would have. However as there was programmed time to brief/read up then I foresee this just taking a bit longer to explain and would not affect the time scale of the CR workup. Lack of flying may have more of an impact for a less experienced pilot but this may just require an extra few trips to get current again. In my opinion I can't see a less experienced pilot struggling more than I did as I have very little AD background and the A-S work is different enough to require additional learning.

Flt Lt Jenkins

XI Sqn MRCR pilot

Appendix L Officer Commanding 29 Squadron Statement of Impact

ROYAL AIR FORCE

29 Squadron

RAF Coningsby
Coningsby
Lincolnshire
LN4 4SY
Tel: 01526 346330
Fax: 01526 346312
Email: 29RSqn-OC@mod.uk

Date: 18 Jul 16

Dear Alistair,

In response to your request for a statement of the impact your work has had on the Typhoon Force I am delighted to provide the following, in order that your work is recognised for the fundamental changes it has brought.

Allsop's research has been the definitive work in the RAF with respect to simulation and the manner and extent of its employment. His work has informed the use of synthetics across the spectrum of training but particularly on the Typhoon Operational Conversion Unit (OCU), by pioneering the understanding of the capability of simulation when training our trainee pilots. This knowledge has been directly used to construct the new Combat Ready syllabus that will now, for the first time, enable us to graduate students in Multi Role combat operations; a role normally devolved to the Front Line (FL) training units. This role previously took the FL units on average 3-4 months to complete. Due to the increased use of simulators for training and consolidation, this additional training is now conducted in only 4 weeks. The increase in the OCU task has also been achieved within the extant resources available, which would just not have been possible without Allsop's research. Additionally, the OCU has also be able to realise major efficiencies; his research has allowed us to transition from a live flying to simulator ratio of 50:50 to 25:75, allowing us to train more pilots per year that ever before. As a second order effect, the research has been in no small way responsible for the increased availability of the simulators, from below 50% five years ago to above 95% presently.

I would be happy to answer any enquiries from your supervisory staff or examining body.

Yours Sincerely,

James Bolton
Wing Commander