Behavioural Markers of Air Traffic Controller Development

David John Thompson

A thesis submitted to

University College London

For the degree of

Doctor in Engineering (EngD)

Communications Engineering Doctorate Centre

Department of Electronic and Electrical Engineering

UCL

October, 2013

Sponsoring organisation:

THESIS DECLARATION

I, David John Thompson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date: ______________________________

All photos contain in this thesis are from the NATS press office: http://www.flickr.com/photos/natspressoffice/

For my parents, and the memory of my grandparents

For Alice and Lewis

For Rebecca

ACKNOWLEDGEMENTS

I am very grateful for the support, guidance, and expertise generously given by so many to me whilst undertaking this Engineering Doctorate.

My UCL Supervisors Dr Nadia Berthouze and Prof. Ann Blandford thank you for your kindness, generosity, warmth, and unwavering support through my research. It has been a pleasure working within the UCLIC centre. I am grateful to all the Communications Engineering Doctorate Centre staff that helped me throughout my studies. In particular, I would like to thank Paul McKenna, Dr Bob Sutherland, Prof. Andy Valdar, Prof. Ian Moffet, and Vicky Coombes.

I would like to sincerely thank Liz Skinner, Nic Turley, and Neil May for supporting my ambitions to undertake doctoral study in conjunction with my employment at NATS. Warm thanks especially to my industrial supervisors Dr Brian Janes and Simon Skinner.

I wish to thank the support of my Human Factors colleagues more generally at NATS. In particular I would like to thank Dr Martin Curry for his dry wit which has been a source of tremendous support to me through my studies. The professionalism and swan-like grace demonstrated by Andy Lowrey has been a helpful inspiration in challenging situations. Dr Steve Shorrock, thank you for the research concept and the useful insights provided to me from half way round the world.

To all the air traffic controllers, operational supervisors, safety observers, and other technical staff at NATS who have help with this research; thank you for your support, time, and expertise. Finally to my wife Rebecca, you have steered me through rocky shoals and helped me to achieve my dreams, I simply could not have done this without you.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AC	Area Control (En-route)
AAC	Air Arrivals Controller
ACC	Area Control Centre
ADV	Aerodrome
ANSP	Air Navigation Service Provider
ANTS	Anaesthetists Non-Technical Skills
ATC	Air Traffic Control
AT-SAT	Air Traffic Selection And Training
ATCO	Air Traffic Controller (Air Traffic Control Officer)
ATM	Air Traffic Management
BARS	Behaviourally Anchored Rating Scale
BM	Behavioural Marker
BOOM	Behaviourally Oriented Observation Method
BOS	Behaviour Observation Scale
CAA	Civil Aviation Authority
CRM	Crew/Cockpit Resource Management
CPD	Continuous Professional Development
EEG	ElectroEncephaloGram
EFPS	Electronic Flight Progress Strips
FAA	Federal Aviation Administration

FEAST	First European ATCO Selection Test
FPS	Flight Progress Strip
HBP	Human Behaviour Performance
HF	Human Factors
HMI	Human Machine Interface
ICAO	International Civil Aviation Authority
iFACTS	Interim Future Area Control Tool Set
ISS	International Space Station
ITCB	Infrastructure Transition Control Board
LACC	London Area Control Centre
LAS	Local Area Supervisor
LCE	Local Competency Examiner
LOSA	Line Orientated Safety Audit
LTCC	London Terminal Control Centre
MMOP	Multilateral Crew Operations Panel's
NATS	National Air Traffic Services (formerly)
NGT	Nominal Group Technique
NOTECHS	Non TECHnical Skills
NTS	Non-Technical Skills
NOTSS	Non Technical Skills for Surgeons
OJT	On-the-Job Training

OJTI	On-the-Job Training Instructor (Instruction)
OOD	Officer Of the Deck
OTS	Over-The-Shoulder
RSSB	Railway Safety and Standard Board
RTS	Real Time Simulation
SESAR	Single European Sky ATM Research
SME	Subject Matter Expert
STAR	Safety Tracking And Reporting (database)
STCA	Short Term Conflict Alert
SPLINTS	Scrub Practitioners' List of Intraoperative Non-Technical Skills
TC	Terminal Control
TCAS	Traffic alert and Collision Avoidance System
TRACON	Terminal Radar Approach Control (USA equivalent of Terminal Control)
TEM	Threat and Error Management
TMA	Terminal Manoeuvring Area
TRM	Team Resource Management
TRUCE	TRaining in Unusual Circumstances and Emergencies
UCE	Unit Competency Examination/Examiner
WAYSRAYL	Write As You Speak, Read As You Listen

ABSTRACT

A key challenge when introducing new systems and technologies into Air Traffic control (ATC) is to understand levels of emerging controller proficiency ahead of scheduled implementation. Behavioural markers have been used in several complex industries to assess levels of non-technical skill; however these measures invariably focus upon the desired behaviours attained by the end of training. This research has explored how an Air Traffic Controller’s (ATCO’s) overt non-technical behaviour changes in presence and prevalence as they progress their expertise during training.

Through document review, expert engagement, and most extensively direct observation of ATCOs during and after training, a number of non-technical behaviours indicative of varying proficiency have been identified. These markers were placed within a simple three-level learning and development framework. Five categories emerged across the behaviours identified; i) input and interaction with the Human Machine Interface (HMI), ii) interaction with others, iii) physical posture and body Language, iv) attitude and mood; v) communications and verbal commentary.

An observation sheet containing the markers was iteratively developed, tested, and refined in various ATC environments. Both expert ATCOs undergoing system transition training, and ab-initio trainee controllers undertaking aerodrome training were followed through longitudinal study. A capped frequency count was used to record the precise presence of individual markers. Several dual-observations were also undertaken to determine inter-rater reliability and construct validity.

In total, the performance of the individual markers has been evaluated across 129 real-world observations. 30 markers demonstrate reliable correlations for changing prevalence against total system exposure time and provide an original means of tracking and monitoring subtle changes in the behaviour of ATCOs, as their levels of proficiency in the task matures with new ATC systems.

This research has been conducted through a CASE studentship funded by the EPSRC.

CHAPTER 01 – INTRODUCTION & BACKGROUND

"It seems that few realise the complex, sophisticated nature of Test & Evaluation observation and the importance of observation to behavioural measurement."
(David Meister)

1.1 PREAMBLE

On the 1st July 2002 at 23:35, a mid-air collision occurred involving a passenger aircraft and cargo flight above the German town of Überlingen. There were 71 fatalities on board the two aircraft, with no survivors. Both aircraft, a Bashkirian Airlines Tupolev-154 and a DHL Boeing 757, were within controlled airspace under the jurisdiction of the Zurich Area Control Centre (AAC). The Boeing heading north was climbing to the same level as the Tupolev heading west, with both aircraft on converging trajectories. The control centre at the time was operating in a fallback mode (due to routine maintenance) and the controller's Short Term Conflict Alert (STCA) was therefore non-functional. In addition, the controller was dividing attention between two radar displays (one focused on guiding an aircraft into Friedrichshafen airport, the other en-route control provision).

To compound matters the phone communications with adjacent centres were not available, which prevented the Upper ACC at Karlsruhe from informing the Zurich AAC controller of the conflict (despite repeated efforts). The Zurich controller gave collision avoidance instructions to the Tupolev to climb – and considered the conflict resolved; whilst on-board the two aircraft both Traffic alert and Collision Avoidance Systems (TCAS) raised the alarm to the flight crews of a potential collision. However there was confusion on the Tupolev between the pilot and co-pilot as to whether the controller's instructions to descend should be followed or the TCAS instructions to ascend (in Russia the controller has authority over TCAS, but in Europe the opposite was the case).

It is clear that the misunderstanding between the pilot and co-pilot on the Tupolev along with the preoccupation of the controller with a telecoms system that was not fully functional whilst dividing attention across two very separate tasks and workstations were all causal factors within the accident. Further to these specific factors were the unsafe general operational working practices during periods of low traffic with the Zurich Centre (Nunes & Laursen, 2004; Brooker, 2008).

Human error represents the causal factor in 70-80 percent of all accidents (Reason, 1990). Poor Non Technical Skills (NTS) such as teamwork, leadership, situation awareness, and communications have repeatedly been shown to be contributory factors within complex system safety failures, (McElhatton, 1993; Helmreich et al, 1995; Brooker, 2005; Flin et al, 2008).

Much work has been undertaken to understand i) what areas does NTS encompass, i) what are the ways we assess and evaluate them, and iii) how can we best improve them through training. Behavioural Markers is a structured method of observation used in a variety of complex technical domains to assess the Non-Technical Skill use by individuals and teams (Klampfer et al, 2001; Helmreich et al, 1994). Although often an implicit element in ATC training and assessment there is little in the way of a stand-alone NTS assessment system for Air Traffic Control (ATC).

1.2 THESIS OVERVIEW

This University College London (UCL) Communications Engineering Doctorate (EngD) Research has been undertaken at the UK Air Traffic Control (ATC) organisation NATS from 1st January 2008 to present (Appendix A01). This thesis takes principles found within the Behavioural Markers field and explores how NTS behaviours develop and change during phases of learning and development as a person moves towards competency with a system. Specifically, this research explores how Air Traffic Controller's (ATCO) behaviour changes during training, and when learning to use new ATC Systems. The research in this thesis has undertaken circa 200 observation sessions, in several UK ATC Towers and training facilities. This research has shown that overt behaviour does change over time, and measurement can help to evaluate an ATCO's level of learning and development.

During the implementation of a new system (or procedures), it is important that both the system is fit for use, and that the users are fit to use the system at desired operational levels. For complex systems users may require extensive training, and their competency is likely to take time to fully develop. Through structured observation of both technical and non-technical behaviour, it is possible to assess certain developed competencies. Structured Observation using Behavioural markers is a method of assessment used to evaluate Crew Resource Management (CRM) NTS training effectiveness in complex domains such as medicine and aviation. Behavioural markers are indicators of NTS performance, and allow the assessor to determine skill competency due to the prevalence of certain pre-identified behaviours. These observations are predominantly focused at the end of training, in order to evaluate learned behaviour.

This research expands upon the observation of specific learned non-technical skills (at the end of training), and explores how ATCO's non-technical behaviour develops and changes, when using a new Air Traffic control System. This encompasses the first unfamiliar steps through to proficient expert use. The theory of this research is that as a user learns to use a new system or procedure, they display overt behaviours which indicate their current level of skill development and competency. By examining behaviour over time, the presence and prevalence of various behavioural markers may therefore be used to determine a user's level of development. The benefits of this research are further useful insights into a user population's current state of development with a new system; providing relevant and useful information to training and validation teams when transitioning to a new system in Air Traffic Control. Extensive observation of ATCOs during and after training with a new flight strip system has revealed a variety of non-technical behaviours which indicate a ATCOs level of development with a system. Identified behaviours have formed a structured framework and observational checklist to monitor ATCO development over time.

1.3 RESEARCH QUESTIONS

Through an extensive review of the literature, a clear research gap has been identified concerning the identification, development, and testing of a robust and best practice derived set of behavioural markers for the assessment of Non-Technical Skill proficiency in Air traffic control. Research question 1 emerges as a result.

1.3.1 Research Question 1

- What non-technical behavioural markers may be used to evaluate ATCO performance?

During preliminary research (Chapter 4) it became evident that a more fundamental research question existed, and this has been the overwhelming focus of this thesis. The fundamental question amounts to what are the differences in Non Technical Skill (NTS) behaviour between a novice and expert ATCO, and can changes in these behaviours indicate developing competency. As a result the following further research questions are identified:

1.3.2 Research Question 2

- What phases of development are there, including transient stages?

1.3.3 Research Question 3

- How might the presence or prevalence of certain non-technical behaviours be used to indicate how well a user is engaging and developing with a system?

1.3.4 Research Question 4

- What situational factors may impact the presence and prevalence of certain behaviours?

1.4 ORIGINALITY

- Although there is overlap between behavioural observation systems across different domains, behavioural markers must be separately identified for each new domain. There is very little published work concerning the development of Behavioural Markers in the Air Traffic Control domain. What limited work has been produced provides very little explicit markers of behaviour, focusing instead upon competencies. This research has identified a number of relevant observable behaviours in the domain of ATC, indicative of emerging ATCO proficiency.

- Existing behavioural marker systems invariably use a high level frequency or performance rating. This research has taken a novel approach and employed a capped frequency tally, in order to gather nominal data with a high degree of sensitivity to changing prevalence of certain behaviours. Critical evaluation of this approach has also been undertaken.

- Behavioural observation systems rarely use the source of direct observation to elicit potential NTS behavioural markers. This research has used observation extensively to identify behaviours, eliminating recollection and other biases.

- The majority of behavioural observation systems developed have been evaluated within the simulation environment, often using scripted scenarios with actors deliberately displaying certain behaviours. This research has evaluated the behavioural markers identified in the field, during training and live operations.

1.5 STRUCTURE OF THESIS

This thesis is structured into eleven chapters. Table 1.1 provides a high level summary of each chapter. Figure 1.1 provides an illustrative depiction of the research structure, particularly the iterative cycle of behavioural markers development, testing, refinement and reapplication.

Chapter 1	Current Chapter
Chapter 2	An overview of Air Traffic Control and the role of the controller within the UK. This chapter includes details of NATS operations, and various electronic strip systems which have been introduced to support systemised operations.
Chapter 3	A background review of the literature exploring the observation of Non-Technical Skills in safety critical domains.
Chapter 4	A preliminary study which identified a number of NTS applicable to ATC, and gathered data using a method derived from this material.
Chapter 5	A study which identified through observation behavioural markers to assess different levels of Non-Technical Skill (NTS) proficiency of Air Traffic Controllers (ATCOs) engaging with their primary flight strip system, and placed them within an observation sheet
Chapter 6	An observational study using the method developed in chapter 5 to track changes in ATCO behaviour whilst transitioning from paper to electronic flight progress strips.
Chapter 7	A study which explored inter-rater reliability of the behavioural marker set developed in Chapter 5, and observer feedback regarding the methodology employed in this
Chapter 8	An observational study exploring changes in trainee ATCOs NTS behaviour as they undertake an aerodrome ATC course.
Chapter 9	A study which gathered observational data within a non-benign high workload Air Traffic environment in order to evaluate the impact upon NTS behaviour
Chapter 10	This chapter reviews the findings across the four study chapter (5-9) and derives a final set of behavioural markers based on the significant findings. The chapter then discusses the research undertaken with regards to the literature detailed in chapter 3.
Chapter 11	The final conclusions derived from this research, and suggested areas for future research within this domain and research topic.

Table 1.1 - Overview of Thesis Chapters

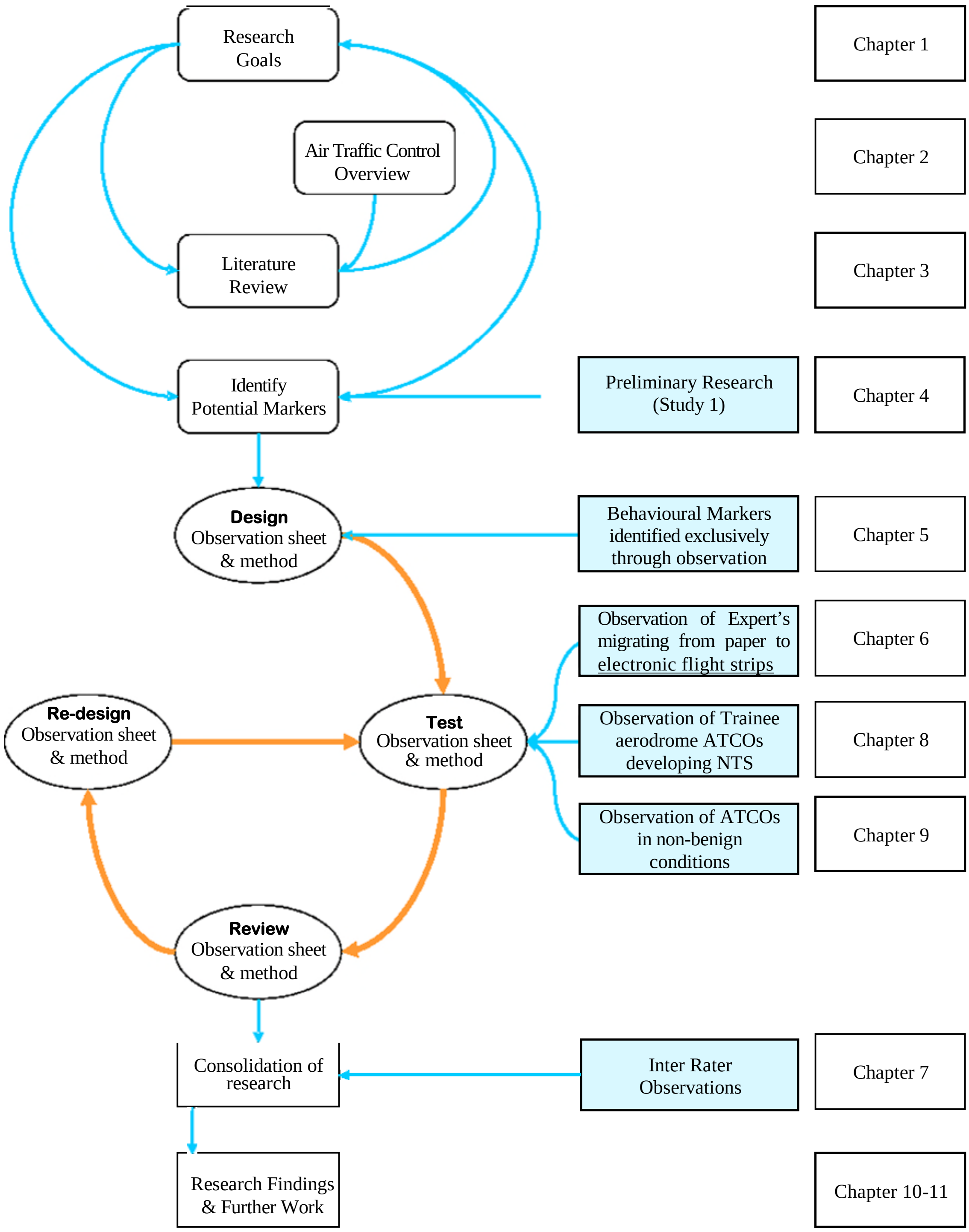


Figure 1.1 - Diagram of thesis structure

1.6 PUBLISHED WORKS

The following publications and presentations have been produced under this research:

- Thompson D.J. (2011). Behavioural development of Air Traffic Control Trainees during Aerodrome training. UCL London Communications Symposium. Robert's Building, Torrington Place, WC1E 7JE. 8th September 2011.
- Thompson D.J. (2011). Behavioural Markers of User Development: during ATC training. UCL Communications EngD Poster Competition. 25th March 2011.
- Thompson D.J. (2010). Behavioural Markers of Controller Development with Electronic Flight Progress Strips. UCL London Communications Symposium. Robert's Building, Torrington Place, WC1E 7JE. 10th September 2010.
- Thompson D.J. (2010). Behavioural markers of user development with a new Air Traffic Control System. UCL Communications EngD Poster Competition. 25th June April 2010.
- Thompson D.J (2010). Behavioural markers of user development with a new Air Traffic Control System. Proceedings of the Institute of Ergonomics and Human Factors 1st Doctoral Consortium. 19th May, 2010. University of Nottingham. http://www.ergonomics.org.uk/events/doctoral4consortium
- Thompson D.J. (2009). Predicting Behavioural Change Resulting from System Change. UCL Communications EngD Poster Competition. 3rd April 2009.
- Thompson D.J. (2008). The Development and Use of Behavioural Markers of Performance for Air Traffic Control (ATC). UCL Communications EngD Symposium, 18th September 2008

1.7 ENGD TAUGHT COMPONENT

An Engineering Doctorate (EngD) comprises both a research component (75%) and taught component (25%); as a consequence it is a four year course. The UCL Communications Engineering doctorate has a modular requirement of 16 masters level modules, however other qualifications and experience may be taken into account, reducing the required number.

The author's modular requirements as stated by the Communications EngD exam board has been the study of ten non-compulsory masters modules. The following UCL modules were undertaken to satisfy this requirement between 2008 and 2012. The modules were selected by applicability and utility to the thesis, areas of knowledge that merited development, modules of utility and benefit for the development and maturation within the industrial business context, and finally for general interest.

- Personal & Professional Management Skills
- Usability Evaluation Methods
- Project Management
- Applied Cognitive Science
- Advanced Experimental Design and Analysis (Open University)
- Strategy, Marketing and the Business Environment
- Finance and Product Management
- Customer Service, Operations and Planning
- Global Aspects, Innovation Management, People Management and Organisational Design

Appendix A02 contains further detail on these modules, and Appendix A03 details the Roberts points gathered during this research which fulfils the final taught element of the EngD course.

SUMMARY OF CHAPTER 1

In this chapter, an overview of the research contained within this thesis is provided, including the structure and linkage between chapters. Detail of the taught component for the Engineering Doctorate (EngD) is presented, in addition to the presentation and publication of this work.

CHAPTER 02 – OVERVIEW OF THE AIR TRAFFIC CONTROLLER ROLE IN THE UNITED KINGDOM

"Improvise, Adapt and Overcome"
(Unofficial motto of the US Marine Corps)

The purpose of this chapter is to provide the reader with a short introduction into the ATC domain including the types of services an Air Traffic Controller (ATCO) undertakes, specific information regarding NATS (the Industrial sponsor of this research), and the technical challenges the ATCO community in general is facing through the introduction of next generation systems and technologies.

Air Traffic Control provides a variety of different services to the aviation community in order to deliver the safe and efficient use of the UK's airspace. The research contained in this thesis concerns the behaviour of Air Traffic Controllers (ATCOs) within this complex safety critical domain.

Although this research focuses upon Air Traffic Control within the United Kingdom, there is a high degree of consistency across ATC operations globally, with a variety of standards maintained by the International Civil Aviation Authority (CAA), European Aviation Safety Agency (EASA), and the International Civil Aviation Organisation (ICAO). Therefore much of the NATS operational description contained within this chapter is generic and is applicable to other Air Navigation Service Providers (ANSPs).

2.1 OVERVIEW OF UK AIRSPACE

There are essentially two types of airspace within the UK, controlled airspace where maintaining vertical and lateral separation is the responsibility of the ATCO; and uncontrolled airspace where pilots within this airspace maintain their own separation using visual flight rules (unless in receipt of a supplementary advisory service from ATC). As shown in Table 2.1, there are various classes of airspace. These classes of airspace are used in different geographical areas and phases of flight.

Within these classes of airspace, different separation standards exist, largely reflecting differences in speed attained at higher altitude and limitations in technologies such as radar provision (Eurocontrol, 2012).

Airspace Class	Detail of airspace classification
Class A	Airways with a maximum 'flight level' of FL195 (approximately 19,500ft)
Class B	No UK airspace has been classified within this category.
Class C	All UK Airspace above FL195, with a few specific exceptions.
Class D	Airspace surrounding principal airfields (known as control zones)
Class E	Scottish TMA airspace below 6,000ft, Belfast TMA and the Scottish control zones (outside Glasgow and Prestwick control Zones).
Class F	UK airspace with an ATC advisory service only.
Class G	Unregulated airspace, where ATC advisory or information services may be available optionally on request.

Table 2.1 - Classification of UK Airspace. (Adapted from Eurocontrol, 2012)

2.1.1 Flight Categories

A wide variety of different aircraft fly within the UK airspace. Each of these aircraft is assigned a dynamic flight category, which prioritises the service received by ATC when in controlled airspace. Table 2.2 presents a definition of each of these flight categories.

Flight Category	Definition of Flight Category
Category A:	Aircraft in emergency (e.g. engine fault, fuel shortage, seriously ill passenger).
Category B:	Flights operating for search and rescue or other humanitarian reasons. Other flights, including Open Skies Flights, authorised by the CAA. Police flights under normal operational priority.
Category C:	Royal Flights, notified flights carrying visiting Heads of State.
Category D:	Flights notified by the CAA carrying Heads of Government or very senior government ministers.
Category E:	Flight check aircraft engaged on, or in transit to, time or weather critical calibration flights.
Normal Flights	i) Flights which have filed a flight plan in the normal way and conforming with normal routing procedures. ii) Initial instrument flight tests conducted by the CAA Flight Examining Unit.
Category Z:	Training, non-standard and other flights.

Table 2.2 - UK Flight Categories – CAP 493 Part 1. (Adapted from CAA, 2012)

2.1.2 Phases of Flight

It is useful to understand the various ATC services from the perspective of a typical flight from a UK airport, flying through controlled airspace. These phases are pre-flight, take-off, en-route, approach, and landing.

Pre-flight

With the aircraft on the ground the flight crew will check the aircraft systems status, any significant weather or considerations en-route; in order to finalise and file a flight plan with ATC. This flight plan is processed by ATC systems in order to generate a Flight Progress Strip (FPS), which includes the aircraft type, assigned call sign, airline, destination and route; the FPS is then used by ATCOs during the flight (Figure 2.1).

When ready to depart, the flight deck will contact ATC for permission to start and push back; traffic level permitting clearance will be given, which includes an estimated departure time. GROUND movement controllers will instruct the aircraft towards the departure runway following numbered taxi ways. The departure route will be usually be confirmed by the pilots at this stage, generally following a predefined Standard Instrument Departure route (SID). Once the aircraft is at the final hold-point prior to the runway, the FPS is passed to the AIR controller, who manages the departure (and with single runway operations the arrival) of aircraft.

There are up to five positions within an air traffic control tower, and these positions may be grouped together during low traffic, or divided for periods of high aerodrome activity. These are the GROUND positions (Ground Delivery Manager, Ground Manager), the AIR positions (Arrivals Manager, Departures Manager), and finally the Tower Supervisor. There may also be one or more Air Traffic Control Assistants to support the activities of the controlling staff.

Figure 2.1 - Edinburgh Tower ATCO using paper strips

Take-off

The AIR controller will instruct the departure aircraft to line up on the runway when a wake vortex and separation minima gap is available. The aircraft will then be instructed to take-off and is provided the new frequency to contract radar control (approach) once airborne. With the exception of the London area which has a combined control centre for Heathrow, Gatwick, Stansted, Luton and London City, most UK airports have their own approach radar facility.

When safe, the pilot will power up, accelerate, and take off, with an instruction to activate a transponder device inside the aircraft, and to contact Approach radar. The transponder sends a unique signal which helps pair the radar track with the flight identity of the 'blip' on the radar screen. The pilot will contact the approach radar ATCO using their radio. Communications between air and ground employ the use of radio frequencies to transmit spoken messages, following standardised phraseology (Duke, 1997).

The pilot will receive instructions from the radar controller which include changes to speed, heading and altitude. Aircraft will generally follow pre-defined departure routes through an aerodrome control zone or terminal manoeuvring areas receiving instructions from Air Traffic Control as they look to climb into higher airspace and receive an en-route service.

En route

Unless the aircraft departs controlled airspace during its flight, the flight will receive instructions from the ATCO as it travels through one or more en-route 'sectors', towards its destination (a sector is a three dimensional volume of airspace managed by one or more ATCOs). Details of the aircraft including its current heading, speed, and altitude will be passed from ATCO to ATCO across sector boundaries with the details recorded on flight strips (either paper or electronic). The ATCO will provide the pilot with important weather and traffic information, and aim to provide a safe but expeditious service.

When the aircraft is approximately 150 miles from arrival at its destination, the aircraft will begin to receive instructions to descend. For large airports and busy airspace the aircraft may be instructed to join a hold, which vertically stacks a number of aircraft in a confined area and permits the smooth and continuous feed of aircraft to an airfield to maximise capacity. Streams of aircraft will be merged into a single flow, where the aircraft will make the last few turns onto approach.

Approach

Approach radar control may be performed by one or more ATCOs depending on the complexity of the airport (Figure 2.2). Their role is to use the large radar screen and flight strips as tools to help sequence arriving aircraft into an efficient landing order taking consideration of the type of aircraft they are, and their wake vortex spacing requirements. The pilot will deploy gear and flaps to efficiently manage the residual energy of the aircraft reducing speed and height as instructed. Once the aircraft has turned onto final approach and established a stable approach either by using visual means or Instrument Landing System (ILS), the aircraft will be instructed to contact the tower ATCO on a different radio frequency. This is usually around 5-10 miles from touch down.

Figure 2.2 - Swanwick Terminal Radar Controller using Paper Strips

Landing

The final phase of flight is managed by the tower ATCO. Their role is to ensure that the runway is clear of vehicles and aircraft and safe to land upon. The tower ATCO will inform the pilot of the meteorological conditions including surface wind direction and strength, and the condition of the runway (wet, dry etc.). In good visual conditions the tower ATCO will look out of the tower to monitor the situation, but a surface and air radar display is provided to assist them at night or in poor visual conditions. The aircraft will then be given approval to land by the ATCO. The pilot will perform final adjustments for landing speed and guide the aircraft into land. Once safely landed the pilot will be instructed to depart the runway using a specific taxi-way; and instructed to contact the ground ATCO on a different radio frequency who will then direct the aircraft safely to the gate.

2.2 OVERVIEW OF NATS OPERATIONS

NATS (formerly known as National Air Traffic Services) was formed in 1962 to manage traffic in the UK. In 2001 it was privatised from the CAA, with a number of owners that comprise a group of seven UK airlines, Her Majesties Government, BAA, and employee ownership.

NATS delivers two fundamental services within the UK; the first is the sole provider of en-route ATC operations which it manages from two locations (Prestwick, and Swanwick). The second is the provider of approach and aerodrome operations service provision to a large number of UK airports including London Heathrow, Gatwick, Birmingham, and Manchester (Figure 2.3). Other organisations such as SERCO manage a number of other UK airports such as East Midlands, Exeter, Leeds International, and Bournemouth.

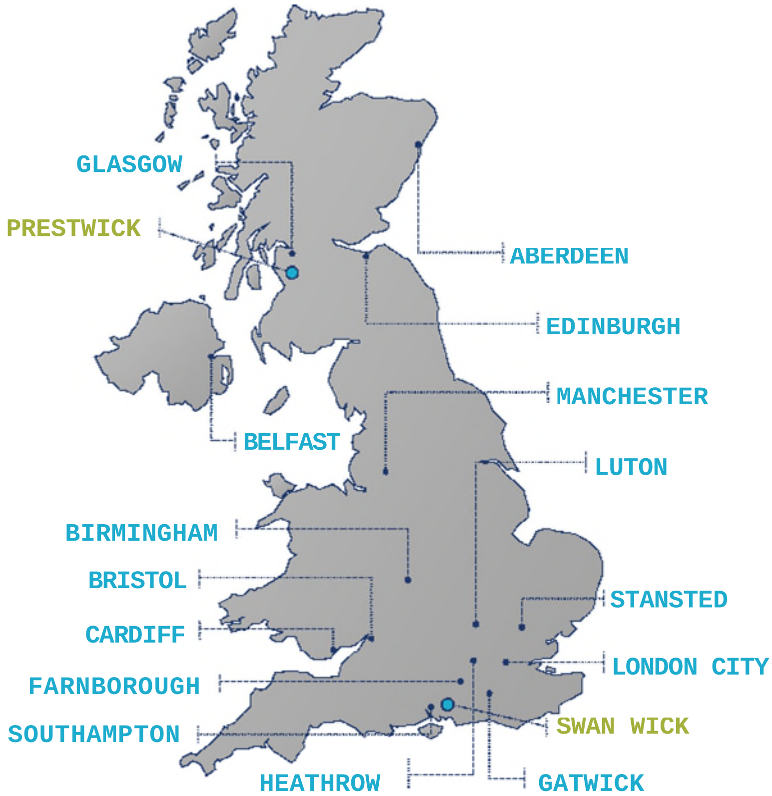


Figure 2.3 - Overview of NATS UK operations[1]

There are two control centres which operate from the Swanwick facility. Firstly, there is the London Terminal Control Centre (LTCC) which is responsible for the arrival and departure radar support to the London airports. Given the close geographical proximity of these airports and their holds, their control zones combine to form a 'terminal manoeuvring area' (TMA). Airspace is controlled by

[1] Correct as of April 2013.

a single ATCO within Terminal Control, who is supported by assistants and supervisors.

Secondly there is the London Area Control Centre (LACC), which provides the en-route area control service for England and Wales. There are two ATCO positions within area control, for each sector of airspace. The tactical ATCO controls aircraft within their sector, and they are supported by the planner ATCO who co-ordinates aircraft into and out of the sector. The two sector ATCOs are supported by an assistant.

The Prestwick centre contains TMA functions (similar to LTCC) for the Edinburgh TMA, and Manchester TMA, As well as the en-route operation for Scottish airspace. A partnership with Ireland provides a joint oceanic service which accounts for 40% of transatlantic air traffic. There are other functions NATS provides, which include the provision of equipment, systems and facilities for the military. Finally a flight information service is delivered to support general aviation from Swanwick.

2.2.1 Live versus simulation environments

Notwithstanding the quality of NATS real time test and evaluation simulation facilities, there are differences between the simulation and live operational environments. The simulation environment lacks fidelity and does not result in deleterious safety outcomes when errors are made, which reduces the safety criticality of decisions and actions and the ultimate pressure and level of intensity experienced. In addition the simulator is very ‘clean’, and does not contain the richness of real world factors which may invoke different behavioural responses, such as the behaviour of foreign pilots who may be difficult and slow to communicate with.

2.3 AN OVERVIEW OF FLIGHT STRIPS

The primary source of information that an ATCO uses in order to manage Air Traffic is paper flight progress strips. Each piece of paper contains all the core information needed to provide an ATC service to that aircraft. This information includes the call sign, and the route or destination. In flight, the strip will indicate the heading, speed, and altitude that it enters the sector’s airspace.

Where a change is made to the aircrafts journey (e.g. heading, speed, altitude), the amendment is written on the paper strip by the ATCO using a set of shorthand numbers and symbols. Different ATC positions will have different configurations for the paper strip, depending on their requirements (Figure 2.4, Figure 2.5).

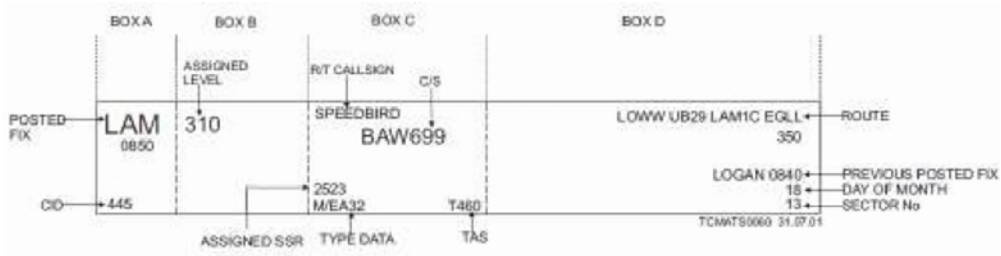


Figure 2.4 - Arrival Strip Format – London Terminal Control Centre (LTCC) – NATS (2002)

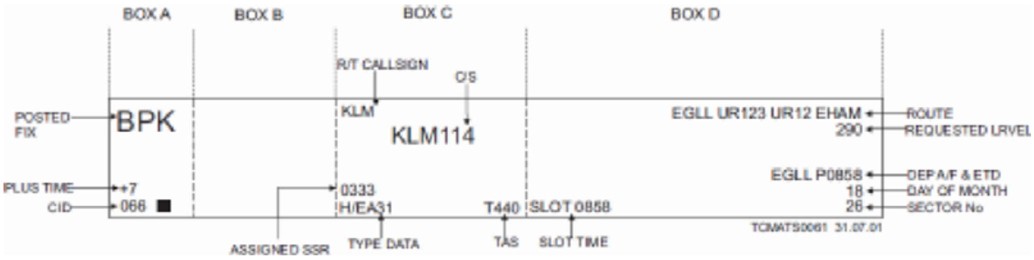


Figure 2.5 - Departure Warning strip - London Terminal Control Centre (LTCC) – NATS (2002)

There are certain ATC positions, for example North Sea helicopter routes controlled from Aberdeen, which rely on paper strips without the use of a radar screen. In a loss of radar fallback scenario, radar ATCOs are trained to manage the airspace using their paper strips.

Paper strips are contained within plastic strip holders, and placed within a strip board in front of the ATCO (Figure 2.6). The ATCO will maintain the currency of the strip board, organising them as appropriate, and adding new strips and removing old strips when no longer required.

Figure 2.6 - Farnborough Tower ATCO using Paper Flight Progress Strips in the centre of the ATCO working position

2.3.1 Electronic Flight strips

There are three electronic flight strip systems employed within NATS operations, two of which have been systems used by ATCOs observed over the course of this research. Electronic Flight Progress Strips (EFPS) is a system used within a number of NATS towers, including all London airports, and Aberdeen, Glasgow and Edinburgh (Figure 2.7). The EFPS system has been designed to replicate closely the paper flight strip structure and format, as a consequence the transition period between the two is accomplished in a short time when compared to other electronic flight strip systems.

Since 2005, EFPS has been introduced into several UK Air Traffic Control operations to replace paper flight strips. Currently, 9 out of the 15 NATS towers operate with EFPS[2]. At the time of this research, 4 NATS control towers had already had this system introduced into their operations (Heathrow, Gatwick, Stansted, Luton); whilst 5 control towers were undergoing paper to electronic flight strip system transition (Glasgow, Edinburgh, Aberdeen, London City, and Manchester).

2 Correct as of February 2013.

Figure 2.7 - Electronic Flight Progress Strips (Heathrow Arrivals Runway ATCO)

The interim Future Area Controller Tool Set (iFACTS) is an electronic flight strip system which was introduced into Swanwick AC operations in 2011. Unlike EFPS which recreates paper flight strips with fewer changes in layout and design, iFACTS distributes the flight data information across the iFACTS main window (which is positioned where the paper strip board would appear), and the radar display. iFACTS contains decision support tools which include medium term conflict detection, and 'what if' probes which allow changes of flight parameters to be tested for suitability prior to issuing any instructions to the aircraft.

2.4 CHALLENGES FACING THE INDUSTRY

Within the European Union, wide scale harmonisation of systems, airspace and procedures is being undertaken under the EU programme Single European Sky Air Traffic Management (ATM) Research (SESAR). SESAR's main drivers are to improve efficiency and safety through the implementation of a variety of new systems and processes. These include trajectory based controlling, the wider use of electronic data and strips, the provision of decision support and conflict detection tools, and the efficient dynamic re-allocation of airspace (e.g. civil usage of military danger areas etc.).

This is a challenging time in the industry, and for the ATCO. SESAR introduces increased automation and reduced manning, in addition to a variety of new systems which must be designed, trained for, and seamlessly implemented into operations. Within NATS, iFACTS and EFPS represent two such systems. Methods and techniques to evaluate the human performance of ATCOs using these new systems are extremely important and of great benefit. The research undertaken in later chapters explores the utility of behavioural observation to provide insight into the maturation of ATCO Non-Technical Skills when learning the ATC task as a trainee, and for qualified ATCOs transition training to electronic flight strips.

SUMMARY OF CHAPTER 2

In this chapter, a basic explanation of air traffic operations has been provided as background information to the research domain of this thesis. An overview of the types of ATC positions, and the roles performed has been included. Finally, contextual detail regarding the use of flight strips has been provided. In the next chapter, an exploration of behavioural observation methods is undertaken, which includes key work undertaken within the ATC domain.

CHAPTER 03 – A REVIEW OF THE BEHAVIOURAL MARKERS LITERATURE

"There are known knowns; there are things we know that we know. There are known unknowns; that is to say there are things that we now know we do not know. But there are also unknown unknowns; there are things we do not know, we don't know"
(Donald Rumsfeld)

This chapter explores the origins of behavioural markers observation and provides the rational for the development and understanding of Non-Technical skills in a variety of safety critical domains. It presents the principal systems which have been developed over the last 15 years including where possible an examination of the method used in their development and application. Finally, a discussion section at the end of the chapter explores the commonality and learning points which may be taken from these systems, and the opportunities for further application within the ATC domain.

3.1 OBSERVATIONAL RESEARCH

The breadth of observation is diverse. Humans can make direct observations themselves or use technology to provide perspective into areas too extreme for our senses to probe. Within the human sciences, observation is used to study man's interaction with equipment and interfaces. Observation can examine the individual, a team, even crowds. More widely observation is used to select and recruit, to train, to monitor; evidence collected can be used for discipline and dismissal or reward and promotion.

When undertaking the assessment of human performance, observation may be employed in a variety of ways. Meister's (1985) hierarchy of observational methods conveys three principal forms of observation, namely: self-observation; and the qualitative; or quantitative observation of others.

Meister (1986) goes on to indicate that the selection of an appropriate method of observation comes about when answering a number of fundamental questions such as; who will be doing the observation, where will the observations be made, what is to be observed, what the resources available are. There is also the consideration of key elements such as reliability, accuracy, frequency. Figure 3.1 depicts the three principal forms of observation, and the subcategories of observation structured within them. Behavioural markers observation, which is discussed in detail within this chapter and thesis falls under '*direct observation*[0] within this hierarchy.

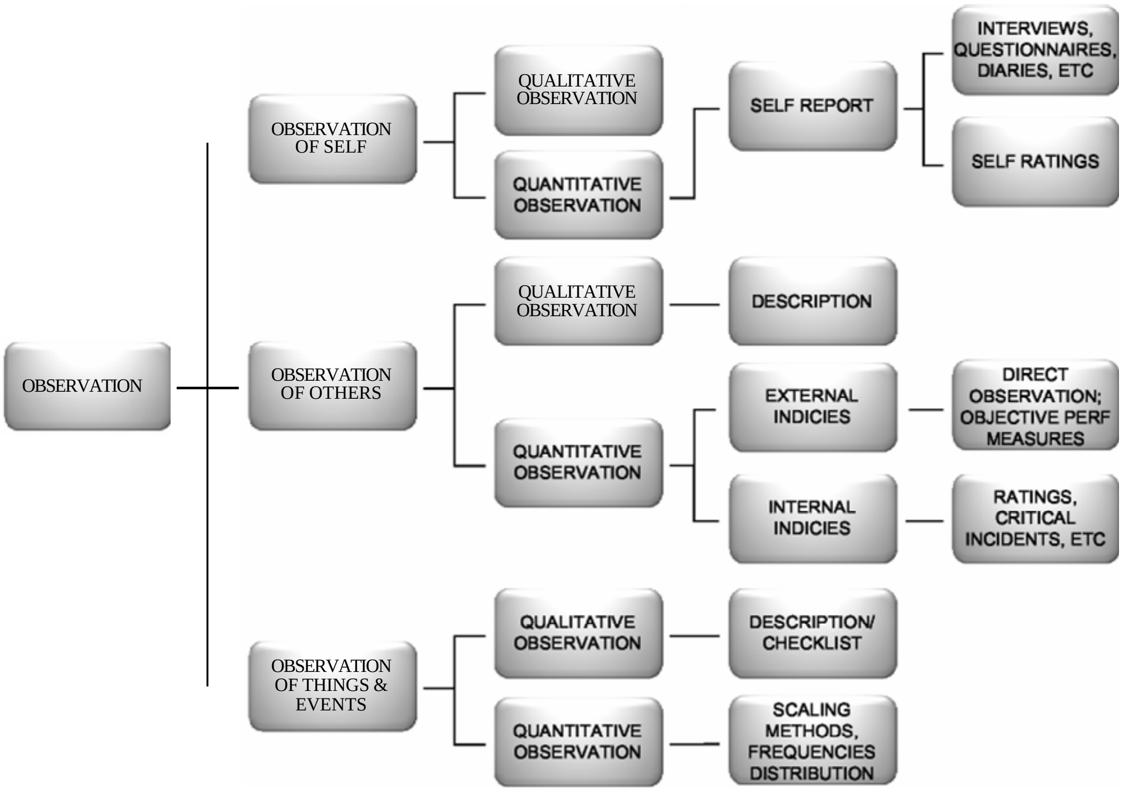


Figure 3.1 - A hierarchy of observational methods (Meister (1985, p.298) Permission to reproduce this diagram has been granted by John Wiley & Sons, Inc.

3.1.1 The impact of Observation on Performance

From 192441932, a series of experiments were conducted at the Western Electric's Hawthorne plant. The purpose of these studies was to identify social and environmental conditions which may impact on upon productivity. One area investigated was ambient lighting within the factory and social areas used for coffee breaks. When light levels were increased, productivity against the control condition increased. However when light levels were decreased, productivity against the control condition also increased. Finally when the light levels were returned back to original levels (same as the control condition), productivity levels also increased. This curious affect could not be explained by early researchers (Kornblum, 2011; McCarney et al 2007; Fox et al 2008). In the 1950's the results were revisited by Henry A Landsberger who concluded that the reason for the increase in productivity was the increased attention shown in the workforce as a consequence of the observations (Khurana, 2009). Put simply, the workers enjoyed the attention given to them, and this reflected in their levels of effort exerted to complete their tasks.

As an observational researcher, one must always therefore be mindful that your very presence may have an impact on the behaviour and task performance of those whom you are observing. Within the Air Traffic Control domain, ATCOs are extremely familiar with the use of observation. Observation is the primary means of measuring competency during training, licensing and validation. It is used during the testing of new systems, procedures, and airspace design; and observations are made both in the simulated environment e.g. for Training in Unusual Circumstances and Emergencies (TRUCE), and the live operation environment e.g. Unit Competence Examination. Although, therefore there is the potential for impact on performance, it is somewhat more modest given that this is normal practice within the domain of ATC.

3.1.2 Overt Versus Covert Observation

The research in this thesis considers the observation of behaviour purely in terms of ‘overt’ observation. In this context the person observed is explicitly aware of the purpose of the observation, and has given their prior consent. However it is useful to recognise that observation may be used in other circumstances on a ‘covert’ basis. In these situations, it may be employed in order to assess an individual’s compliance with rules (van de Mortel et al, 2000), evaluate discrepancies between subjective and objective viewpoints regarding performance (Brokaw et al, 2004), or for the evaluation of passenger behaviour in order to identify potential transport security threats (Gordon and Fleisher, 2011).

3.1.3 Crew Resource Management

The origins of Crew Resource Management (CRM) date back to a series of significant aviation disasters in the late 1970’s (Helmreich et al 1999; Thatcher, 2007). Investigation into several significant accidents, had revealed critical errors due to poor team working, leadership, and communications were the main causes; rather than technical airmanship proficiency (Salas et al, 2001). These findings solidified efforts across the industry to learn, understand, and improve these ‘*1on-Technical Skills'*, in order to reduce the likelihood of further accidents occurring (Helmreich & Foushee, 2010).

Over many years the discipline of CRM has been enhanced and expanded, and has gained increasing acceptance across the airline industry. Indeed CRM skill training and refreshment is now a mandatory training component for UK pilots (CAA, 2002). The CAA’s definition of the skills that commonly fall within the domain of CRM are those of communications, situation awareness, problem solving, decision making, and teamwork (CAA 2006a; CAA, 2006b).

Within Air Traffic Control, the Non-Technical Skills categorised under CRM are considered an important element to delivering optimum performance. Within the ATC domain, the concept is referred to as Team Resource Management (TRM), a term which reflects the importance of the team unit in the successful delivery of Air Traffic Management (Woldring et al, 2005).

3.1.4 The Origins of Non-Technical Skills

Non-Technical Skills (NTS) is a generic non domain specific definition of the skills that are synonymous with CRM. An early definition of NTS is *"the cognitive and social skills of flight crew members in the cockpit, not directly related to aircraft control, system management, and standard operating procedures"* (Flin & Martin et al, 2003 p.96). Over time the definition has broadened and expanded to include other complementary NTS. The most current definition of NTS is Professor Rhona Flin’s *"the cognitive, social and personal resource skills that complement technical skills, and contribute to safe and efficient task performance"* (Flin et. al, 2008, p.1).

Table 3.1 presents the examination of 14 significant incidents which have occurred over the last 30 years, each of which have revealed poor NTS as a contributory factor (Flin et al, 2008). These incidents ranged across safety critical areas such as nuclear power (3 mile island, Chernobyl), maritime (Herald of Free Enterprise), police (Hillsborough), oil and gas (Piper Alpha), aviation (Kegworth), and healthcare. NTS failures include poor teamwork, situation awareness, communications, leadership, fatigue, and decision making.

Year	Industry	Incident	Non-Technical skills failures
1979	Nuclear Power	Three Mile island nuclear power plant release	Problem solving, teamwork Situation awareness
1986	Nuclear Power	Chernobyl Nuclear power plant release	Decision-making, Situation Awareness, personal limitations
1987	Maritime	Herald of Free Enterprise Ship sails with bow doors open and capsizes	Team co-ordination, situation awareness
1988	Oil and gas production	Piper Alpha Oil platform explosion	Communication, leadership, decision-making, team handover
1988	Military	USS Vincennes Warship destroys passenger plane	Team co-ordination, decision making
1989	Police	Hillsborough Police response to football crowd being crushed	Communication, situation awareness, leadership
1989	Aviation	Kegworth Plane crash Wrong engine shut down	Situation Awareness, decision-making
1990	Maritime	Scandinavian Star Response to ship fire	Teamwork, leadership
1994	Health Care	Betsy Lehman Chemotherapy overdose	Situation Awareness, Decision making
1996	Transport	Channel Tunnel Response to fire in tunnel	Communication, stress, teamwork co-ordination
1998	Petrochemical	Esso Longford Refinery explosion	Communication (shift handover), situation awareness
2000	Healthcare	Graham Reeves – wrong kidney removed	Situation awareness, teamwork, leadership
2001	Healthcare	Wayne Jowett Chemotherapy site error	Decision making, Situation Awareness, Communication
2005	Petrochemical	BP Texas City Refinery explosion	Leadership, decision making, fatigue, communication

Table 3.1 - Safety Incidents and Non-Technical Skill Failures (Flin et al, 2008)

Specifically within the transportation domain, a study of 1020 work-related traumatic driving deaths occurring throughout Australia between 1982-1984, 91.2% were considered to have behavioural factors as underlying causes of the incident (human error, poor work practices, poor supervision, poor training). In 42.0% of fatalities, unsafe work practices were identified as major factors in the cause of the incident (Williamson & Feyer, 1990).

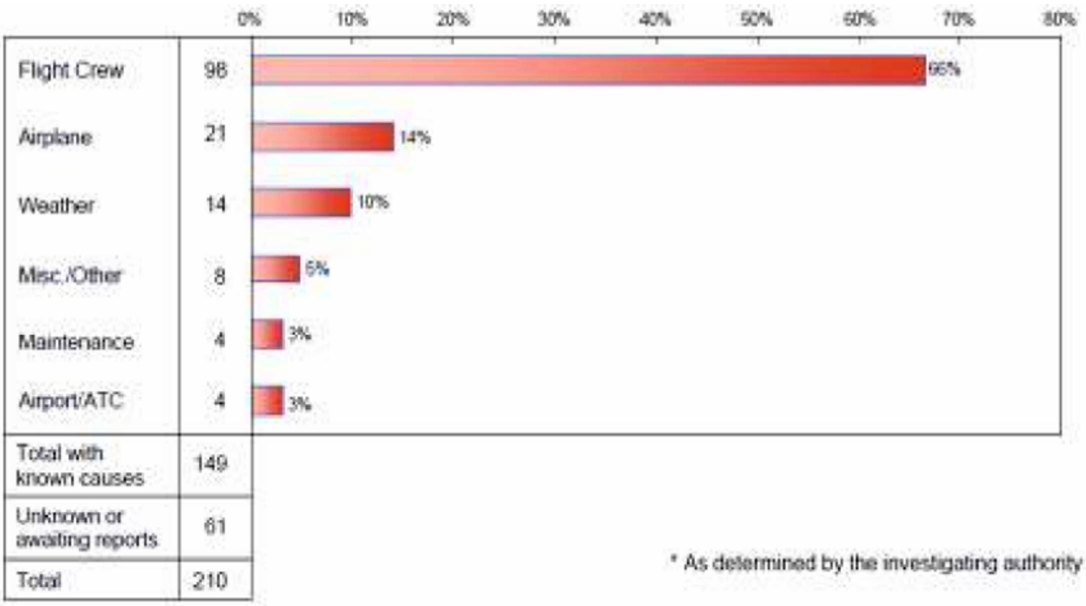


Figure 3.2 - Worldwide Commercial Jet Hull Loss Accidents (1 992-2001): Primary cause (Boeing, 2001)

Within the aviation industry, Boeing has undertaken extensive analysis of all hull losses occurring since the 1950s. Figure 3.2 presents an analysis of accident primary causes for worldwide commercial jet fleet hull losses occurring through the years 1992 to 2001. Of the 149 accidents with known causes, 96 (56%) have been attributed to the actions of the flight crew as primary cause, (Boeing, 2001). In a separate review of aviation incidents, the University of Texas found 51 separate events where poor CRM NTS were a contributory factor (Helmreich et. al, 1995).

3.1.5 Factors which impact performance

When observing the performance of others engaged in a task, the majority of measures invariably focus upon the technical performance of the individual. However as shown in the CRM literature, performance on complex tasks often does not purely rely on technical skills e.g. motor skills, but also may require complementary abilities. Oprins et al ‘s (2006) *'ATC Performance model'* of ATC recognises the impact and relationship of additional performance shaping factors and NTS under the term *'influencing factors'* (Figure 3.3). Although largely focusing upon the non-observable cognitive processes of an ATCO undergoing training, it is clear that these factors are considered to impact both the decision making and action components of task processing.

In addition to technical and non-technical skills, there are other factors which may also impact the task such as the physical and state of the ATCO. It is well known that fatigue has a significant impact on task performance, and has led to many accidents, (Kumashiro, 1990; Sanders & McCormick, 1992). When specifically concerning the psychological and physical state of an individual, and the impact upon behaviour and performance, Oprins et al (2006, p.299-300) states *"A person can be competent but performance can still be insufficient due to environmental or personal influences. A temporary or long-term personal or psycho-physiological state may influence regulation of performance, caused by factors such as stress, motivation, or fatigue".*

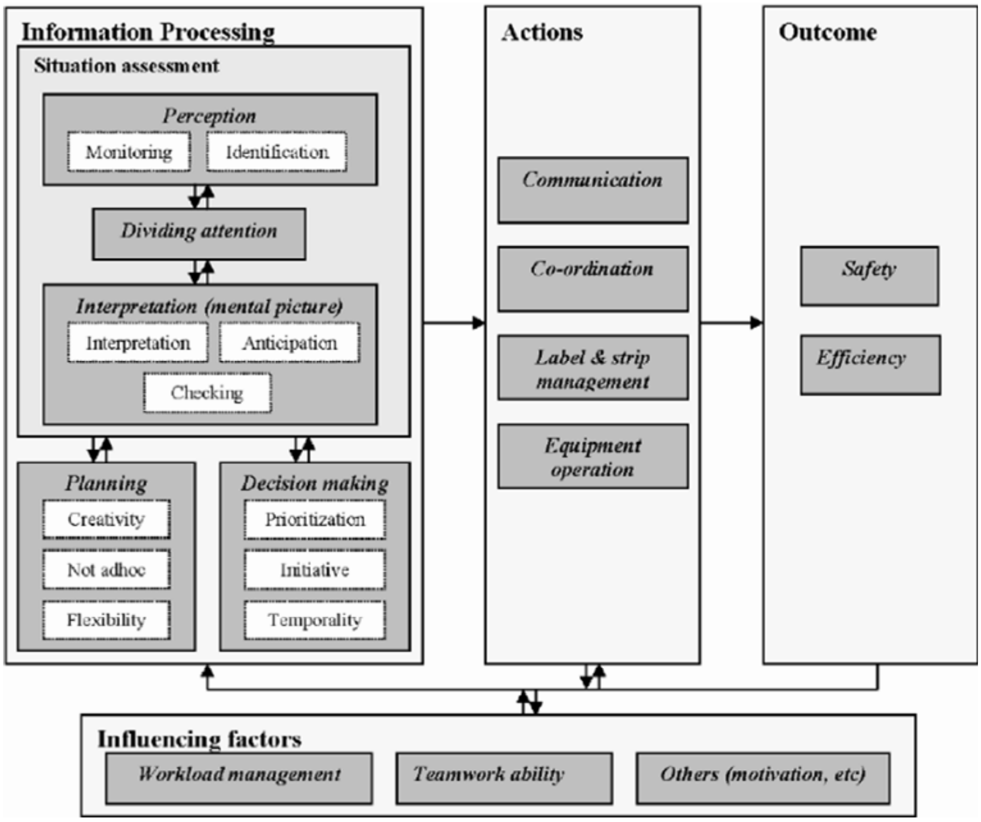


Figure 3.3 - 'ATC Performance Model' (Oprins et al, 2006) Permission to reproduce this diagram has been granted by Taylor & Francis

The design of the systems in use, and the environmental and task situation may also impact the delivery of the task. Therefore it is important to be aware of any performance shaping factors when undertaking observational research. Figure 3.4 illustrates the technical and non-technical skills which contribute to overt behaviour (action) and task performance. However Figure 3.4 also presents other contributory factors such as physical and psychological state; the task situation, and environment which may also impact.

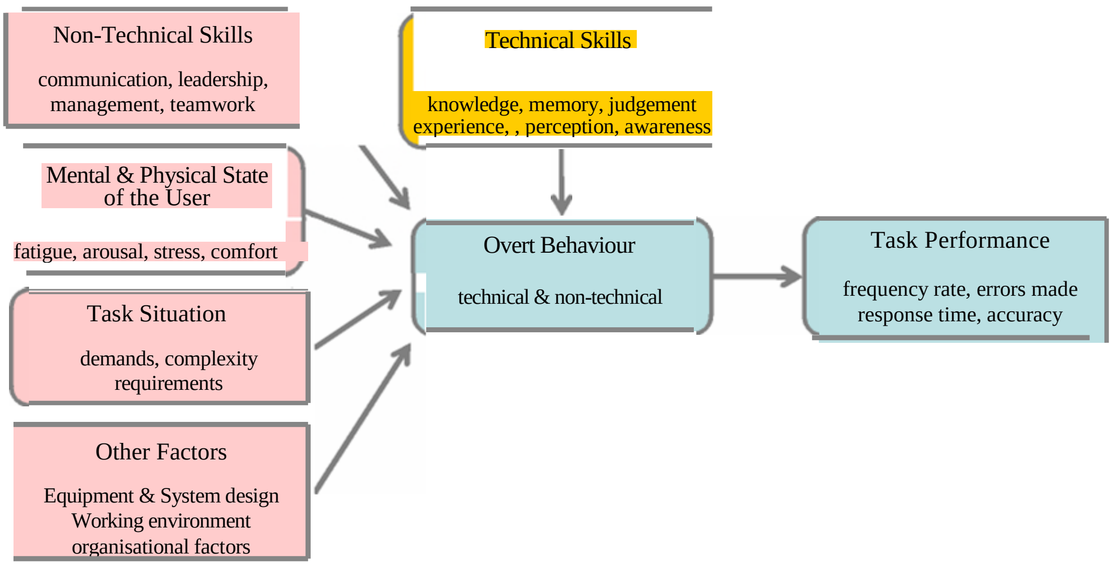


Figure 3.4 - Factors contributing to Task Performance

3.1.6 Observation of Psychological and physical State

Within other domains such mental health care and education, behavioural observation (a.k.a direct observation) is a popular assessment technique. Assessments are undertaken in order to assess either physical or mental aspects of the person observed. The behavioural observation systems are predominantly either checklists or Behavioural Observation Systems (BOS) (Stanton, 2005). Application areas include the assessment of pain in patients (Ahles et al, 1990, Labus et al, 2003); the evaluation of child behaviour problems (Nock & Kurtz, 2005); and the assessment of comfort in users across various workplace layouts and working postures (Lindegård et al, 2005).

An interesting piece of observational research undertaken by Jensen and Asren (1998) examined the subjective ratings of workload made by operators in automated plants against observations of workload made through behavioural observation. These observations focused upon the amount and type of upper body movements and lower body movements displayed, in addition to details regarding communications which took place. These observations were made in order to determine objective levels of task workload. The authors suggest that a significant correlation was found between the subjective and observational measures, unfortunately there is little further work published in this area.

With regards to behavioural marker systems (discussed in detail later in this chapter), the systems reviewed do not readily focus upon observable characteristics such as the mental or physical state of the user. Indeed key design principles and guidelines which have been produced specifically preclude the inclusion of markers which focus upon the attitude of the person observed (Klampfer et al, 2001). However, the inclusion of psychological and physical assessment may provide further insight as to the current state of the user; and in turn their likely potential performance, and is an area that warrants further exploration.

3.1.7 Introducing Behavioural Markers

Poor CRM has been identified as a key impact in many incidents including those in ATC; as a result CRM is taught within the wider training syllabus within many safety critical domains. The Behavioural Markers observational technique is a structured method of observation that was developed in order to assess these CRM skills both for individuals and teams, (Helmreich et al, 1994). A behavioural markers observation system allows training instructors to evaluate the effectiveness of their training, and how well the trainees demonstrate the CRM skills and competencies associated with high levels of performance and safety.

Klampfer et al (2001, p.10) define behavioural markers as *"Observable, non-technical behaviours that contribute to superior or substandard performance within a work environment (for example, as contributing factors enhancing safety or in accidents and incidents in aviation)"* Over time the scope and use of behavioural markers observation has expanded and developed.

A collective of pre-eminent behavioural markers researchers within a 2001 Workshop entitled *'Enhancing Performance in High Risk Environments[0]* defined their role and usage of behavioural markers as follows (Klampfer et al, 2001, p.11):

- *"To enable performance measurement for training and assessment, evaluation of training, safety management, and research*
- *To highlight positive examples of performance*
- *To provide a common vocabulary for training, briefing and debriefing, communication, regulation, research and to connect different domains of safety (e.g., incident analysis and performance tracking)*
- *To build performance databases to identify norms and prioritise training needs*
- *To compare sub-groups in organisations (e.g., aircraft fleets, etc.)*
- *To give feedback on performance at individual, team, organisational, and system level*
- *To establish co-operation between safety/quality, training, and operations"*

3.1.8 Benefits of observing behavioural markers

As a human performance measure the technique of Behavioural Markers observation has a number of distinct benefits in contrast to other measures.

Firstly, it is a largely unobtrusive method by which performance may be evaluated (unlike techniques such as the electroencephalogram (EEG)). Secondly, the technique requires low effort on behalf of the person observed. In the domain of ATC, where high workload is often a significant component to the task, techniques that do not add to this burden are welcome. Thirdly, minimal equipment other than a pen and check sheet is required (although video capture may enrich analysis). This makes the technique potentially suitable for use in a live environment, although due precaution and care is required when observing in this environment (Flin el al, 2008). Fourthly, observation allows best practice to be assessed – which can be missed when only examining objective data (Meister, 1986).

3.1.9 Limitations of observing behavioural markers

There are a number of limitations that exist with the behavioural markers technique. For instance, not all aspects of performance and behaviour can be assessed due to the rareness of occurrence of some behaviours (e.g. crisis management behaviour), and by its nature this technique focuses upon only those things that can be observed (therefore decision making and planning is largely not possible to observe). The use of scripted crisis scenarios facilitated through actors playing and demonstrating certain behaviours associated within the crisis management environment can enable non-technical skill performance to be evaluated (Gaba et al, 1998; Gatfield, 2008).

In addition the human observer has limitations and can be distracted or overloaded; both of which will dilute the quality of captured data, and they bring their own biases and personal perceptions (Klampfer et al, 2001). Where possible the use of video to record behaviour and analyse in detail later, can mitigate against the effects of live observation, and may also be useful to train observers and to evaluate inter-rater reliability (Sollenberger et al 1997; Yule et al, 2008).

3.1.10 General Principles in the Design of Behavioural Markers

The participating experts of Klampfer et al’s (2001, p.10) workshop identified a number of principles which make for well-designed behavioural markers, and which provide a robust insight into NTS proficiency:

- *"It describes a specific, observable behaviour, not an attitude or personality trait, with clear definition (enactment of skills or knowledge is shown in behaviour).*
- *It has demonstrated a causal relationship to performance outcome.*
- *It does not have to be present in all situations.*
- *Its appropriateness depends on context.*
- *It uses domain specific language that reflects the operational environment.*
- *It employs simple phraseology.*
- *It describes a clear concept.*
- *It must capture the context in which the assessment is made (e.g., crew dynamics and experience, operating environment, operational complexity)"*

Further detail as to why attitude should be precluded is not provided. The most likely reason is to mitigate the potential for individual difference and personality which can impact attitude. When introducing new systems into operations, attitude is very important; as it forms part of the test process in change management. For example, someone who is actively engaged in change management process is unlikely to make derogatory remarks. Negative behaviours may therefore suggest elements of user frustration and dissatisfaction. It is therefore considered prudent to consider attitudes on a case-by-case basis; in order to determine if they may provide additional insight into the overall attitudes of the user population that is being observed.

3.1.11 Design of Behavioural Markers for Surgery

Yule et al (2006b) have produced a useful set of design guidelines for the development of a behavioural markers set within the surgical domain. These design guidelines are broadly similar to those developed by Klampfer et al (2001), although point 5 which concerns the size of the observation tool is new. Although the language and phraseology of Yule et al’s (2006b, p1101) guidelines are focused upon the surgical domain, these guidelines are broadly transferable to alternative domains with a little modification:

1. *"The skills must be applicable to a surgeon's behaviour during the intraoperative phase of an operation.*
2. *The system should comprise specific, observable behaviours that are well defined and contribute to superior or substandard performance.*
3. *The skills and behavioural markers should either be directly observable in the case of social skills or inferred from observing communication or other behaviours, in the case of the cognitive skills.*
4. *The system should be parsimonious and encompass the most important behaviours in the least number of categories and elements possible.*
5. *The rating tool will need to fit on one page, not larger than A4 paper to be of practical use in the operating theatre or high-fidelity simulated environment. This will limit the number of categories and elements.*
6. *The categories and elements should have the maximum mutual exclusivity possible. It is understood that this is only achievable to a certain degree, given the interdependence of the non-technical skills.*
7. *The terminology used should reflect everyday, domain-specific language for surgeons' behaviour, rather than psychological jargon."*

Table 3.2 combines both Yule et al's (2006b) design rules (adapted by the author to form generic non-domain specific statements) with Klampfer et al's (2001) guidelines on behavioural marker design:

Design Characteristic	Detail
Task Centric	The Non-Technical Skills must be applicable to a users' behaviour whilst engaged and employed in the task.
Overt Behaviour	The system should comprise specific, observable behaviours, with a clear definition, that are not attitudes or personality traits; not all skills will be observable on all occasions. The skills and behavioural markers should either be directly observable in the case of social skills or inferred from observing communication or other behaviours, in the case of the cognitive skills.
Performance link	The Non-Technical Skills should demonstrate a causal link to superior or substandard performance.
Concise	The system should be concise and encompass the most important behaviours in the least number of categories and elements possible.
Mutual Exclusivity	The categories and elements should have the maximum mutual exclusivity possible. It is understood that this is only achievable to a certain degree, given the interdependence of the non-technical skills.
Terminology	The terminology used should reflect everyday, domain-specific language for user's behaviour, rather than psychological jargon.
Context	The context surrounding the behaviour should be noted (environment, task factors, etc.)

Table 3.2 - Consolidated Behavioural Marker Design Characteristics (Adapted from Klampfer et al 2001, Yule et al, 2006b)

3.1.12 Identifying Behavioural Markers

Flin et al (2008) identify several methods by which domain specific NTS and underlying markers of behaviour may be identified. Flin et al (2008, p.216) states that the process is in essence two key phases, the first being to "*identify the skills and related behaviours deemed to influence safe and efficient performance*". The second part being to *"refine the resulting list and to organise it into a concise, hierarchical structure or taxonomy"*. Flin et al (2008) goes on to provide a series of five sources that have potential for the identification and development of a BM system for the assessment of NTS (Table 3.3):

Sources for identifying markers	Detail
Source 1:	Published literature on studies examining behaviours that contributed to safety and performance
Source 2:	Documented analysis from organisations in different domains covering competency frameworks, job assessments, task analyses, training programmes, and assessment/appraisal systems
Source 3:	Review and analyses of incident data and incident reports
Source 4:	Questioning of users and other domain experts. This may be through one-to-one discussions, semi or structured interviews, or knowledge capture with a group of experts.
Source 5:	Observation of the users engaged in the task; either in simulation or the real environment
Phase 1:	Identify the skills and related behaviours deemed to influence safe and efficient performance.
Phase 2:	Refine the resulting list and to organise it into a concise, hierarchical structure or taxonomy.

Table 3.3 - Sources of identifying behavioural markers and Phases of consolidation, (Flin et al, 2008)

3.1.13 Principles for Developing a Behavioural Marker System

Table 3.4 presents a number of principles to consider when developing and designing a behavioural marker system all of which greatly affect the accuracy of the technique, and the quality of the data that is collected (Klampfer et al, 2001). In addition to the principles identified in Klampfer et al's (2001) workshop, Flin et al. (2008) identify a number of additional principles which must be considered in order for the design of a behavioural marker system to be effective.

Principles for effective behavioural markers	Detail
Training	Raters require extensive training (initial and recurrent) and calibration.
Domain Specificity	Behavioural marker systems do not transfer across domains and cultures without adaptation (e.g., western markers in eastern cultures, or from aviation to
Implementation	Behavioural marker systems need proper implementation into an organisation, and need management and workforce support. Phased introduction of behavioural marker systems [are] required to build confidence and expertise in raters and ratees.
Purpose	Application of the behavioural marker system must be sensitive to the stage of professional development of the individual, and to the maturity of the organisational and professional culture (e.g., whether used as a diagnostic, training, and/or assessment tool).
Environment	Use must consider context (e.g., crew experience, workload, operating environment, operational complexity)

Table 3.4 - Principles of an effective Behavioural Marker System (Klampfer et al, 2001)

Principles for effective behavioural markers	Detail
Sensitivity;	The system should be based on detectable behaviours that differentiate performance. So for example, raters can distinguish between behaviours indicating poor leadership from behaviours indicative of good leadership.
Reliability;	This relates to the consistency or stability of the measurement.
	Test-retest – assesses stability over time. Raters would be asked to make the same judgements on two occasions and these would be compared (correlated).
	Internal reliability or consistency – tests the level of inter-correlation between a set of items intended to be measuring the same construct (e.g. the inter-correlation of scores on elements of a skill category called decision-making).
	Inter-rater reliability – measures whether the raters using the system are applying it in the same way and are showing agreement in their ratings.
Validity;	Refers to the extent to which a measure really assesses the construct. The behaviour ratings should accurately reflect real differences in the skills being measured. The skills and behaviours being assessed should also be related to the performance outcome of interest (e.g. safety)
	Face validity – is whether the items look to practitioners as if they are measuring the appropriate construct. This is not a true measure of validity but if face validity is low (i.e. content of scale does not look relevant), then this can influence practitioner acceptance of a measure.
	Construct validity – is whether the rating scale is actually measuring what it claims to measure. This can be assessed in different ways, such as comparing the new test with an established measure of the same construct (convergent validity) or by testing whether scores on the test actually relate to the key outcome measure (criterion validity). So for non-technical skills, this would be whether higher ratings actually relate to better safety and efficiency of practice.
Structure:	Minimal overlap between components (e.g. categories). Transparency: those being rated understand the performance criteria against which they are being rated. The reliability and validity data should be available to show the system properties.
Usability:	The system needs to be usable – i.e. the framework is simple, easy to understand, has domain-appropriate language, is sensitive to rater workload, the target behaviours are easy to observe, and raters can be trained to use it.
	Baselines: for performance criteria are used appropriately for the experience level of rate (i.e. ab initio/trainees vs. more experienced practitioners).

Table 3.5 - Further Principles of an effective Behavioural Marker System (Flin et al, 2008)

3.1.14 Who should be an observer?

Behavioural marker observation systems are designed for trained observers to assess Non-technical skills either during training, or as on-going competency assessment. Klampfer et al (2001) provide a number of guidelines in terms of the training of observers (Table 3.6).

What are prerequisites to be a trainer for a Behavioural Marker course?
• Qualifications required of the persons who will deliver a formal course to train, calibrate and qualify raters (evaluators) using the behavioural marker system. • Commitment to human factors principles • Domain knowledge • Formal training in applicable aspects of human factors or non-technical skills (e.g., Crew Resource Management) • Formal training in the use and limitations of performance rating systems • Formal training in the use of the specific behavioural marker system
What are prerequisites for evaluators using a Behavioural Marker system?
• Entry requirements for personnel who will serve as evaluators: • Commitment to human factors principles • Domain knowledge • Formal training in applicable aspects of Human Factors or non technical skills (e.g., Crew Resource Management)
What are necessary qualifications of evaluators?
• Complete initial training on behavioural marker systems • Formal assessment as competent and calibrated following behavioural marker system-training in classroom • Calibration in operational environment (e.g., training, simulator, work environment) • Periodic re-calibration for continuing use of the behavioural marker system

Table 3.6 - Training guidelines for Behavioural Markers (Klampfer et al, 2001)

3.1.16 Structure of a Behavioural Markers System

Behavioural marker systems are invariably hierarchical in structure, with NTS and behaviours presented in an ordered taxonomy. Using the example of the Non-Technical skills Checklist (NOTECHS), four NTS areas are presented within a hierarchical structure of the competencies and behaviours required for best performance (van Avermaete et al, 1998). NOTECHS has been used as an example as it has been used as the underpinning structure to a number of other behavioural marker systems. The structure of NOTECHS is presented in Figure 3.5, in addition, the diagram annotation highlights the three levels of hierarchy the system provides; *category* of NTS, the skill '*elements'* that underpin this skill, and specific observable positive '*behaviours'*.

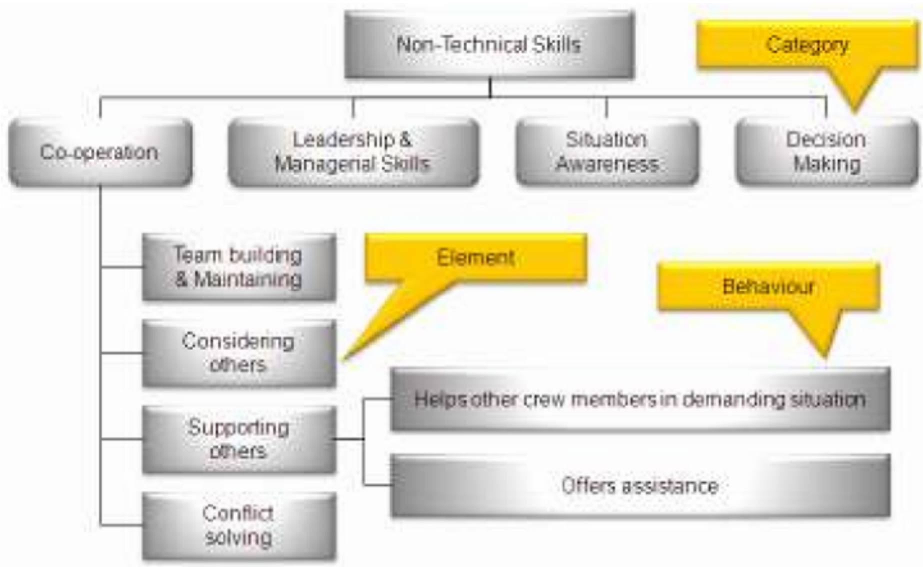


Figure 3.5 - NOTECHS Marker system - hierarchical structure (annotated), van Avermaete et al (1998)

3.1.17 Definition of terms

For clarity purposes and to aid the reader, the following definitions have are used for this thesis:

Behaviour	The overt display of physical action and reaction to undertaking the task, and the general physical and emotional state of the person being observed.
Behavioural Marker	A behaviour which provide insights into aspects of task performance and overall levels of experience in undertaking the task.
Category	A group of behaviours or behavioural markers which share commonality (for example physical body movements and posture). This is a broader use of the term than other behavioural marker systems, where 'category' is used exclusively to denote classes of Non-Technical Skill (NTS) behaviour (for example leadership) (Flin et al, 2008).

3.2 GENERAL CONTENTS OF BEHAVIOURAL MARKERS

Although various marker systems have been developed for different purposes in different domains, there are some broad topics which are common across them (alongside some differences which are highlighted later in this chapter):

- Situation Awareness
- Decision Making & Task management
- Leadership
- Teamwork & Communications

3.2.1 Situation Awareness

Situation Awareness (SA) has a multitude of definitions although the most widely recognised and currently accepted is Endsley's definition of Situation Awareness, which refers to *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* (as cited in Endsley, 1998, p.1).

In terms of ATC, SA translates to an awareness of aircraft positions and flight plans in order to predict future states and resolve possible upcoming conflicts (Jenault et al, 2000).

There are many factors that influence the operator's process of acquiring and then maintaining SA. Cognitive abilities vary between individuals, and this could reduce their ability to acquire SA. This may be the result of lack of skills, experience, and training, (Endsley, 1995).

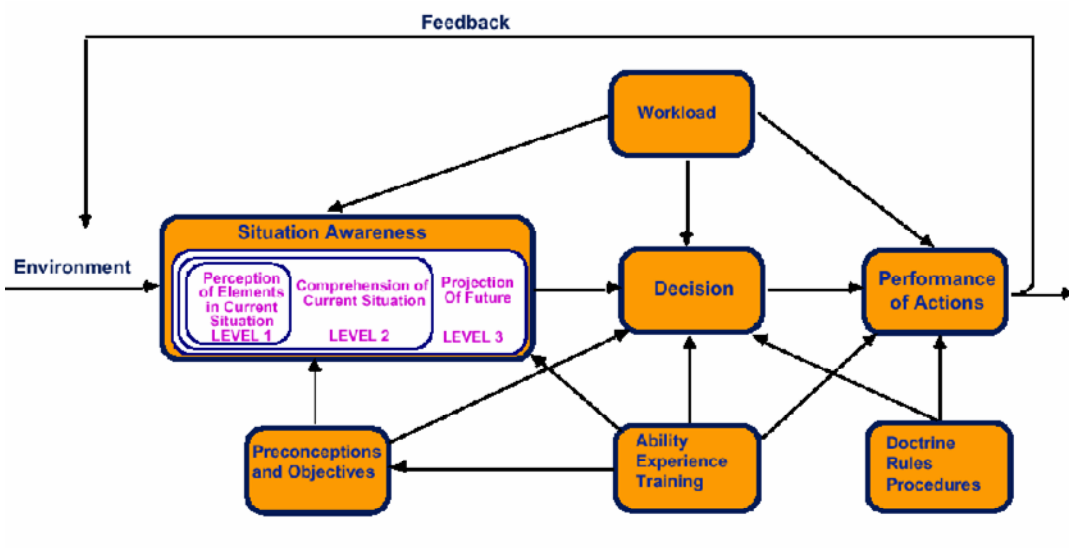


Figure 3.5 - Model of SA in dynamic decision making (Adapted from Endsley, 2000) Permission to reproduce this diagram has been granted by SAGE Publications

Endsley's (2000) model of SA outlines further the internal and external factors that impact upon SA (Figure 3.6). The fundamental concept of this model is that operator perception of all the relevant environmental elements forms the basis of SA. Endsley (2000) argues that SA is a separate preceding action to both decision making and the performance of actions (Figure 3.5).

With regards to behavioural observation, it is the *'performance of the actions'* through which Situation Awareness is evaluated; both in terms of building and maintaining an accurate picture. Overt behaviour within this category concerns looking around an environment at the state of the situation, and probing through actions and questions specific elements for further detail (Flin et al, 2008).

3.2.2 Decision Making

Key to both technical and non-technical performance is decision making, where an incorrect or poor decision can have significant deleterious results, as demonstrated in a number of incidents (Table 3.1). Decision making is extensively an internal cognitive process that incorporates both the gathering and processing of information, the referencing of memory and knowledge, in order to derive a suitable response. Wickens and Holland's (2000) model of decision making succinctly illustrates information processing involved. However it is only the aspects of decision making regarding 'sense making' and 'response execution' which may generate overt observable action (Figure 3.6).

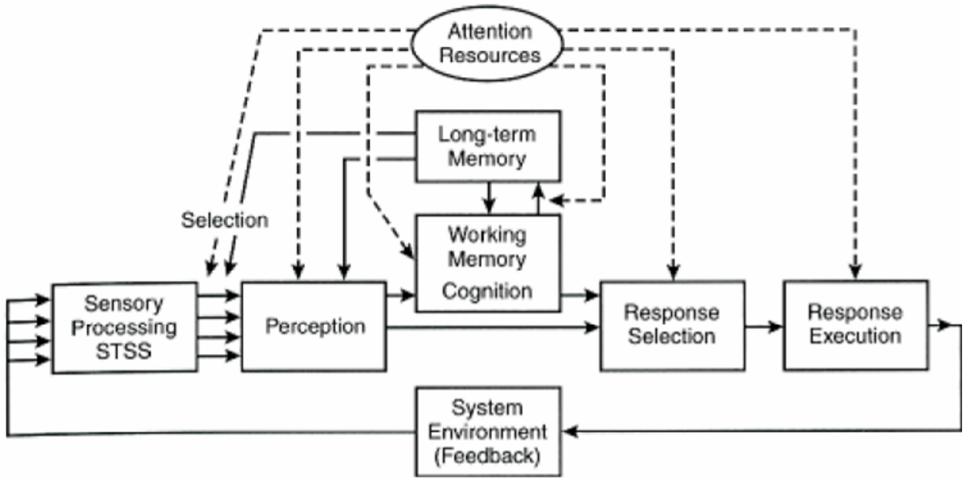


Figure 3.6 - Information Processing Model of Decision Making (Wickens & Holland, 2000)

Decision making may be impacted by a number of external pressures, such as uncertainty, familiarity and expertise, and time. Therefore noting and understanding the context and environment in which a decision has been made may afford greater understanding of the situation.

Response execution forms the basic structure for task delivery, from the product of attention resource management. A number of actions may be made following the result of a decision (including inaction). Figure 3.7 depicts the core tasks an ATCO is required to provide in their role (Eurocontrol, 1996).

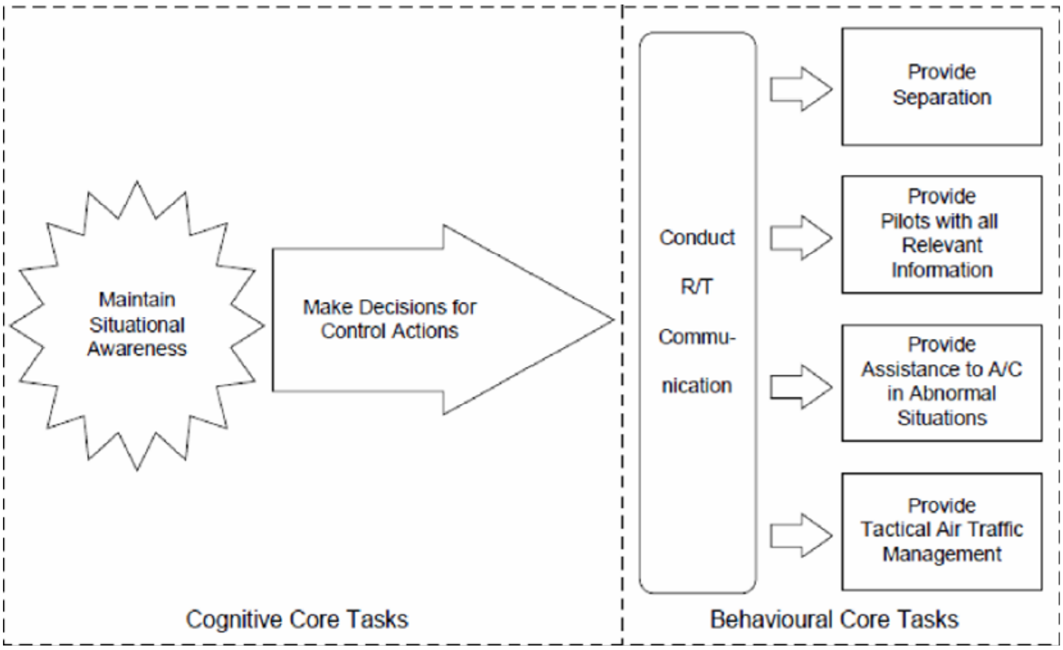


Figure 3.7 - Process model of ATCO core tasks (Eurocontrol, 1996)

Figure 3.8 presents a representative breakdown of the cognitive and task demands placed upon the generic ATCO control positions (Roske-Hofstrand & Murphy, 1998). Figure 3.8 succinctly indicates the differences in skill, and the degree of dynamism required within the various ATC environments.

Non-Radar (Oceanic Control Position)	Radar (En Route Control Position)	Radar (Approach / Departure Control Position)	Non-Radar (Tower Control Position) Good weather/Day Time
No direct representation of the traffic situation	**Symbolic representation of the traffic situation**	**Symbolic representation of the traffic situation**	**Direct perception of the traffic situation**
High memory demand	**Moderate memory demand**	**Little memory demand**	**High memory demand**
High demand for mental simulation	**Moderate demand for mental projection**	**High demand for mental projection**	**Visual estimation**
High strategical planning	**Combination strategic and tactical planning**	**High tactical planning**	**High tactical planning**
Long delays in air-ground communications	**Normally no delays in air-ground communications**	**Normally no delays in ground communications**	**Normally no delays in ground communications**
Slow event development	**Moderately quick event development**	**Quick event development**	**Rapid event development**

Figure 3.8 - ATCO Cognitive task parameters (Adapted from Roske-Hofstrand & Murphy, 1998)

3.2.3 Leadership

Chemers (1997, p.1) defines leaderships *as "a process of social influence in which one person can enlist the aid and support the others in the accomplishment of a common task"*. In order to ensure that decisions made and activities assigned, a leader must be effective in order to ensure these are satisfactorily completed. Unsforth and West (2000) identify several input and process factors which can impact effective teams, key in this process is effective leadership:

- Input factors
  - Task
  - Team Composition
  - Organisational Context
  - Cultural context
- Process factors:
  - Leadership
  - Decision Making
  - Cohesiveness

Reason (2008) demonstrates through the exploration of several dramatic emergency situations in various domains, that effective leadership 4 in addition to knowledge, skills and experience 4 enables a successful safety outcome in the most extreme of circumstances. Leadership therefore can be the key driving force within a team or organisation's approach to safety. Leadership is therefore an important NTS category to consider in any behavioural marker system.

3.2.4 Teamwork & Communications

Salas et al (2008, p.541) defines teamwork as "*a set of interrelated cognitions, attitudes and behaviours contributing to the dynamic processes of performance*". Salas et al (2008) go on to state that a number of factors impact teamwork including the personalities of team members, individual's cognitive abilities, motivators within the team (drivers and goals) in addition to cultural, and organisational factors, and the demands (workload) placed upon the team.

Within the ATC domain, Malakis et al (2010), identify elements which are essential for effective teamwork, namely; Team Orientation , Team coordination Information exchange (Communication), Error management and Task distribution (Change management). These elements form the foundation of the Taskwork and Teamwork strategies in Emergencies in Air traffic Management (T2EAM) model depicted in Figure 3.9.

Team communication may be direct (verbal, or non-verbal) or indirect for example by electronic means. Cushing (1997) identifies 6 core aspects of communications which have resulted in injury or death within Aviation, these aspects concern problems with language; differing frames of reference (distance, height, speed, locations), repetition, radio equipment issues, and compliance. The implementation and adaptation of standardised phraseology is an important mitigation against potential communications errors within ATC (Duke, 1997; CAA, 2012).

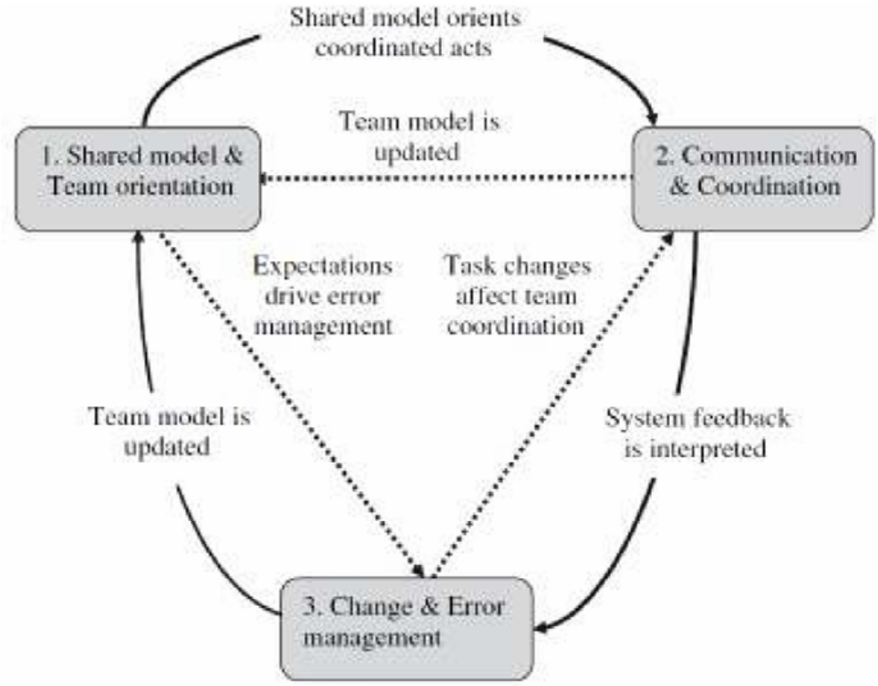


Figure 3.9 - Taskwork and Teamwork strategies in Emergencies in Air traffic Management (T2EAM). (Malakis et al 2010)

3.3 KEY BEHAVIOURAL MARKER SYSTEMS

Over the last 20 years, a number of observational systems have been developed in order to assess NTS competency. Due to the relatively high cost to develop and utilise such systems, these systems are found predominantly within safety critical domains such as the flight deck (Klampfer et al, 2001).

The systems presented in this chapter are arguably the most significant and underpinning in this field of research, although it is recognised that there are other 'in-house' systems unpublished outside of the airline industry (Flin, 2001). These systems have been developed across various industries including the flight deck, space, medicine, maritime, and the railway industry. Key observational research undertaken within ATC is also explored, although this is largely focused upon the competency assessment of task performance. The 19 behavioural observation and marker systems which have been reviewed in this chapter are presented in Table 3.7. Appendix B presents additional content for several of the Behavioural Observation Systems contained in this chapter.

Domain	Developer of Behavioural Observation System / Name of System
Aerospace:	- University of Texas / AC 120-51A
	- Line Oriented Safety Audit (LOSA)
	- Non-TECHnical Skills Checklist (NOTECHS)
	- NASA – Human Behaviour & Performance Competency Model
Medicine:	- Anaesthetists Non-Technical Skills (ANTS)
	- Non-Technical Skills for Surgeons (NOTSS)
	- Non-Technical Skills for Scrub Practitioners (SPLINTS)
Maritime and Rail:	- Maritime Crisis management - Non Technical Skills for officers of the deck (NTSOD) - Railway Safety and Standards Board (RSSB)
ATC:	- EUROCONTROL - Behaviourally Orientated Observation Technique (BOOM) - EUROCONTROL – Behaviour Observation Scale) - Luchtverkeersleiding Nederland – Executive & Planning Controller Observer competency test for ACC - FAA: Separation and Control Hiring Assessment Project (SACHA) Observation Sheet - FAA: Modified SACHA Observation Sheet - FAA: Air Traffic Selection And Training (AT-SAT) Rating Form - NATS: Heathrow Tower Safety Markers - NATS: Day-to-day Safety Observations - NATS: Individual's Positive Team Behaviours

Table 3.7 - Behavioural Marker Systems reviewed in this chapter

Where available, background has been provided to each key NTS behavioural marker system including by whom and for what purpose they have been developed. However it is important to state that by the nature of these in-house systems, there is often limited detail on their development, and limited detail on their subsequent testing and evaluation.

3.4 KEY BEHAVIOURAL OBSERVATION WORK IN AEROSPACE

3.4.1 University of Texas (UT) Behavioural Markers

The earliest documented behavioural marker system to be developed for the assessment of CRM was the University of Texas (UT) Behavioural Markers System (Helmreich & Wilhelm, 1987). This US government funded work, initially titled the 'NASA /UT Project' had two clear goals, the first was to evaluate the effectiveness of CRM training as measured by observable behaviours, while the second was to aid in defining the scope of CRM programmes.

Under this project, an observational checklist was developed. Very little information has been published regarding the development of this tool however it has laid the foundation for other tools produced in subsequent research. The UT checklist contains a number of behavioural 'anchors' used by the observer to rate performance. Details of how these were derived is unavailable, however the influence of in-house CRM checklists from aviation companies (America

West Airlines, Continental, American Airlines) is one of the sources reviewed (Helmreich & Foushee, 2010).

The UT Observation sheet requires the observer to make ratings (using a 4 point rating scale of crew performance) against 13 CRM skill areas (Table 3.8). Following successful early trials the design was partitioned, so that the ratings are made across the various phases of flight. P = Pre-departure/Taxi, T = Takeoff/Climb, D = Decent/Approach/Land, G = Global. Of significant interest are the behavioural anchors presented in Table 3.7 which provide an early illustration of specific observable behaviours which may provide insights into superior or inferior NTS performance.

Markers	Definition	Anchors (examples)	Phase
SOP Briefing	The required briefing was interactive and operationally thorough	Concise, not rushed, and met SOP requirements. Bottom lines were established	P-D
Plans Stated	Operational plans and decisions were communicated and acknowledged	Shared understanding about plans – 'Everybody on the same page'	P-D
Workload Assignment	Roles and responsibilities were defined for normal and non-normal situations	Workload assignments were communicated and acknowledged	P-D
Contingency Management	Crew members developed effective strategies to manage threats to safety	Threats and their consequences were anticipated. Used all available resources to manage threats	P-D
Monitor / Crosscheck	Crew members actively monitored and cross-checked systems and other crew members	Aircraft position, settings and crew actions were verified	P-T-D
Workload Management	Operational tasks were prioritised and properly managed to handle primary flight duties	Avoided task fixation. Did not allow work overload	P-T-D
Vigilance	Crew members remained alert of the environment and position of the aircraft	Crew members maintained situational awareness	P-T-D
Automation Management	Automation was properly managed to balance situational awareness and/or workload requirements	Automation set-up was briefed to other members. Effective recovery techniques from automation anomalies	P-T-D
Evaluation of plans	Existing plans were reviewed and modified when necessary.	Crew decisions and actions were openly analysed to make sure the existing plan was the best plan	P-T
Inquiry	Crew members asked questions to investigate and/or clarify current plans of action	Crew members not afraid to express a lack of knowledge. 'Nothing taken for granted' attitude	P-T
Assertiveness	Crew members stated critical information and/or solutions with appropriate persistence	Crew members spoke up without hesitation	P-T
Communication Environment	Environment for open communication was established and maintained	Good cross-talk – flow of information was clear and direct	G
Leadership	Captain showed leadership and co-ordinated flight deck activities	In command, decisive and encouraged crew participation	G

1 = Poor	2 = Marginal	3 = Good	4 = Outstanding
Observed performance had safety implications	Observed performance was barely adequate	Observed performance was effective	Observed performance was truly noteworthy

Table 3.8 - University of Texas (UT) Behavioural Markers Scale (Klampfer et al, 2001)

As part of the FAA's continued support to the area of CRM, a series of Advisory Circulars (AC 120-51) have been produced over a number of years under the title "Crew Resource Management Training". The earliest version AC 120-51A (FAA, 1993) contains an expanded and quite extensive set of behavioural markers against the 13 CRM topics found in the UT behavioural marker scheme. The full set of behavioural markers presented in AC 120-51E (which is the most recent and up-to-date version of this Advisory Circular) is presented in Appendix B01 (FAA, 2004). An example taken from this set, '*Communications/Decisions'*, is presented in Table 3.9. The behavioural markers contained within AC 120-51E are all examples of desired behaviour, and provide additional detail upon which observers may ground their judgements of NTS performance. AC 120-51 is the first published example of an extensive set of behavioural markers for use in the assessment of NTS competency.

1. COMMUNICATIONS PROCESSES AND DECISION BEHAVIOUR CLUSTER.	
Communications/Decisions. These behaviours relate to free and open communication. They reflect the extent to which crewmembers provide necessary information at the appropriate time (e.g., initiating checklists and alerting others to developing problems). Active participation in the decision making process is encouraged. Decisions are clearly communicated and acknowledged. Questioning of actions and decisions is considered routine.	(1) Operational decisions are clearly stated to other crewmembers. (2) Crewmembers acknowledge their understanding of decisions. (3) "Bottom lines" for safety are established and communicated. (4) The "big picture" and the game plan are shared within the team, including flight attendants and others as appropriate. (5) Crewmembers are encouraged to state their own ideas, opinions, and recommendations. (6) Efforts are made to provide an atmosphere that invites open and free communications. (7) Initial entries and changed entries to automated systems are verbalized and acknowledged.

Table 3.9 - Communications Behavioural Markers: Advisory Circular 120&51E (FAA, 2004)

3.4.2 Line Oriented Safety Audit (LOSA)

The University of Texas, under a follow up FAA research project continued to explore the domain of flight Deck CRM. Through a number of practical applications of the original behavioural markers system, the need was identified for a system which focuses specifically upon Threat and Error Management (TEM). As a result, a methodology we developed to specifically evaluate TEM. This system is called the Line Operations Safety Audit (LOSA).

The method, which takes many design principles from the UT observation sheet, involves trained observers being present within the cockpit, in order to evaluate several aspects of crew performance (Klinect et al, 2003). The categories and elements covered under LOSA assessment are presented in Table 3.10. LOSA observers record the following:

- *The various threats encountered by aircrew*
- *The types of errors committed, and most importantly,*
- *They record how flight crews manage these situations to maintain safety.*

In addition to the observational data collected, a LOSA trained observer also undertakes a structured interview in order to ask pilots for any safety improvement suggestions. The combination of direct observation and interview provide airlines with a diagnostic snapshot of safety strengths and weaknesses in normal flight operations.

LOSA is the recommended safety assessment process by the International Civil Aviation Organisation (ICAO, 2002) and by the FAA under Safety Program Advisory 120-90 (FAA, 2006).

Category	Elements
Planning	Briefing
	Contingency Management
	Workload Assignment
	Plans Stated
Execution	Monitor/Cross-check
	Workload Management
	Vigilance
	Automation Management
Review/Modify Plans	Evaluation of plans
	Inquiry
	Assertiveness
Overall Markers	Communication environment
	Leadership
	Flight Attendant briefing on first leg
	Captain Contribution to crew effectiveness
	First officer contribution to crew effectiveness
	Overall Crew effectiveness

Table 3.10 - Line Oriented Safety Audit, (ICAO, 2002)

LOSA represents the first published example of a NTS observation system developed specifically in order to evaluate threat management, instead of a broader set of NTS competencies. LOSA has recently been adapted for the ATC environment. The Normal Operations Safety Survey (NOSS) enables trained observers to capture data during normal 'safe' operations in order to evaluate levels of 'Threats, Errors, and Undesired States' within an ATC organisation, and how they are managed routinely during normal operations (Barbarino & Patterson, 2007).

3.4.3 Non-Technical skills Checklist (NOTECHS)

The Non-Technical skills Checklist (NOTECHS) was developed by the Non-Technical Skills Project, under the European Joint Aviation Authorities (JAA) Project Advisory Group on Human Factors. The project group consisted of a researchers from four institutions; DLR, NLR, IMASSA, and Aberdeen University. The purpose of the project was to develop an assessment method for flight crews CRM skills during training and assessment. Hörmann and Neb (2004) state the key objectives of the NOTECHS checklist are:

- *To assess the skills of an individual pilot, rather than a crew*
- *For the assessments [to be] based on observable behaviours*
- *Behaviours are precisely defined, culturally robust, reliable, and practicable*

Flin et al (2003) describe an iterative process of development for NOTECHS, which began with a literature review of existing behavioural observation systems used to evaluate pilot's non-technical skills proficiency (as used by KLM, Air France, and Lufthansa, as well as The University of Texas NTS system). This review was under taken in order to i) identify common categories and elements of behaviour and ii) obtain relevant findings relating to key categories of non- technical skills identified in existing systems.

Next a series of ongoing discussions and workshops was undertaken with Subject Matter Experts (SMEs) from KLM who had prior experience in evaluating NTS. A number of design principles were used in the process of developing the NOTECHS system (van Avermaete & Kruijsen, 1998; Flin et al., 2003).

- It should be suitable for use across Europe by both large and small operators
- It should use a two-point rating scale to assess the level of CRM skill proficiency (acceptable or unacceptable).
- It should contain the minimum number of categories and elements in order to cover critical behaviours whilst maintaining the maximum possible mutual exclusivity.
- The terminology used should avoid psychological jargon and reflect everyday language for behaviour,
- If a crew member fails the observation, this should indicate that negative consequences would likely result in the future. i.e. a poor CRM skill that was deemed to have had a direct impact on flight safety.
- An explanation must be provided for any CRM skill deemed to be deficient
- Repetition of the observed skill deficiency is required in order to establish repeated poor performance.

- Only observable behaviour shall be assessed with limited interpretation of cognitive processes (pilot personality and emotional attitude were excluded), which is a design principle advocated by Klampfer et al (2001).

A prototype system was developed from non-technical skill sets identified in the literature, and through a series of meetings and subsequent workshops a draft taxonomy was produced and reviewed by the project team, KLM SMEs and other stakeholders. Revisions to the components and structure were made in a final meeting, addressing any comments arising from the review.

The developed marker system contains two cognitive skills (decision making, situation awareness) and two social skills (co-operation, leadership & managerial); with a number of skill elements (Table 3.11). A five-point rating scale was used to rate each element of NTS performance (‘very good, good, acceptable, poor, very poor’) in addition to an overall rating of ‘acceptable/unacceptable’ performance.

A number of observable markers of both bad and good behaviours were also identified for each skill element and are presented in Appendix B02. This is significant, as this is the first published example where behavioural markers of both good and bad NTS have been included (all previous examples having contained examples of only good behaviours). This precedent has impacted the design of several subsequent behavioural observation systems. The NOTECHS system has been used as the structural design for a number of other behavioural marker observation systems, most notably those developed for the medical domain by the Industrial Psychology Research Centre, Aberdeen University (ANTS, NOTSS, SPLINTS presented later in this chapter).

Category	Element
Co-operation	Team building and maintaining
	Consideration of others
	Support of others
	Conflict Solving
Situation Awareness	Awareness of aircraft systems
	Awareness of external environment
	Awareness of time
Leadership & Managerial Skills	Use of authority and assertiveness
	Providing and maintaining standards
	Planning and co-ordination
	Workload Management
Decision Making	Problem definition and diagnosis
	Option generation
	Risk assessment and option selection
	Outcome review

Table 3.11 - NOTECHS Categories and Elements, van Avermaete et al (1998)

The NOTECHS observational technique was tested using 105 instructor pilots across Europe. These instructors were tasked with assessing the CRM skills displayed by individual flight deck team members using pre-recorded video scripted scenarios acting out a variety of alternative safety and normal situations (O'Connor et al, 2002). Internal consistency, accuracy, inter-rater agreement, and user acceptance were evaluated.

3.4.4 NASA – International Space Station Human Behaviour & Performance Competency Model

A collaboration of researchers from NASA, the Canadian, European, and Japanese Space agencies, and the Russian Gagarin Cosmonaut Training Centre have developed a series of Human Behaviour Performance (HBP) competencies for long duration international missions. These were developed in response to the Multilateral Crew Operations Panel’s (MMOP) request for Human Behaviour Performance requirements for astronauts aboard the International Space Station (ISS), (Bessone et al, 2008b).

The competency model builds upon the fundamental concept that *"observing astronaut behaviour in simulation and comparing it to observed performance in space would provide feedback to training personnel about the effectiveness of the simulation, as well as feedback to psychologists about the validity of on-orbit observational techniques" (*Musson, 2000, p.152).

No details are available as to how the competency model was developed, or whether an observational tool has been produced as a result. The collaborative nature of the project suggests a workshop and Subject Matter Expert (SME) type arrangement as per the JARTEL group who produced NOTECHS. This work has been included within this chapter principally for two reasons. The first is that it represents the only published behavioural marker set within the domain of space science. The second is that the coverage of the marker set is somewhat broader than the CRM type skills contained within other systems. Given space missions are much longer in duration than that of an aircraft flight, the observation of team dynamics must also consider the longer time frame and incorporate aspects such as group living and cultural considerations.

Table 3.12 provides the set of competencies produced by the project team, for use in the assessment of the training curriculum's delivery for long duration crewmembers. Appendix B03 presents a further set of identified positive Behavioural markers with regards to the NTS category and elements identified.

Category	Competency (Element)
Self-Care, Self-Management	Refine accuracy of self image
	Manage stress
	Care for oneself
	Maintain efficiency
Communication	Optimize communication
	Ensure Understanding
Cross Cultural	Demonstrate respect towards other cultures [national, organisational, professional]
	Understand culture and cultural differences [national, organizational and professional]
	Build and maintain social and working relationships
	Intercultural communication and language skills
Teamwork & Group Living	Active team participation
	Interpersonal relationships
	Group living
Leadership	Execution of designated leader's authority
	Mentoring skills
	Followership
	Workload Management
Conflict Management	Conflict prevention
	Conflict resolution
Situational Awareness	Maintenance of an accurate perception of the situation
	Processing of information
Decision Making & Problem Solving	Problem solving and decision making methods
	Preparation of decision
	Execution of decision

Table 3.12 - NASA Behaviour Competencies Long-Duration Missions, (Bessone et al, 2008a)

3.5 KEY BEHAVIOURAL OBSERVATION WORK IN MEDICINE

3.5.1 Anaesthetists Non-Technical Skills (ANTS) Checklist

Developed by Aberdeen University's Industrial Psychology Research Centre, the Anaesthetists Non-Technical Skills (ANTS) checklist follows a similar style and construction of format to the NOTECHS system (Fletcher et al, 2003). With Aberdeen University a key contributor in the JARTEL project (who developed

NOTECHS), the ANTS technique builds on this pillar of research knowledge, adapting and expanding into the medical domain (Flin, 2004).

The ANTS checklist was developed through review of literature and underpinning material gathered through a survey of anaesthetists' attitudes to safety and teamwork (Flin & Fletcher et al, 2003). In addition a semi-structured interview of experts was undertaken. This multi-part interview first involved the interviewee recounting their experience managing a difficult incident (as per the critical incident technique), with supporting probing questions from the interviewer, (Flanagan, 1954). The second part of the interview involved a more general discussion with the interviewee on the NTS they felt important for good practice in anaesthesia. During the last part of the interview the interviewees were asked to sort a number of cards into groups of related items (these items having been identified as common across other marker systems).

A final aspect of knowledge elicitation involved a review of available anaesthetists incident reports, however this review was non beneficial to the marker system's development due to a lack of fidelity to incident data, particularly in the coding of human factors information (Fletcher et al, 2004).

The categories and elements identified and assessed through the development of ANTS are presented in Table 3.13. These are taken from version 1 of the ANTS Handbook, available from http://www.abdn.ac.uk/iprc/ants/. A large set of example positive and negative markers, from which the observers derive their NTS competency ratings, are presented in Appendix B04.

Category	Element
Task Management	Planning and preparation
	Prioritisation
	Providing and maintaining standards
	Identifying and utilising resources
Team Working	Co-ordinates activities with team members
	Information exchange
	Use of authority and assertiveness
	Assessment of capabilities of team and self
	Supporting others
Situation Awareness	Gathering information
	Understanding and recognition
	Anticipation
Decision Making	Identifying
	Balancing risks and selecting options
	Re-evaluation

Table 3.13 - ANTS Categories and Elements, ANTS v.1 (Fletcher, 2006)

The utility, reliability, and validity of the ANTS system was evaluated by 50 consultant anaesthetists, who viewed the video footage of surgical team undertaking scripted anaesthetic scenarios. Performance ratings (using an anchored 4-point scale) were made against the categories and elements contained within ANTS. Inter-rater reliability, accuracy and external consistency all received favourable scores (Mean inter-rater agreement between 0.5 – 07). With regards to usability, the observation sheet was deemed as both acceptable and well designed (Fletcher et al, 2003; Fletcher 2006). Minor amendments were made as a result of this study; this principle of 'iterative development' is reflective of other behavioural marker systems contained in this review (e.g. NOTSS, SPLINTS, NOTECHS).

3.5.2 Non-Technical Skills for Surgeons (NOTSS)

The second Non-Technical Skills observational checklist developed by Aberdeen University's Industrial Psychology Research Centre, is the Non-Technical Skills for Surgeons (NOTSS). The NOTSS checklist follows a similar format to the ANTS and NOTCHS checklists. The system focuses upon the cognitive and interpersonal skills that complement surgeons' technical skills.

This work initiated through a literature review examining previous research into non-technical skills assessment in surgery (and other domains), including analysis of observational studies, surgical adverse events analysis, and surgical training and competence assessment (Yule et al, 2006a). A series of structured interviews were then undertaken with 27 surgeons from 11 hospitals using the Critical Incident Technique. The interviews were regressive, involving repeated exploration of the incident using probes and cognitive cues designed to elicit deeper-held tacit knowledge (Yule et al, 2008).

Four categories of NTS were identified and are contained within the NOTSS checklist, and 12 skill elements (Table 3.14). As per the ANTS system, *Situation Awareness* and *Decision Making* are categories; *Teamwork* is also a category although it has been classified as *Communications and teamwork*. A new category, *Leadership* is contained within the NOTSS system, which partially replicates the skill elements contained in the ANTS Task Management Category (i.e. 'providing and maintaining standards', and 'supporting others'). Appendix B05 contains a set of example positive and negative markers, from which the observers derive their NTS competency ratings.

Category	Element
Situation Awareness	Gathering information
	Understanding and recognition
	Projecting and anticipating future state
Decision Making	Considering options
	Selecting and communicating options
	Implementing and reviewing decisions
Leadership	Setting and maintaining standards
	Supporting others
	Coping with pressure
Communications & Teamwork	Exchanging Information
	Establishing a shared understanding
	Co-ordinating team

Table 3.14 - Categories and Elements, NOTSS Handbook Version 1.2 (Yule et al, 2008)

A study was undertaken with 44 surgeons participating from five Scottish hospitals. Following training in the NOTSS system, they were asked to rate the behaviours of consultant surgeons within a simulated operating room, enacting various positive and negative NTS behaviour. Ratings, as per the ANTS system, were made using a 4-point performance anchored rating scale (where behaviours which potentially endanger patient were rated as 1 (poor)).

The results show that inter-rater reliability was generally acceptable (IRR (>0.6) for all categories except Task management, with strong Intra-class correlation coefficients (0.95 - 0.99). Yule et al (2008) go on to suggest that further training would likely improve the accuracy of the system across observers, and that testing in the operational environment is warranted. Once again this finding indicates that in order to develop a robust system it must be tested on several occasions, with a variety of well-trained observers; and reviewed and amended as appropriate.

3.5.3 Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS)

The Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS) is the latest Non-Technical Skills observational checklist developed by Aberdeen University's Industrial Psychology Research Centre, for use in the medical domain (Mitchell et al, 2011). It builds upon the principles and knowledge obtained through the ANTS and NOTSS systems, and follows a similar format. The system focuses upon the cognitive and interpersonal skills that complement surgeons' technical skills.

The SPLINTS system has been developed in order to support the training of junior team members in the scrub role, by senior perioperative practitioners.

Assessments are made either as is the case for ANTS and SPLINTS through peer rating or it may be used as the basis for self-assessment (Mitchell et al, 2011).

Focus groups with expert practitioners were held, in order to develop a preliminary taxonomy from a review of NTS related behaviours identified from interviews. Skill categories and elements were labelled, in addition, examples of good and poor behaviours were recorded. Through an iterative process the preliminary taxonomy of non-technical skills and 28 underlying elements was reduced to three categories (situation awareness, communication and teamwork, task management), each with three underlying elements (Table 3.15). This constitutes version 1 of the SPLINTS Handbook (available from http://www.abdn.ac.uk/iprc/splints/). A number of positive and negative markers associated with the NTS are contained within Appendix B06.

Category	Elements
Situation Awareness	Gathering information
	Recognising and understanding information
	Anticipating
Communication and Teamwork	Acting assertively
	Exchanging information
	Co-ordinating with others
Task Management	Planning and preparing
	Providing and maintaining standards
	Coping with pressure

Table 3.15 - Categories and Elements, SPLINTS handbook – Volume 1

The SPLINTS system was evaluated through expert practical application of the observational system. 34 Experienced scrub practitioners received one days training in the design and application of the system. They were then asked to use SPLINTS to rate the scrub practitioners' NTS in the video recordings of seven standardized simulated, surgical scenarios. A 4-point performance anchored rating scale was employed (1-poor, 2-marginal, 3-acceptable and 4-good, and NR-Not Required). The recordings (approximately 2 – 4 minutes duration) were developed through the input of a SME steering group, and were considered to contain a reflective number of the routine and non-routine 'surgical events' a scrub practitioner is likely to face, Mitchell et al (2012a).

A two-part questionnaire was used to 1) gather background information on observers experience of training junior members of perioperative staff, their knowledge of NTS, and indication if they had previously been involved in the development of SPLINTS, along with basic demographic data (sex, years of experience); 2) gain feedback on the design of the system. Questions covered the completeness, observability, usability, the design of the rating scale, the training, and the design of the scenarios, (Mitchell et al, 2012a).

Inter Rater Reliability was generally favourable (within-group agreement: r_{wg} for the three skill categories and six out of nine elements, > 0.7). In addition, ratings were within one scale point for over 90% of skill categories and elements. These findings are consistent with the qualitative questionnaire feedback (Mitchell et al, 2012b).

Flin & Patney (2011) indicate that several further works are underway with the medical domain in order to develop NTS tools to support anaesthetic assistants, obstetric anaesthetists, and histopathologists.

3.6 KEY BEHAVIOURAL OBSERVATION WORK IN MARITIME AND RAIL

3.6.1 Non-Technical Skills for Navy Crisis Management

A number of CRM crisis management skills have been identified by the Warash Maritime Centre. Observations within training scenarios identified certain patterns of interaction between engine room team leaders and other team members. An analysis of this data was undertaken to identify particular patterns and behaviours that lead to successful management of crises (Table 3.16), Gatfield (2005).

Category	Behavioural Marker
Situation Awareness	Ratio of the degree of feedback control to the degree of predictive control.
Teamwork & Shared Mental Models	The number of alternative hypothesis and actions communicated to team members.
Decision Making	Considering only as many alternatives as needed to discover one that satisfied – level of satisfaction exhibited.
Communication & Shared Mental Models	Building, maintaining and refining the accuracy of the shared mental model of the team
Shared Mental Models	Relevance and timeliness of unsolicited information passed between team members.
Situation Awareness	Level of anticipation of other team members
Situation Awareness	Level of anticipation of future action and task requirements
Too focused on reducing uncertainty	Indication of a tendency towards analytical decision-making, and away from naturalistic decision-making
Lack of situation overview [attention tunnelling]	Tendency to focus on one system at a time, thereby ignoring the dynamics of the complete system
Situation Awareness & Mental Models	Amount of sampling behaviour exhibited
Uncertainty	Number of unfinished sentences
Workload Management	Delegation of work tasks
Situation Awareness	Patterns of movement [around the control room?]

Table 3.16 . Naval Crisis Management Observable Markers (Adapted from Gatfield, 2005)

Gatfield (2008) explores these preliminary research findings in greater detail, in order to identify a robust set of behavioural markers for use in the assessment of maritime crisis management. Gatfield (2008, p.105) states the following "*methodological attributes*[0]" required for the system:

- *"Unobtrusiveness — so as not to adversely affect the natural behaviour ofthe participants*
- *Repeatability — so that data from a number of experiments could be collected in the same way*
- *Holistic — so that all relevant behaviours were recorded*
- *Permanence — so that the data could be analysed as often as required*
- *Contextual - so that the data reflect the setting in which the behaviourstook place, as well as the behaviours themselves*
- *Accuracy — so that the data are a precise record of both the audio andvisual elements of the observed behaviour*
- *Speed —so that none of the observable behaviour is missed*
- *Objectivity — so that the data captured is an objective record of the behaviours observed".*

Through a review of the literature, and subsequent identification of suitable NTS markers (through observations made by trained CRM observers within scripted crisis management simulations), a number of markers considered important were identified. A further offline review of the video recordings taken during the simulation exercises was also undertaken in order to ensure no additional markers had been missed. A review of the data collected was undertaken in order to determine the strongest and most suitable markers for use in evaluating crisis management; four metrics were considered (ease of observation, ease of evaluation, frequency of occurrence, relevance to maritime crisis management). This review was achieved through two SMEs reviewing one of the recordings made, and rating each of the markers.

A final study was undertaken with trained observers for the purpose of evaluating the down selected observational markers set. Later reliability and the overall observability of the markers were tested. For the purposes of this study, a single amalgamated video was created from the original footage, which contained as many examples of the various CRM behaviours captured as possible. Three groups of assessors (6, 6, 7) used the behavioural markers framework and a 4-level rating scale (Good, Towards Good, Towards Poor, Poor) to rate the performances demonstrated in the video. No training was provided to these observers, only the framework and a list of positive and negative markers associated with each behaviour (Appendix B07). Gatfield, (2008) states that the results indicated many behaviours which demonstrated moderate to strong levels of inter-rater agreement (Mean $_{rwg}$ for 13 out of 19 markers < 0.5), the observability of behaviours also scored extremely highly (Mean percentage of total observed responses was 95.7% for all 19 markers).

3.6.2 Non Technical Skills for officers of the deck (NTSOD) - US Navy

A prototype NTS observation system has been developed by Long (2010) for the purpose of assessing NTS for the officer of the deck (OOD) of a US Navy ship (NTSOD). The OOD is responsible for the ship's safe operations and is accountable to the commanding officer for any event which occurs during their watch period (O'Connor & Long, 2011). The research started with a review of the NTS CRM literature, Table 3.17 presents 17 NTS categories which were identified and considered of importance to the OOD role (Long, 2010).

Non-Technical Skills Identified		
Situational awareness	Coping with fatigue	Energy
Decision making	Forehandedness	Co-operation
Communication	Vigilance	Management skills
Team working	Judgment	Task management
Leadership	Intuition/experience	Workload management
Managing stress		Assertiveness

Table 3.17 - Non-Technical Skills relevant to Officers Of the Deck (Long, 2010)

A focus group of four qualified OODs used the skills identified from the literature review in order to develop an initial taxonomy containing 5 NTS categories (Leadership, Decision Making, Situational Awareness, Communication, Managing Stress), and 14 corresponding behavioural elements. The next phase of research concerned the undertaking of 16 separate interviews with OOD in order to identify specific behavioural markers of relevance to the OOD role. In total, 149 behavioural marker statements were captured through interview.

The next step in the development of NTSOD was for two SMES to separately classify the 149 behavioural marker statements under the five NTS categories and 14 elements. The purpose of this study was to confirm levels of inter-rater agreement converting the categorisation markers under the preliminary framework. This exercise was repeated on two further occasions as an iterative process, consolidating and adjusting the frame work to suit the qualitative and quantitative results. As a result of this process, the framework was reduced from five to four skill categories, covering 10 skill elements (Table 3.18), with a Cohen's kappa of 0.91 resulting from the final classification of the remaining markers (O'Connor & Long, 2011).

The NTSOD system incorporates a performance rating scale of four levels (Unsatisfactory, Marginal, Satisfactory, Outstanding, and Not Observed). Appendix B08 presents a number of positive behavioural markers associated with the NTS contained within the NTSOD system (Long, 2010).

Category	Element
Leadership	Managing Watch Team
	Coping with Stress
Communications	Providing Information
	Issuing Orders
Situational Awareness	Gathering Awareness
	Understanding Awareness
	Anticipating Future Events
Decision Making	Analytical Decision Making
	Following Orders & Procedures
	Intuitive Decision Making

Table 3.18 - NTSOD, Final Framework (Long, 2010)

3.6.3 Railway Safety and Standards Board (RSSB) NTS for the Rail Industry

As part of the Railway Safety and Standards Board (RSSB) T869 research project, commissioned by the Rail Industry Skills Forum, a framework has been developed in order to evaluate NTS during a rail industry training course. The focus of the research was upon the train driver, although recognition is made that the content could be adapted to suit other roles, (Bonsall-Clarke, 2012).

Through a series of six task analysis workshops, held with trainers and standards managers, a draft generic task analysis to break down the role of the driver into goals, tasks and sub-tasks was generated. Using this framework, the workshop attendees were asked to consider what skills and knowledge the driver would need for each of the goals, tasks and subtasks, a draft list of NTS was used as a prompt in order to help elicit information from the participants.

An iterative process of review of the NTS marker set was undertaken during these workshops, in order to remove overlap and duplication. A further workshop (with SMEs from Northern Rail) was undertaken where small groups of experts were asked to review a single skill category and associated markers. Each of these groups was also asked to list examples of good and bad behaviour for each skill area. Workshop participants were asked for their comments regarding the observability of the existing and new behavioural markers, and to provide suggestions for improvement where warranted (Bonsall, 2012). The final framework (Table 3.19) covers 7 NTS areas and 26 skill elements. Positive and negative behaviours associated with the NTS identified are presented in Appendix B09.

Category	Element
Situational awareness	Attention to detail
	Overall awareness
	Maintain concentration
	Retain information (during shift)
	Anticipation of risk
Conscientiousness	Systematic and thorough approach
	Checking
	Positive attitude towards rules and procedures
Communication	Listening to people (not stimuli)
	Clarity
	Assertiveness
	Sharing information
Decision making and action	Effective decisions
	Timely decisions
	Diagnosing and solving problems
Cooperation and working with others	Considering others' needs
	Supporting others
	Treating others with respect
	Dealing with conflict/aggressive behaviour
Workload management	Multi-tasking and selective attention
	Prioritising
	Calm under pressure
Self-management	Motivation
	Confidence and initiative
	Maintain and develop skills and knowledge
	Prepared and organised

Table 3.19 - Non Technical Skills for the Rail Industry, (Adapted from Bonsall, 2012)

3.7 KEY BEHAVIOURAL OBSERVATION WORK IN ATC

As has been illustrated, the technique of NTS assessment through the employment of Behavioural markers has been undertaken in a number of high critical safety domains; however their published use within the domain of ATC is limited. However, Non-Technical CRM type skills are recognised as important factors within Air Traffic Management (ATM) Safety. Within the ATC domain, the concept is referred to as Team Resource Management (TRM) (Woldring et al, 2005). The behavioural markers technique is still in a state of transference into the ATC domain, with limited effort undertaken to develop this technique to-date; although the use of observation is used extensively as a tool to assess task performance and technical competency.

Contained in the next section are a number of specific observational techniques which have been developed and employed within the ATC domain. The purpose of these observations differs from the assessment of task performance and competency. There are however many principles of value which may be learned from each system, and it is on this basis that they have been included in this chapter.

3.7.1 EUROCONTROL - Behaviourally Orientated Observation Technique (BOOM)

EUROCONTOL, with inputs from several Air Navigation Service Providers (ANSPs), has developed a Behavioural Observation technique (BOOM) for use in simulations and live operations, BOOM is a structured procedure and tool designed in order to train facilitators, simulation instructors and OJTIs with the skills needed in order to undertake NTS observation in ATC (Woldring and Patterson, 2003).

Unlike a structured technique such as NOTECHS, the BOOM technique is less structured involving the observers making semi-structured observations of actions and events; whilst recording detail on additional Performance Shaping Factors such as workload. No set of behavioural markers or indeed NTS competencies is provided, instead the onus is on the observer to identify and assess these using their own internal matrix. Woldring et al. (2005, p.240) state that "*early attempts at developing the BOOM method using TRM facilitators and controllers to generate a list of desirable non-technical skills (embedded in the TRM framework) and the behavioural descriptions associated with these were unsuccessful. Such prototype lists were shown to be not comprehensive to be useful in all contexts and practical experience showed that the experts could not agree on the items that should be included, or the interpretation of these"*. A reproduction of the BOOM observation sheet is in Appendix B10, or available on request from Eurocontrol (http://www.Eurocontrolint/articles/behavioural-oriented-observation-method/).

The observation that such prototypes were non-comprehensive or not useful in all contexts is an important statement, as Klampfer et al (2001, p.10) state in their summation of what makes a good behavioural marker, "*It does not have to be present in all situations"*. This is consistent with Flin's (2010, p197) remark that NTS marker sets should not just be developed simply for crisis scenarios, *"but also routine aspects of safety-critical jobs"*.

With regards to the concept of BOOM, it is through prototype testing, that a prototype list of behaviours be updated and expanded, adopting an iterative approach advocated by Fin et al, (2008). The BOOM technique is mentioned as part of the *'Summary of 7 years of TRM support to Eurocontrol'*, (Woldring et al, 2005). In this summary, which does not provide a great deal of detail on the development or application of the technique, the reader is given the impression that there was mixed opinion as to the content and design of the NTS assessment system; however detail regarding this mixed opinion is not provided.

The process of using the BOOM technique is as follows:

1. A trained observer makes an observation either in live operations or simulation environment – noting any significant TRM behaviour, and their surrounding context.

2. The observer will then generate a series of questions for use during a debrief session. These questions are designed to help interpret the behaviours and underlying NTS (Table 3.20).

3. The observer will undertake a debrief session with the person observed, using the list of behaviours observed, and focusing the discussion on how the person observed might improving team performance through changes in their behaviour.

The BOOM technique differs significantly from other behavioural marker systems. Instead of a comprehensive set of markers for the observer to base their judgement upon, the observer is free to note down any point they consider significant. This is a highly flexible approach, although it is open to the potential for considerable bias (e.g. recollection bias, halo bias),and the debrief interview provided an opportunity to ground the observations made before final judgement is provided. BOOM is one of a handful of publicly available behavioural observation techniques within the ATC domain, notwithstanding the potential weaknesses highlighted it is useful to include within this review.

I saw/heard the following behaviour (observation)	I think (Impression, interpretation of specific or general behaviours, implied NTS)	Generally the observed person (observed several times, or explained as such by the observed controller)	Related TRM domains (depending on the TRM characteristic of the member state)	Safety impact (context dependent)
Her/his assistant interrupted her/him and she/he answered her/him quickly while she/he was speaking to a pilot	Difficulties in managing communication	Allows her/himself to be disturbed by interruptions from colleagues	Communication, Teamwork	4
	She/he makes quick decision concerning priorities	Prompt to decide	Decision Making, Team Building,	+
She/he gave an instruction to her/his colleague without checking her/his availability	Difficulties in taking into account the workload of her/his colleague	Never manages her/his interruption of colleagues when she/he communicates	Communication	4
	Good management of time pressure and own workload.	Good stress management	Stress	+
Did not recover an error in her/his instruction to the pilot	Forgot very essential aspects of the situation without detecting it	Never checks the results of her/his actions	Error Management	4
	BOOMEE may have her/his reason to do that	However, does not recover her/his error systematically	Error Management	4
Non-standard strip position on the board	Use personal tricks to prevent error	Develops her/his own defences	Error Management	+
	Makes a procedure violation	Does not respect working procedure	Procedure	4
Enquires to the knowledge her/his assistant has of the traffic situation	Realised the loss of situational awareness and tried to avoid it.	She/he is attentive to her/his SA	Situational awareness	+
	Checked if the assistant shares the same SA as her/him	Questions situational awareness of team members	Teamwork	+
	Checked the situation before making a decision	Makes decision when having a good understanding of the situation	Decision Making	+
Asked her/his assistant for information in a hurried way	Inappropriate reaction	Shows her/his stress easily	Stress	4
	Takes into account the emergency of the situation	Communicates in the appropriate way	Communication, Teamwork	+
Make a “rhrhrh” noise while looking at the screen	Express an irritation	She/he expresses easily his/her stress	Stress	+/4
	Express lassitude	She/he is bored	Stress	+/-
No sector splitting when the assistant is asking for this	Needs of the colleague not taken into account	Never asks for help	Decision making, Teamwork,	4
	Better Situational awareness than her/his colleague	Not easily influenced	Teamwork, Situational awareness	+
Did not take into account the reservations (suggestions) from her/his colleagues	Did not question her/his own point of view	Never questions her/his own point of view	Teamwork	4
	The suggestion was inappropriate	Not easily influenced	Teamwork	+

Table 3.20 - BOOM example behaviours for debrief discussion (Available on request from Eurocontrol)

3.7.2 EUROCONTROL – Behaviour Observation Scale (BOS)

The Behaviour Observation Scale (BOS) has been developed as part of EUROCONTROL’s pan-European test programme for the selection and monitoring of trainee ATCOs (First European ATCO Selection Test (FEAST)). The BOS checklist was created through a review and amalgamation of various checklists sourced from a number of ANSPs. An initial set of 140 competency criteria were compared against an integrated FEAST task analysis model. Table 3.21 presents the results of an iterative process which reduced the initial set of competency criteria down to 38 (Rathje et al, 2004).

Inter rater consistency and reliability was evaluated through several paired observations of one subject (a Pearson score of between 0.286 and 0.648, for 34 out of 35 aspects contained within the prototype system found to be significant). A factor analysis revealed Teamwork, then Attention, Ambition, Working Under Stress, Communication, Aviation Context; demonstrated strong relationships with trainee ATCO competency using other measures taken during training (Eiβfeldt, 2003). The BOS employs a six point Behaviourally Anchored Rating Scale (BARS), grouped into three levels of overall performance (Inferior 4 to most trainees in similar stages of training, Comparable 4 to most trainees in similar stages of training, Superior 4 to most trainees in similar stages of training).

FEAST BOS Competency Elements	
1. Organises and maintains an efficient traffic flow	20. Constantly checks available information on incoming and outgoing data
2. Detects conflict early	21. Ensures traffic safety while at the same time taking account of economic aspects
3. Resolves conflicts effectively	22. Takes decisions quickly and in a responsible manner with due regard for priorities
4. Detects deviations	23. Improvises in situations requiring unconventional approaches or solutions
5. Corrects deviations	24. Communicates in a clear, unambiguous and to-the point manner
6. Operates technical systems	25. Adjusts tone of voice for messages in special situations
7. Masters the required ATC technical terminology	26. Profound knowledge and use of English language
8. Maintains attention over the entire shift	27. Shows stability (emotional control) in crisis situations
9. Insight into own limitations	28. Gives support to others if needed
10. Adapts to own limitations	29. Shows identification with the job
11. Teamwork skills	30. Shows initiative and motivation
12. Ability to identify with the pilots and understand their needs	31. Demonstrates leadership
13. Willingness to work according to company rules	32. Accepts and deals constructively with criticism
14. Works in an orderly way under pressure	33. Demonstrates self-confidence
15. Ability to detect and correct own mistakes	34. Demonstrates authority and decisiveness
16. Does not give in to pilot's demands when they are in conflict with own view	35. Demonstrates Flexibility
17. Controls in a way that does not create problems for other controllers	36. The trainee shows ambition to reach training goals
18. Shows consideration for colleagues	37. The trainee develops ATCO skills in appropriate time
19. Thorough knowledge of aircraft and their characteristics	38. The trainee is easy to handle for coaches

Table 3.21 - ATCO Behaviour Observation Scale (EUROCONTROL, 2005)

In addition to the 38 individual BARS ratings (Appendix B011), the observer is asked at the end of the observation sheet to provide one overall judgement of performance '*compared to other trainees in similar stages of training'*: using a 10 point scale equally divided into the following levels (Eurocontrol, 2005):

- Far below average
- Below average
- Average,
- Above average
- Far above average

3.7.3 Luchtverkeersleiding Nederland (Air Traffic Control the Netherlands)

The Dutch ATC organisation Luchtverkeersleiding Nederland (LVNL) has developed an in-house ATC competency assessment system for the assessment of Trainee ATCOs. This is an observational method for use by ATC examiners. The full competency system is commercial in confidence (confirmed through email communications to Esther Oprins, LVNL), however an example of the ATC planning task which is part of the LVNL competency framework including is presented in Table 3.22 (Oprins et al, 2006). This example includes a number of behavioural markers relating to planning.

Category	Element/Behavioural Marker
Planning	Makes a plan, executes the plan, and adapts the plan to (changed) circumstances
	Reverts to standard procedures if necessary
	Adapts his/her own plan to requirements and wishes of others
	Deals with procedures in a flexible way
	Works concentrated on his/her routine, but is able to interrupt the process at once, following another plan
	Makes a collective plan in collaboration with a colleague, while making concessions
	Proposes alternatives
	Is able to create order in chaos
	Presents the traffic situation schematically to realize an orderly flow of traffic

Table 3.22 - Planning Behavioural Markers, Oprins et al (2006)

The LVNL competency assessment system is an observational system for use for both training and examination purposes. Although the system is largely focused on technical skills and proficiency, there are a number of NTS embedded within it. Appendix B12 presents the competency framework for i) progression monitoring and ii) the final examination (Oprins, 2008).

The final examination form (Appendix B12) splits the competencies between physical and cognitive skills. In addition, the list of competencies includes attitude. Once again this is at odds with the collective opinion expressed in Klampfer et al (2001) that attitude should be excluded; however this opinion is in the context of assessing NTS rather than technical competency.

Table 3.23 presents a number of NTS identified by the author extracted from the two observation frameworks. These behavioural markers principally concern clear and effective communications and team working, task management and delivery, and attitude and confidence.

Category	Competencies
Situation Awareness & Attention	Keeps a good overview of the situation by scanning regularly
	Looks, observes and takes action if necessary
	Checks available information to be correct
	Guards the identification process of the label presentation
	Anticipates on future and variable traffic situations
	Can divide attention between several situations sufficiently
	Can perform several actions simultaneously
	Can park information in his/her memory without forgetting it
Decision Making & Planning	Is creative in inventing various solutions
	Can plan according to valid procedures and agreements
	Is flexible in adjusting plans
	Takes initiative and acts
	Shows confidence and takes the lead
	Acknowledges priorities correctly
	Shows confidence in taking the lead
Workload Management	Adapts work tempo to traffic load
	Stays calm, also during hectic moments
Communications	Expresses him/herself concisely, to the point, unambiguously and firmly
	Has a clear, quiet pronunciation and intonation
	Expresses him/herself in the Dutch language to conform with ICAO level 2
	Is easily approachable for others at the sector
Teamwork & Coordination	Communicates his/her plan concisely and to the point
	Makes clear arrangements and acknowledges these correctly
	Collaborates with others easily
	Is willing to adapt to common standards and values
Attitude	Shows responsibility during work
	Takes his/her training seriously
	Is eager to learn

Table 3.23 - NTS identified from the LVNL progression and examination competency frameworks (Adapted from Oprins, 2008)

Inter rater agreement was evaluated by Oprins (2008) through 22 dual instructor observations. These dual instructor observations involved two instructors observing the same trainee ATCO, and using the developed competency assessment sheet to rate individual aspects of performance. Dual instructor observations were made at two milestone points along an ATC training course (pre-OJT, and 'real' ACC training). The observational checklist employed a 6 point performance anchored scale divided equally between Insufficient and sufficient (Oprins, 2008). Additional qualitative information was collected from the instructors regarding the design and utility of the observational system.

Analysis of the results revealed moderate levels of inter-rater agreement for the dual observations made at the second milestone point (Pearson correlation of 0.56), however the observations made at the first observation milestone were found to be somewhat less reliable (0.46 – 0.6). Oprins (2008) concludes that 'some assessors give low ratings on some competencies while others give high ratings on the same competencies for the same trainee'. Oprins (2008) suggests that this is due to observers' different views on the required performance standards, reflecting the importance of training and standardisation to overall reliability.

3.7.4 Federal Aviation Administration (FAA)

The Federal Aviation Administration (FAA) over a period of 20 years and through a series of consecutive and independent projects has developed several over-the-shoulder (OTS) observational techniques to assess ATCO performance:

1. SACHA - a Behaviourally Anchored Rating Scale (BARS) design
2. A modified version of the SACHA observation sheet employing the use of Likert scales.
3. AT-SAT OTS - a Likert rating sheet for high level and low level task competencies

These observational systems have all been applied on several occasions, for different purposes, and have received minor amendments in order to answer application specific questions. This section provides the key detail of the three measures but does not provide an exhaustive chronological narrative of the various individual applications and minor iterative developments:

It is important to state that these observational systems focus predominantly on technical skills, and use ATC competency and training experts to undertake the assessment. They have been included in this review for the following reasons

- Development – how they were developed, how the competencies were identified.
- Design – several examples of observational sheets incorporating different design principles and design iterations.

- Domain – all of these were developed for use within the AT domain, and incorporate some NTS elements.

3.7.5 FAA - Separation and Control Hiring Assessment Project (SACHA) Observation Sheet

The first recorded behavioural observation sheet detailed in this section was developed by the FAA under the Separation and Control Hiring Assessment Project (Hedge et al, 1993). Although the original report is unavailable, the detail of the observation sheets' design and development is referenced in Borman et al (2001), and presented here.

The observation sheet was designed to focus on objective behaviours of performance, rather than subjective judgement. The design uses behavioural statements anchored to the performance scales with different descriptions for different levels of effectiveness, much akin to the design of a Behaviourally Anchored Rating Scale (which employs anchors at each end of the scale).

Borman et al (2001) states that the observation sheet was developed through a series of workshops undertaken with a total of 73 controllers teaching at the FAA academy, who generated 708 examples of effective, midrange, and ineffective controller performance. Hedge et al (1993) then identified eight categories across these performance areas:

- Communicating and Informing
- Managing Multiple Tasks
- Technical Knowledge
- Reacting to Stress
- Maintaining Attention and Vigilance
- Prioritising
- Maintaining Safe and Efficient Traffic Flow
- Adaptability and Flexibility

A further 24 controllers in five smaller workshops reviewed the categories and definitions identified from the initial workshops and a further two categories emerged (Teamwork, Coordinating), in addition an 'overall effectiveness' category was also added.

A final exercise involved SMEs from Tower, TRACON (Terminal Radar Approach Control), and en-route operations assigning performance behaviours to one (and only one) of the 10 performance categories, and then rating the level of effectiveness (from 1-7, very ineffective - very effective) for each individual performance element. The results of this work generated a three level performance framework for the 10 behaviours (high effectiveness, middle effectiveness, low effectiveness) with an anchored behaviour summary statement.

Appendix B13 presents the full ten categories contained within the SACHA observation system, including the behaviourally anchored high, middle, and low performance behaviours. Table 3.24 presents an example from a single category (Communicating and Informing).

<table>
<tr><th colspan="3">Communicating and Informing</th></tr>
<tr><td colspan="3">Uses clear concise accurate language to get message across unambiguously, talking only when necessary and appropriate; employing proper phraseology to ensure accurate communication; notifying pilots/controllers/other personnel of information that might affect them as appropriate; issuing advisories and alerts to appropriate parties; listening carefully to requests and instructions and ensuring that they are understood; attending to read backs and ensuring they are accurate.</td></tr>
<tr><td>Low Performance:
Is consistently too wordy, imprecise in phraseology, or uses slang inappropriately during transitions to pilots and other controllers
Is careless about informing pilots concerning circumstances that affect them such as weather, nearby traffic etc.
Often fails to ensure that own instructions are understood; is not very good at picking up on errors in pilot read backs of clearances, course changes etc.</td><td>Middle Performance:
Radio and interphone communications are usually easy to understand; at times, may be somewhat wordy or use unambiguous phraseology on the air
Is normally good at informing pilots about situations and conditions that affect them (e.g. safety related items)
For the most part checks to be certain that own instructions are understood; only occasionally fails to pick up on inaccurate read backs from pilots</td><td>High Performance:
Always uses clear, concise phraseology when talking to pilots or other controllers; is very easy to understand
Consistently provides pilots with the information they need such as timely safety alerts, weather advisories, warnings about unpublished obstructions
Always ensures that own instructions are clearly understood; pays careful attention to pilot read backs of clearances</td></tr>
<tr><td>Rating Scale:
1, 2, 3</td><td>Rating Scale:
4, 5</td><td>Rating Scale:
6, 7</td></tr>
</table>

Table 3.24 - SACHA observation sheet: Communicating & Informing (Borman et al, 2001)

3.7.6 FAA – Modified SACHA Observation sheet.

In 1997 the FAA developed a modified, complementary observation sheet using the Borman et al's (2001) SACHA work as a starting framework (Sollenberger et al, 1997). The fundamental difference with this observation sheet is that the behavioural anchors were removed, instead a series of technical competencies assessed through the use of a 8-point rating scale were added. Given the technical judgements which are required, for example managing traffic flows – the observation sheet can only be used by a certified ATCO.

The modified SACHA observation sheet focuses on the following technical competencies:

1. Maintaining Safe and efficient traffic flow
2. Maintaining Attention and Situation Awareness
3. Prioritising
4. Providing Control Information
5. Technical Knowledge
6. Communicating

An example section from the sheet is presented in Table 3.25, with the full framework and rating scale presented in Appendix B14 (Sollenberger et al, 1997).

Maintaining Safe and Efficient Traffic Flow	
Maintaining Separation and Resolving Potential Conflicts: Using control instructions that maintain safe aircraft separation Detecting and resolving impending conflicts early	1, 2, 3, 4, 5, 6, 7, 8, NA
Sequencing Arrival and Departure Aircraft Efficiently: Using efficient and orderly spacing techniques for arrival and departure aircraft Maintaining safe arrival and departure intervals that minimize delays	1, 2, 3, 4, 5, 6, 7, 8, NA
Using control instructions effectively: Providing accurate navigational assistance to pilots Avoiding clearances that result in the need for additional instructions to handle aircraft completely Avoiding excessive vectoring or over-controlling	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall safe and efficient traffic flow scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA

Table 3.25 - modified SACHA observation sheet: Communicating & Informing (Sollenberger et al, 1997)

The sheet was developed for the TRACON environment, and has been modified and applied to the en-route environment (Vardman & Stein, 1998). Minor modifications were made to the sheet for a subsequent study exploring reduced manning in the en-route environment (Willems & Truitt, 1999).

3.7.7 FAA – Air Traffic Selection And Training (AT-SAT) Rating Form

The third observation system developed by the FAA and contained within this literature review is a form developed by the AT-SAT project and used as part of a test battery to assess controller performance during ATCO selection and training. The principles of the form's design are that it should assess controller performance across broad dimensions of performance, as well as at a more detailed step-by-step level (Manning et al 2001).

The design of the form followed two phases. Firstly, technical performance dimensions were identified through review of previous existing on-the-job-training forms (including the SACHA form), performance verification forms, consultation with FAA training academy instructors, as well as surrounding ATSAT project work. The users of this form were FAA academy instructors.

A 7-point effectiveness scale was chosen, divided into three levels of performance (1-2: below average, 3-5: fully adequate, 6-7: exceptional). Initially eight performance dimensions were identified, along with behavioural descriptors for these dimensions:

- Maintaining Separation
- Maintaining Efficient Air Traffic Flow
- Maintaining Attention and Situation Awareness
- Communicating Clearly, Accurately, and Concisely
- Coordinating
- Performing Multiple Tasks
- Managing Sector Workload
- Facilitating Information Flow[3]

The full AT-SAT OTS form can be found in Appendix B15, however the *Maintaining Separation and conflict resolution* dimension and behavioural descriptors is presented in Table 3.26.

Maintaining Separation:
• Checks separation and evaluates traffic movement to ensure separation standards are maintained. • Detects and resolves impending conflictions • Applies appropriate speed and altitude restrictions • Analyses pilot requests, plans and issues clearances • Considers aircraft performance parameters when issuing clearances • Establishes and maintains proper aircraft identification • Properly uses separation procedures to ensure safety • Issues safety and traffic alerts

Table 3.26 - FAA AT-SAT– Maintaining Separation & Conflict Resolution (Manning et al., 2001)

3 This was later dropped following rater feedback gathered during pilot testing.

3.7.8 National Air Traffic Services (NATS) Behavioural Markers Work

The technique of observation to assess ATC performance is a technique that has been used within NATS for over 30 years (Smith & Stamp 1973, 1974). Predominantly, observation has been used in the training environment (to assess the competence of trainee controllers), and the simulation environment (to assess ATCO interaction with new airspace designs, operational procedures, HMI, workstations and other equipment). Within NATS observation is a core component to ATCO continued training such as annual competency checks, and TRaining in Unusual Circumstances and Emergencies (TRUCE).

With regards to the employment of behavioural markers to evaluate NTS, there have been three individual development activities undertaken by NATS prior to the research undertaken in this thesis. The three systems contained within this review are:

- NATS – Heathrow Tower Markers
- NATS – Day-to-day Safety Observations
- NATS – Individual's Positive Team Behaviours

Each of these applications has been undertaken for an individual purpose. Unfortunately there is even less information available as to the way these systems were designed and developed than the other behavioural observation systems contained in this review. However their technical content is of value, and has provided additional insight into NTS assessment within ATC. However, the lack of depth in both their development and validation highlight the large research gap which exists within this domain.

3.7.9 NATS – Heathrow Tower Markers

In late 2006, a new control tower at Heathrow was completed, and a training and implementation programme was undertaken to effectively transfer control staff across. A key challenge for Dr Steve Shorrock (a former NATS Human Factors expert working on the new Heathrow Tower installation) was to demonstrate user confidence and competence in order to provide human performance safety assurance. The project was extremely challenging, as not only was there a new tower location and internal physical layout but some of the key equipment also changed (Most notably Electronic Flight Progress Strips replacing Paper Flight Strips). Shorrock's (2007) greatest challenge was that although ATCOs were demonstrating task competency with the new equipment, their user attitudes remained low, and there was a degree of resistance from unit staff regarding the transfer to the new tower.

Table 3.27 presents a set of markers developed by Shorrock (2007) through a review of internal documentation within NATS (sources unspecified). This set was used in a tower simulation and in the live operational environment where ATCOs were 'shadowing' live operations; during the transition of ATC staff from the old control tower to a new tower.

A five-point frequency scale (always, mostly, sometimes, rarely, never, n/a) was used to structure the observation of 26 behavioural markers, structured under 3 NTS categories (Workload, Situation Awareness, Teamwork). The purpose of observing these behavioural markers was to assess the levels of safety and good practice demonstrated, in order to report and present levels of proficiency back to the unit, and raise levels of confidence in the new system across the ATCO population. At the end of each observation, a short debrief was undertaken with the ATCO in order to ask specific questions concerning arrival and departure runway sequencing and management.

Shorrock's (2007) work provided the unit with the additional confidence and assurance information which was needed. Shorrock's (2007) work provided the research concept for this EngD thesis, to develop a set of behavioural markers for ATC.

Category	Element/Behavioural Marker
Workload	Controller interacts with FDE during transmission.
	Clearance information is recorded.
	Controller keeps on top of RT loading.
	Controller remains calm whilst carrying out role, with no obvious signs of overload or frustration.
	Controller talks socially during less busy periods.
	Rate of speech allows for clear communication.
	Debrief: Bearing in mind the nature of practice simulations, was workload within comfortable limits during this session?
Situation Awareness	Controller looks up during each transmission.
	Controller performs regular visual scan of airfield.
	Controller able to find FDE when aircraft calls.
	No obvious signs of confusion.
	Controller detects and corrects any readback errors.
	Controller detects and resolves conflictions.
	Controller shows no signs of becoming tunnelled into EFPS.
	Controller avoids or immediately rectifies runway bay scrolling
	Bearing in mind the nature of practice simulations, was your level of situation awareness reasonable during this session?
Teamwork	Verbal coordination is clear, concise and timely.
	Verbal coordination and prompts from others are ack'd.
	Controller requests clarification, if unsure.
	Controllers alert each other to problems or erroneous info.
	The mood of the team is positive.
	Bearing in mind the nature of practice simulations, did the level of team interaction enable you to perform your job appropriately?
Planning (questions for debrief only)	Controller is able to integrate departing traffic on arrivals runway effectively.
	Controller is able to plan departure aircraft intervals and sequence effectively.
	Controller is able to remember sequence of departing aircraft.
	Bearing in mind the nature of practice simulations, were you able to plan ahead adequately?

Table 3.27 - NATS markers of favourable NTS performance (Shorrock, 2007)

3.7.10 NATS – Day-to-day safety Observations

The Day-to-Day safety observation sheet was developed by a number of SME ATC and safety experts within NATS (NATS, 2007). This behavioural observation system is designed as a ‘litmus test’ to collect information on a regular basis, in the live operation. This data is then used in order to determine levels of day to day safety, via the prevalence of specific behaviours associated with positive safety behaviour.

The Day-to-Day system is used across NATS, and was developed and refined through the consultation and contribution of ATCOs, and observational experts, although the detail of this is un-documented and unpublished; information on the system had been gathered through internal NATS correspondence and discussion with colleagues.

The format used to undertake a Day to Day observation is to observe an individual ATCO on three separate sessions, in order to capture the prevalence of individual behaviours (Table 3.28) displayed across the 5 safety areas contained in the system (visual scanning, active listening, defensive controlling, multitasking, and strip management). The frequency of individual behaviours is recorded using a three point scale (Always, Sometimes, Never, No opportunity to observe). In addition, contextual detail for each observation session is recorded (Location, date, time, position, observer, traffic level, weather.). Day-to-Day observers are trained in the technique prior to collecting data. It is understood that training comprises a briefing and concurrent dual observation with another trained observer.

Safety Process/Procedure		Behavioural Marker
Part 1	**Visual scanning cycle Frequency**	Memory aids are used as a prompt.
		Scanning cycles are completed.
		Prompts from others are received and acknowledged.
		FPS is used to verify received information.
	Active listening Frequency	Phone calls are deferred when a read back is being received
		One party is signalled to stand by whilst paying attention to other information.
		Headsets are used when on position.
		The intercom or telephone is used for interposition coordination where applicable.
		Incorrect aural information is corrected.
		Standard phraseology is insisted on and used.
		Incomplete read backs are not accepted.
Part 2	**Defensive controlling Frequency**	Clarification is requested on clearances given by the observed position.
		Aircraft on headings are initially given safe levels.
		A full and correct read back is insisted upon.
		Situations which require extra monitoring for safety are not created.
		Initial headings are correct.
		Conflicting traffic is positioned in such a way that it will miss would a plan fail.
		Incorrect information is always corrected.
		A common handover procedure is used.
Part 3	**WAYSRAYL Frequency (Write As You Speak, Read As You Listen)**	The controller consults FPS during transmission or reception of information.
		The controller amends or annotates FPS during transmission.
		Where relevant, transmitted information is consistent with FPS information.
		Where relevant, received information is consistent with FPS information.
		Clearance information is recorded on the FPS.
	Strip management Frequency	The FPS are moved in order to display the relative position of the traffic.
		Tactile methods are used to highlight specific events.
		Highlights are removed when no longer needed.

Table 3.28 - NATS Day-to-Day Behavioural Safety Markers (NATS, 2007)

The Day-to-Day observational system shares similarities in design and structure to other behavioural observation systems, although the choice of a shorter frequency scale differs in approach from benchmark systems developed in the medical domain (ANTS, NOTSS, SPLINTS). The differences between the Day to Day observational system in comparison to other benchmark systems extends to the detailing of the method used to identify and develop the behavioural markers, the depth of critical evaluation undertaken (most notably the testing of reliability and validity), and the presentation and documentation of these findings.

3.7.11 NATS – Individual's Positive Team Behaviours

Developed by Dr Anne Isaac (Directorate of Safety, NATS), these markers of positive behaviour (Table 3.29) are intended for use in the assessment of an individual's level of team working (Isaac, 2007). Thereby its use is to identify and assess if someone's behaviour is negative within a team; and whether a decision on their operational status should be undertaken (e.g. someone who is demonstrating bullying behaviour.). No further development of this marker set has been undertaken, including practical application within an observational tool. However it represents a useful set of informative behaviours of relevance to include within this review, despite the limited detail regarding development.

Category	Element	Behavioural Marker
Situational Leadership	Leadership	Ensures all other staff involvement when required
		Considers suggestions from staff
		Visible and accessible
		Positive and enthusiastic
		Shows appreciation
		Gives feedback
		Explains own position
		Takes control of the situation when required
		Seeks and shares information when required
	Maintenance of Standards	Intervenes when other operational staff deviate from standards
		Demonstrates desire to achieve high standards
		Encourages vigilance in high risk and safety critical situations
	Workload Management	Recognises signs of stress, fatigue and overload in self and others
		Responds effectively to high, low and fluctuating workload
		Shares tasks when required
Teamwork	Teambuilding	Relaxed, supportive and approachable
		Creates atmosphere for open communication
		Encourages inputs and feedback from others
		Polite and friendly
	Support of Others	Helps other operational staff in difficult or demanding situations
		Offers assistance
		Accepts assistance
		Recognises the needs of other staff
	Conflict Resolving	Keeps calm
		Suggests solutions
		Addresses conflict with persons concerned and concentrates on facts not personalities
		Non confrontational
Situation Awareness	Communication	Briefings are open, interactive and identify risks, hazards and errors if known
		Listens actively
		Gives and receives feedback in a respectful manner
		Challenge others appropriately and constructively in a non confrontational manner
	Decision Making	Evaluates own performance
		Clarifies uncertainty
		Seeks help when workload increases in both self and others
		Avoids distractions
		Reports unsatisfactory and safety critical procedures, documents or equipment

Table 3.29 . NATS Individual's Positive Team Behaviours (Isaac, 2007)

3.8 COMPARISON OF THE BEHAVIOURAL OBSERVATION SYSTEMS REVIEWED

Within this chapter, 19 separate observational systems for the assessment of technical and Non-Technical Skills have been presented. These observational sheets and behavioural frameworks have been developed for different purposes, and reflect different ideas of how to capture and record observational data within different operational and non-operational domains. Within the domain of air traffic control, only preliminary work has been undertaken in order to derive a comprehensive set of behavioural markers to evaluate ATCOs NTS performance.

This next section provides three summary Tables (Tables 3.30-3.32) which facilitate the comparison of the key learning points taken from the systems which have been reviewed and presents the key properties and information regarding their design, evaluation, and practical application. The comparison is structured into three parts;

1. Information sources used in the identification of the markers themselves,
2. General content and categories of the system,
3. Design and testing of the developed marker system.

3.8.1 Comparison of sources used to identify behavioural markers

Table 3.30 presents the five principal sources of information Flin et al (2008) recommend for the identification of NTS, and indicates which of these have been used by the 19 behavioural observation systems reviewed. An ‘x’ in Table 3.30 indicates that a marker system has used this source, as part of its identification and development.

Across the 19 systems, the source of domain experts was used to elicit potential markers without exception. With several of the systems reviewed, experts were used to provide context validity regarding their scope and utility of the product developed.

The least common method used to identify markers is through observation. It is likely that practical limitations of access and time needed to identify behaviours thorough observation accounts for this. Material from incident and investigation reports is also infrequently employed, and is also likely due to limitations of accessibility, and the content constrained within a report.

However it is the NATS Heathrow tower system which is unique amongst the 19 behavioural observation systems, in that it only used a single source to identify suitable markers (although the expertise and contributions of the author must be considered). The lack of diversity of sources employed by the NATS Heathrow tower system reflects the significant time pressures; where the demand for and application of an observational tool outweighed the underpinning method used to produce it.

Observation System	Principal sources used in development, Flin et al (2008).				
	Source 1	Source 2	Source 3	Source 4	Source 5
	Published literature	Technical documents	Incident material	Domain experts	Observation
U. of T./AC 120-51A			x	x	
LOSA			x	x	x
NOTECHS	x	x	x	x	
NASA Competencies	x	x		x	
ANTS	x	x	x	x	
NOTSS	x	x	x	x	
SPLINTS	x	x	x	x	
Maritime Crisis Mgt.	x			x	x
NTSOD	x			x	
RSSB	x	x		x	
Eurocontrol: BOOM	This information is not publicly available.				
Eurocontrol: BOS		x		x	x
LVNL: Competency	x	x		x	
FAA: SACHA				x	
FAA: Modified SACHA				x	
FAA: AT-SAT		x		x	
NATS: Heathrow Tower		x			
NATS: Day to Day		x		x	
NATS: Positive Team	x	x		x	

Table 3.30 - The five sources used by the 19 Behavioural Observation Systems reviewed ('x' indicates the source was used in development of the system)

The majority of systems have been developed through an iterative process of documentation review and consolidation, assimilation of information from other sources such as incident reports, and through the consultation of experts (either individually through interview, or collectively through a workshop forum). Throughout this process the system is consolidated and simplified as needed in order to maximise utility and clarity. Further refinements may be made through prototype application and widespread testing.

3.8.2 Comparison of content

Table 3.31 presents a collated list of NTS areas covered across the 19 behavioural observation systems reviewed. In order to produce the collated list presented in this table, the author has condensed the broader set of categories included in the 19 Behavioural observation systems into eight NTS areas. As part of this process, technical competency areas have been excluded. The resulting eight NTS areas contained within the table are as follows:

1. Teamwork, Cooperation,& Team Building
2. Leadership
3. Stress Management & Self-Care, Workload & Task Management
4. Situation Awareness, Attention & Vigilance
5. Decision Making & Planning
6. Communication
7. Conflict Management, Problem Solving, Error Correction, Adaptability
8. Attitude

Teamwork and team building feature most prominently across the 19 different systems, closely followed by the inclusion of 'Situation Awareness and Attention'. The category which is least covered is 'Attitude'. Klampfer et al (2001) explicitly recommend that attitudes or personality traits are not included. Whether intentional or not, all but three of the systems reviewed adhere to this recommendation.

Table 3.31 also presents in the far right column, the number of elements and where applicable behavioural markers contained within each observational system. Across the 19 systems reviewed, three collective approaches have been employed with regards to the use of behavioural markers:

- The first is that of a completely flexible and open approach which does not provide a set of ridged and specific behavioural markers but instead empowers the observer to use their own schema and expertise to note and evaluate significant behaviours observed (e.g. LOSA, SACHA).

- The second concerns the provision of 'example' markers; however the system stops short of providing a rigid and specific set of behavioural markers (NTSOD, ANTS).

- The final method employed involves the use of a prescriptive set of behavioural markers which the observer must focus their attention upon (LVNL, NATS).

Observation System	Generic NTS Categories								Number of Elements and/or Behavioural Markers Contained
U. of T./AC 120-51A	x	x	x	x	x				13 elements 62 Behavioural markers of positive performance (AC 120-51A)
LOSA	x		x		x				17 elements
NOTECHS	x	x	x	x	x				15 elements 44 good practice behavioural markers 44 poor practice behavioural markers
NASA – Competencies	x	x		x	x	x	x		24 elements 104 positive behavioural markers
ANTS	x		x	x	x				15 elements 58 behavioural markers of good practice 56 behavioural markers of poor practice
NOTSS	x	x		x	x	x			12 elements 52 behaviour examples of good practice 51 behaviour examples of poor practice
SPLINTS	x		x	x		x			9 elements 43 behaviour examples of good practice 41 behaviour examples of poor practice
Maritime Crisis Mgt.	x		x	x	x	x			19 behavioural markers with 23 further examples of positive markers 34 further examples of negative markers
NTSOD	x	x		x	x	x			10 elements 10 example positive behaviours
RSSB	x		x	x	x	x			26 elements 93 positive behavioural markers 93 negative behavioural markers
Eurocontrol: BOOM	x			x	x	x	x		9 behavioural marker examples
Eurocontrol: BOS	x	x	x	x	x	x	x	x	38 negative and positive Behaviourally Anchored Rating Scale (BARS) points
LVNL: Competency	x		x	x	x	x		x	Progression Report: 46 Technical/Non-Technical Competencies Final test ACC: 39 46 Technical /Non-Technical Competencies
FAA: SACHA	x		x	x	x	x	x		3 level (low, medium, high) Behaviourally Anchored rating scales for the 10 technical and non-technical areas covered
FAA: Modified SACHA	x		x	x	x	x	x		42 technical competencies
FAA: AT-SAT	x		x	x		x			51 positive technical competencies
NATS: Heathrow Tower	x		x	x					22 positive behavioural markers
NATS: Day to Day	x		x	x	x	x			27 positive behavioural markers
NATS: Positive Team	x	x		x				x	36 positive behavioural markers

Table 3.31 - The general content of the 19 Behavioural Marker systems reviewed

The final column of Table 3.31 also provides an indication of the overall size of the behavioural observation system. With 186 and 108 markers respectively, the RSSB and NASA behavioural marker sets are the largest of the 19 behavioural observation systems reviewed. A marker set of this size is arguably beyond the utility of a single observation sheet, or single observation session. It is likely that the differences in size and content of the behavioural observation system are reflective of the purpose for which they have been designed, and the level of maturity of the overall system and the level of usability testing and evaluation that has been conducted.

Whilst most systems provide a comprehensive set of behavioural markers for use to assess a collective set of NTS within a specific domain, other systems are more prescriptive focusing upon only safety related NTS (e.g. NATS Day-to-Day), or those which are of importance to crisis or threat management (e.g. LOSA). Therefore the associated task context has a significant framing effect upon what is included and excluded in each system.

3.8.3 Comparison of testing and evaluation

Table 3.32 presents details regarding the testing and evaluation of behavioural observation systems derived from the product of the preceding works (Tables 3.29 – 3.30). The information presented reveals a number of gaps and unknowns regarding the marker systems reviewed. For most of the systems reviewed a significant amount of effort has been taken in the not inconsiderable task of identification selection and incorporation into an observational tool. As a result there has been more limited activity undertaken to practically apply and fundamentally evaluate their real-world functional performance and utility.

With regards to training, most systems that provide detail in this area broadly adhere to Klampfer et al’s (2001) recommendation of two days to train observers. A mixture of class room and practical training has generally been adopted, including the opportunity to ground scores between observers ahead of actual data collection (van Avermaete & Kruijsen 1998; Flin et al, 2003).

Although there is some variability in terms of the number of scale points, all of the observational systems reviewed either require the observer to rate the quality of the behaviour or competency area demonstrated, or the ordinal frequency of occurrence that a certain behaviour is displayed. Several systems, particularly those developed by the FAA which focus on the evaluation of technical performance require the observer to make a final judgment of overall competency based upon the summation of the observations made (Hedge et al, 1993; Sollenberger et al; Manning et al 2001).

Observation System	Training for Observers	Scale Design/Length	Inter Rater Reliability	Application Environment
U. of T./AC 120-51A	Yes	4 point performance anchored scale	Iterative development using SME observers	Simulation Real world
LOSA	Calibrated observation	4 point performance anchored scale	Iterative development using SME observers	Simulation Real world
NOTECHS	Briefing and practice session for trial. 2 days recommended for final system	5 point performance anchored scale	80% IRR agreement of pre-recorded video footage, Flin et al (2003)	Simulation Real world
NASA - competencies	Not developed to this stage			
ANTS	Two 10 minute video clips and briefing (total 4 hours)	4 point performance anchored rating scale	Mean inter-rater agreement (R_{wg}) between 0.5 – 07	Simulation
NOTSS	Three 5 minute video scenarios and briefing (2.5 hours)	4 point performance anchored rating scale	Inter-rater agreement R_{wg} (>0.6) all categories except Task Mgt. Intra-class correlation coefficients (.95-.99).	Simulation Real world
SPLINTS	5 hours of training and 1 hour of practice	4 point performance anchored rating scale	SME observers Within group Rwg Agreement extensively ≥ .7).	Simulation Real world
Maritime Crisis Mgt.	None except an example sheet of good/bad behaviours	4 point performance rating scale	Mean within-group Inter-rater agreement (R_{wg}) between 0.35 – 0.73. Mean percentage of total observed responses was 95.7% for all 19 markers	Simulation
NTSOD	unknown	4 point performance rating scale	Cohen's kappa of 0.91 for the categorisation of markers by 2 SME raters	unknown
RSSB	Not developed to this stage			
Eurocontrol: BOOM	Briefing day including two 1 hour video practice sessions	n/a	Iterative development using SME observers	Simulation Real world
Eurocontrol: BOS	unknown	Thirty eight 6-point behaviourally anchored performance rating scales One 10 point overall performance scale	Limited IRR from two observers and one subject (Pearson 0.286- 0.648, with 34 out of 35 aspects significant)	Simulation Real world
LVNL: Competency	Theory lessons and practical exercises	Six-point performance anchored scale	Dual observations to evaluate IRR revealed moderate reliability (Pearson's q 0.46 – 0.6)	Simulation Real world
FAA: SACHA	unknown	7-point performance scale, divided into three levels of performance	unknown	Simulation
FAA: Modified SACHA	3 days observer training. Briefings, video practice sessions and peer scoring comparison	8 point performance scale	2% inter and intra rater co-efficient above 0.6 for individual sub categories (type of correlation unspecified)	Simulation
FAA: AT-SAT	Briefing session for instructor observers as part of a wider week long rater training programme	7-point performance scale, divided into three levels of performance	2-rater intraclass correlation coefficients (ICC) (between 0.71 – 0.98 for each category)	Simulation
NATS: Heathrow Tower	None	5 point frequency scale	None	Simulation Real world
NATS: Day to Day	Yes	5 point frequency scale	Iterative development using SME observers	Real world
NATS: Positive Team	Not developed to this stage			

Table 3.32 - Application details of the 19 behavioural marker systems reviewed

Where it has been evaluated, inter-rater agreement has been assessed through paired (or more) observations. This has extensively been achieved through the use of pre-recorded video footage as the source material for the observations (Fletcher 2003, Sollenberger et al, 1997). There are several instances where the task of the observer is to observe pre-recorded video footage – rather than being present in the environment. Pre-recoded footage has the benefit of being used on multiple occasions and it may be walked through and explored in detail to ensure nothing is missed. However it is a ‘window’ into behaviour, and does not provide i) the opportunity to absorb the environmental situation and surrounding operational context, or ii) the opportunity to liaise with the person observed and ask them any questions regarding their behaviour and tasks performed.

The significant majority of these systems have been developed for use within the simulation environment, rather than the live operational environment. In addition certain systems have been developed to evaluate video recorded behaviour as opposed to ‘live behaviour’; this may in part be due to the system undergoing reliability testing rather than general application (Fletcher et al, 2003; Kontogiannis & Malakisis, 2013). There is little mention as to whether behaviour differs between simulated and real-world environments. Researchers in the driving domain have found differences in behaviour and performance between the simulation and real-world environments (Santos et al, 2005; Riener, 2010). The implication of these findings highlights the need to evaluate a behavioural marker system in multiple environments in order to ensure reliability.

3.9 DERIVING THE RESEARCH QUESTION

Poor NTS clearly can impact upon overall task performance and can lead to accidents (Helmreich et. al, 1995). Research in the aviation and medical domains has pioneered the development of robust measures to evaluate NTS performance. Within the domain of ATC, observational research has broadly focused upon measuring technical task competency. The FAA has developed several observation tools to measure aspects of task performance (SACHA, modified SACHA, AT-SAT OTS), which on occasion include aspects of NTS (Sollenberger et al, 1997; Borman et al, 2001; Manning et al 2001). However the FAA tools focus upon competencies rather than the extrapolation to specific behavioural markers. This convention remains broadly consistent for the competency observation system developed by Oprins et al (2006) for LVNL; although a handful of example markers (Table 3.22) are disclosed for publication under the confidentiality agreement covering the work (Oprins, 2008).

Rathje et al (2004) produced the Behaviour Observation Scale (BOS) for Eurocontrol to be used as an assessment tool to measure trainee ATCO competence. The BOS tool includes elements of NTS, although the design is orientated around the use of BARS scales, and does not present a set of specific behavioural markers as advocated by Klampfer et al (2001). In addition, the Behaviourally Orientated Observation Method (BOOM) developed by Eurocontrol (2003) presents only a rudimentary framework for observation, leaving the judgement of significant behaviour dependent upon the perception and interpretation of the Observer.

The NATS (2008) work to develop the Day-to-Day safety observation system focused specifically on NTS associated with good safety practice, rather than a wider set of NTS for ATC. Isaac's (2007) preliminary work to identify positive team behaviours is a further example marker set within this domain however detail regarding its development - as is the case for the Day-to-Day system - is unfortunately undocumented. This fact perhaps reflects the resource limitations and constraints that may be experienced when developing NTS observation systems, and in some ways explains the lack of publication within the scientific literature.

The work undertaken by Shorrock (2007) to develop a NTS behavioural observation tool for NATS has adhered to a structure and design in keeping with best practice and convention (Flin et al, 2008). However the method used to identify the markers has been only one source (internal technical documents), and no testing or evaluation of the tool has been undertaken. It is on this basis that the research topic for this thesis received steerage from the Industrial sponsor, NATS, to produce a behavioural markers system for ATC.

Given the approach and product of the works undertaken, it is argued that there is a need to look at the topic from first principles in order to generate a robust and best practice derived set of behavioural markers for the assessment of Non-Technical Skill proficiency in ATC. This extends to i) expanding the depth and number of sources consulted in the identification of markers, ii) the adherence to Klampfer et al's (2001) best practice rules and, iii) Flin et al's (2008) recommended linkage to performance outcomes, and finally iv) the widespread iterative testing of the tool including reliability and evaluation of utility. There is, therefore, a fundamental research question that remains unsatisfied, which has generated research question 1.

3.9.1 Research Question 1

- What non-technical behavioural markers may be used to evaluate ATCO performance?

SUMMARY OF CHAPTER 3

In this chapter, the area the historical context and development of CRM is presented, and the methodology employed to examine the proficiency of Non-Technical Skills within a number of safety critical domains. The behavioural markers observational techniques which have been developed in order to evaluate NTS has been presented, including fundamental design principles and concepts from the literature. A number of NTS and other key observational systems have been catalogued from several safety critical domains, including Air Traffic Control. Key information concerning the design and development of these systems, where available, has been presented. Within the ATC domain, there has been only limited work to develop a comprehensive set of behavioural markers to assess ATCO NTS, with the works undertaken having generally not followed the approach adopted by other benchmark systems. As a consequence a fundamental research gap exists for the identification, development, and testing of a robust and best practice derived set of behavioural markers for the assessment of Non-Technical Skill proficiency in ATC.

CHAPTER 04 – PRELIMINARY OBSERVATIONAL STUDY

"Sometimes a cigar is just a cigar."
(Sigmund Freud)

This chapter presents the preparatory works to develop a preliminary measure for the assessment of Non-Technical Skills (NTS) within Air Traffic Control (ATC). It provides detail on the identification of a behavioural marker set, their integration within a prototype observation framework, the design of an observation sheet, and its preliminary application within an ATC environment. This work was undertaken in order to answer Research Question 1: What non-technical behavioural markers exist to evaluate Air Traffic Controller (ATCO) performance?

4.1 AIMS OF THE STUDY

1) Identify behavioural markers indicative of both good and bad NTS performance.
2) Produced an observation sheet and underlying methodology for NTS observation in ATC.
3) Undertake a number of observations in order to gather 'baseline' data of ATCOs engaged in normal day-to-day operations.
4) Trial the marker set and observational method.
5) Review the findings of the study, and make changes as required.
6) A final aim of this study was for the researcher to gain experience of undertaking the role of an observer within this domain.

4.2 IDENTIFICATION OF BEHAVIOURAL MARKERS

Flin et al (2008) identify five principal sources for the identification of NTS behavioural markers (Chapter 3, Table 3.3). The study presented in this chapter has employed all five of these sources:

1. Published literature
2. Technical documentation
3. Incident material
4. Domain experts
5. Observation

The first four sources were used in the identification phase prior to the creation of an observation sheet. The fifth source (identification through observation) took place during the observational study discussed later in this chapter.

4.2.1 Identifying Behavioural Markers: Sources 1 - 3

In terms of Flin et al's (2008) first two sources Chapter 3 provides an extensive series of behavioural marker sets available in the literature for the assessment of NTS in several safety critical domains including ATC. These have been gathered over the course of the research presented in this thesis.

The work presented in this chapter represents preliminary research. As a result of chronology, a smaller number of NTS systems and competency frameworks had been reviewed (Table 4.1), at the time of this preliminary study.

NTS and other Observational Frameworks	Reference
University of Texas (NASA/FAA) Behavioural Markers	Helmreich et al. (1994)
Line Oriented Safety Audit (LOSA)	ICAO (2002)
Non-TECHnical Skills (NOTECHS)	van Avermaete & Kruijsen (1998)
NASA – Human Behavior & Performance Competency Model	Bessone, et al. (2008a, 2008b)
Anaesthetists Non-Technical Skills (ANTS)	Fletcher (2006)
Federal Aviation Administration (FAA) AT-SAT	Manning et al. (2001)
NATS: Heathrow Tower Safety Markers	Shorrock (2007)
NATS: Day-to-day Safety Observations	NATS (2007)
NATS: Individual's Positive Team Behaviours	Isaac (2007)
NATS: A series of optimum job performance standards of ATCOs graduating from the ATC Training College	Low (2004)
NATS ATCO Competency framework: En route, Aerodrome & Approach Controllers	Thomas and Robinson (2008)

Table 4.1 . Literature sources used to identify markers for Chapter 4 study

Incident reports contained within the NATS Safety Tracking And Reporting (STAR) database (Source 3) were also reviewed. However of the reports reviewed, there was excellent technical detail regarding the nature of the incident, but unfortunately there was very little detail on individual ATCO behaviour; a scenario experienced by other researchers, Fletcher et al (2004). Reports contained within the NTSB Aviation Accident Database were also examined, and were consistently regarded by the author to offer low utility in the further identification of ATC specific NTS behaviours. This experience may account for the lack of use of this type of source for the identification of markers by other behavioural observation systems (Table 3.31).

In order not to consume a great deal of resource further investigating these potential sources, given the utility of the source which had been found thus far, the exercise was discontinued. This does not, however, detract from the value that first hand analysis can provide of an incident, for example the use of the critical incident interview technique employed for the development of the ANTS observation tool, can provide a fruitful means by which potential NTS and other behavioural markers may be identified (Flin & Fletcher et al, 2003).

4.2.2 Behavioural markers identified from Sources 1&3

Sources 143 were used to identify a total of 104 potential behavioural markers (including behaviours, competencies, and skills). Source 1 identified 35 potential markers; Source 2 identified 69; and Source 3 identified none.

The potential behavioural markers were selected from the literature sources where they were considered to be generic or transferable to the ATC domain. All of the potential behavioural markers selected were considered to reflect aspects of superior and inferior NTS performance (positive and negative behaviour). These potential markers covered areas such as communications, interaction task behaviours, as well as posture and body language. Examples of the behavioural markers elicited through Sources 144 are presented in Tables 4.4 and 4.5.

Examples of Positive Behaviours Identified
Ensures team involvement, seeks and shares information
Positive, shows appreciation & gives positive feedback to team
Controller performs regular visual scan of displays
Controller able to find Flight Data when aircraft calls
Controller detects errors and corrects/resolves appropriately
Calm and relaxed when dealing with pressure/stress, in control of emotions, no signs of frustration
Physical 4 sitting up (alert and attentive)
Physical 4 Leaning back/sitting back looking relaxed; crossed legs etc.
Physical 4 Not wearing headset (due to low workload)

Table 4.2 - Example indications of positive behaviour (coping behaviour)

Examples of Not Coping Behaviours Identified
Un-influential; does not exert authority
Very low amount of RT for traffic situation
Preoccupied, distracted, or fixated with specific tasks/minor tasks
Daydreaming/automaton 4 operating on autopilot and switched off
Anger and irritation, excessive swearing
Grumpy, irritable, and snappy; negative comments made
Physical 4 hunched over 4 overly focused on displays (tunnelled into workstation)
Physical 4 leaning into workstation, look of concentration and tense body posture
Physical 4 Looking flustered, rosy cheek
Physical 4 signs of tiredness; rubbing eyes/face, stretching, yawning
Physical 4 Obvious signs of discomfort when moving (e.g. bad back)

Table 4.3 - Example indications of negative behaviour (not-coping behaviour)

4.2.3 Identifying Behavioural Markers: Source 4

With regards to source 4, three Knowledge Elicitation workshops were undertaken, with the purpose of identifying potential markers for inclusion within a behavioural marker sheet. The three groups of Subject Matter Experts (SMEs) employed held knowledge and experience across a variety of relevant areas (Human Factors, Validation Simulations, and ATC), and would therefore provide a broad range of expertise upon which to consult. Table 4.2 shows that the first two workshops had small numbers of participants, however the third workshop had a good number of participants, and represented approximately 70% of the total population of technical validation experts within NATS.

Subject Matter Experts	Number of Participants	Date	No potential markers identified
Air Traffic Controllers (ATCOs)	2	06/05/2008	35
NATS Human Factors Experts	2	27/05/2008	86
NATS Validation Team Experts	6	06/06/2008	103

Table 4.4 - Knowledge elicitation workshops

At the start of each workshop, an introduction to the research topic was provided. Workshop participants were asked to identify behaviours associated firstly with someone who is not coping well with their task (with poor levels of task performance and safety); they were then asked to generate behaviours associated with someone coping (with high levels of task performance and safety). To aid further the elicitation process, participants were presented with a variety of alternative situation and environmental conditions which may evoke certain non-technical behaviours and changes to their psychological or physical state:

- Weather and time of day (day versus night)
- Routine versus unusual events, and emergencies
- Traffic Levels and complexity
- Team familiarity, and levels of harmony or conflict
- Training, experience, and familiarity with new systems
- Changes in status (faulty equipment, systems, fall-back procedures)
- Hunger, tiredness, fatigue, illness
- Corporate and management issues
- Ambient environment (noise, uncomfortable seat, workstation, lighting etc.)

The knowledge elicitation method used for the initial workshop held on the 6th May 2008 was undertaken using standard brainstorming techniques; where participants are all asked to contribute together in an interactive open forum. In order to elicit responses, ATCOs were asked to consider other colleagues' behaviours.

Having undertaken the first knowledge elicitation workshop, it was clear that the lack of structure in elicitation resulted in unequal participation by the experts involved. Reviewing alternative knowledge elicitation methods, the Nominal Group Technique (NGT) was identified as an alternative time efficient method of collecting the views of individuals within a group environment (Liou, 1998). The NGT involves participants anonymously writing down their input into the workshop before all the contributions are shared. A rating of the importance of all the contributions is then made, once again being written down anonymously. NGT was used as the knowledge elicitation method for the second and third workshops, and was found to afford great equality of comments from the expert participants.

4.2.4 Behavioural markers identified from Source 4

The three workshops (Source 4) identified a total of 224 potential behavioural markers. These markers covered both positive and negative behavioural attributes. Examples of the potential markers identified are presented in Table 4.3.

Example	Positive/Negative	Source
Planning Ahead	Positive	ATCOs
Timely organisation and response		
Say nothing versus saying something wrong		
Unfamiliarity with system	Negative	
Miss things due to noisy environment		
Swearing		
Calm and relaxed	positive	HF Experts
Well rested		
Happy		
Frustrated, Angry, irritated	negative	
Unsociable		
Preoccupied or distracted		
Creative	positive	Validation experts
Attention to detail		
Authoritative voice communications		
Over focus	Negative	
Poor team working		
Irrational behaviour		

Table 4.5 . Examples of behaviours identified through the knowledge elicitation workshops

4.2.5 Consolidation of Potential Behavioural Markers

Sources 144 identified in total 330 potential behavioural markers (Source 1 identified 35 potential markers; Source 2 identified 69; Source 3 identified none; Source 4 identified 224).

The 330 potential behavioural markers identified from sources 144 were added into a spread sheet and grouped together into similar categories. This grouping activity allowed a consolidation to be undertaken in order to remove i) duplicates, ii) markers with weak mutual exclusivity, and iii) markers which reflected desirable or undesirable character traits rather than overt observable behaviours. A number of principles for identifying good behavioural markers were used in this process (Klampfer et al., 2001). A mixture of 35 positive and negative behavioural markers emerged as a result of the consolidation process, in the following six categories:

1. Situation Awareness, Attention , Focus
2. User State: Stress Management, Fatigue, and Comfort, Frustration, Morale, Motivation
3. Decision Making, Planning, and Workload Management

4. Teamwork, Teambuilding and Team Support
5. Communications
6. Responding to When Aircraft Calls (Multi-tasking & responding to new inputs)

Four out of the six categories identified are consistent with the areas covered by other key NTS frameworks within the literature (Chapter 3). However there are two categories which represent a significant departure (‘user state’ and ‘responding to when aircraft calls’).

The first is the category of user state, which Klampfer et al (2001) explicitly declare should not be included. No clear explanation is provided by Klampfer et al (2001) as to why user state should be excluded, however it is likely it may be difficult to evaluate user state using common and generic qualities, rather than those concerned with individual difference. The markers identified under the category ‘user state’ were not considered to represent those of individual difference.

The second category of ‘responding to when aircraft calls’ concerns the reactive component of the ATC task, and involves the ATCO locating the aircraft on the radar and flight strip displays and responding as appropriate to the communications.

4.2.6 Consolidated Behavioural Markers Set

Tables 4.6 – 4.11 present the positive and negative behavioural markers identified which were considered to provide potential indications of inferior or superior NTS performance. These markers were allocated to the six topic areas identified, and form the basis of the observational tool used later in this chapter (Appendix C01 and C02).

Situation Awareness, Attention , Focus	
Positive Markers	Focus & Concentration – Fast response to issues requiring action, decisive, clarifies situation, deals well with uncertainty, posture alert and attentive, not easily distracted, attention to detail
	Strip scan 4 Checks through strips – runs finger/hand/pen down strips, cocks strips
	Workstation scan 4 scan of main workstation components; strips, radar, SIS
	Returns to a previously interrupted task
	Where applicable, Electronic Decision Support tools are used (e.g. vector lines)
Negative Markers	?Focus & Concentration 4 Preoccupied, distracted, fixated with specific/minor tasks,
	daydreaming/automaton , on autopilot/switched off, hunched/tunnelled into workstation
	?Awareness 4 appears confused, unable to concentrate, struggling to find aspects of system
	?Very low amount of RT for traffic situation

Table 4.6 - Situation Awareness: Positive & Negative Markers

User State: Stress Management, Fatigue, and Comfort, Frustration, Morale, Motivation	
Positive Markers	Stays Calm & Relaxed 4 unemotional, not hunched into workstation, no verbal frustration
	Adapts to pace of task 4 When quiet: sits back/legs crossed, reads paper/book etc, not wearing headset (due to low workload)
	Positive & responsible attitude: constructive, supportive, approachable, enthusiastic, polite, friendly, relaxed
Negative Markers	Poor concentration 4 , fidgety, restless, distracted, fiddling with systems/stuff, fidgety
	?Angry / Stressed 4 snappy, grumpy, irritated, excessive swearing, verbal frustration, Looking flustered, rosy cheeks
	?Physically Stressed – overly leaning into workstation, tense body posture
	?Uncomfortable / Fatigued 4 rubbing eyes/face, stretching, yawning, signs of discomfort when moving
	?Care and consideration 4 Displays a lack of care and respect for equipment (even vandalism)

Table 4.7 - User State: Positive & Negative Markers

Decision Making, Planning, and Workload Management	
Positive Markers	Maintains strips accuracy (updated by pen or keyboard) – Clearances, level changes etc
	Picks up the pace 4 as traffic level/complexity increases
	Keep on top of RT loading 4 doesn't miss calls, or ask a/c to 'standby'
	Demonstrates multi-tasking and divided attention
	Seeks assistance when workload increases
Negative Markers	?Task Rate/Response Rate 4 Inferior, delayed, or no response to actions and requests, Excessive stalling tactics/hesitation and task dropping, fixated on one task, over focus on easy tasks, Does not keep on top of RT loading, misses calls, asks a/c to standby
	?Excessive/inappropriate levels of help and requests for assistance

Table 4.8 - Decision Making & Planning: Positive & Negative Markers

Teamwork, Teambuilding and Team Support	
Positive Markers	Team member attitude 4 Shows appreciation, happy to receive help, gives positive/constructive feedback, enthusiastic, easily approachable
	Team problem solving 4 Helps others, alert each other to points of interest, problems, erroneous info, acknowledges prompts from others
Negative Markers	?As a Member of team 4 Overly competitive, patronising, negative, dominant, insular, unsociable
	?Team attitude – isolated team members, poor team mood

Table 4.9 - Teamwork: Positive & Negative Markers

Communications	
Positive Markers	Clarity of Comms 4 clear, concise, timely, not rushed, and authoritative, standard/correct phraseology used
	Defers calls 4 Phone calls are deferred, RT parties told to stand by when responding to other information/tasks.
	Quality & Accuracy of Comms 4 Incomplete, or incorrect read backs are not accepted, unclear/ambiguous messages are questioned, non-standard phraseology is challenged
Negative Markers	Verbal/Non-verbal communications 4 unclear, too long, poorly timed, too quick, to quiet/weakly conveyed
	?More than two instructions are given in the same transmission.

Table 4.10 - Communications: Positive & Negative Markers

Responding to Calls	
Positive Markers	Quickly Finds Flight Data when aircraft calls
	Consults FPS during transmission or reception of information.
	Amends or annotates FPS during transmission.

Table 4.11 - Responding to Calls: Positive Markers

Time and resource constraints resulted in the collation and condensation of the identified behaviours into a set of behavioural markers was undertaken solely by the author. Although the resulting marker sheet was presented informally to a couple of HF colleagues and a NATS ATC simulation expert, wider and more formal involvement of experts during this process was not possible. This process therefore has the inherent weakness of encoder bias, and as a consequence is susceptible to reduced inter-rater reliability (Hollnagel and Amalberti, 2001; Creswell, 2003). Chapter 7 explores inter-rater reliability with a marker system that has been developed later in this research. The observational system reviewed in Chapter 7 shares many similarities in terms of format and method with the one developed through the preliminary work presented in this chapter.

4.3 METHOD FOR PRELIMINARY BASELINE OBSERVATIONS

The behavioural markers identified were placed into a simple observation sheet (Appendix C01). This observation sheet was used to structure a series of observations undertaken in order capture the non-technical behaviour of mature, fully competent controllers using paper flight progress strips. This data would serve as a frequency baseline for which certain task behaviours occur during routine operations and enable a comparison when examining data of ATCO behaviour gathered in other situations (using electronic strips, training, adverse conditions, emergencies etc). En-route ATCOs at Swanwick were observed within the live operational environment, whilst undertaking the ATC task using paper flight strips.

The observations were made by the observer (author) sitting behind the shoulder of the ATCO who was seated in front of the workstation. This is a common position for instructors, examiners, and any other person undertaking ATCO observation within NATS. Notes were made on the back of the sheet and recorded potential indicators for ATCO state, reactions and interactions with the interface whilst engaging in the task, task strategies, quotes from the ATCOs, and any specific situation factors which existed to contextualise behaviours observed.

4.3.1 Recording behaviour

Existing methods for behavioural observation invariably involve the use of Likert scales to record frequency or performance judgements, towards or at the end the end of the observation session (Chapter 3). It is argued that this approach lacks sensitivity, and has a degree of subjectivity. In order to generate data with high sensitivity and to help reduce the potential for observer or recollection bias, an uncapped frequency tally was chosen (i.e. for every occasion that a behaviour is observed, a tally was made to record the occurrence). In addition to these frequency notes, the observer (author) noted information regarding the environmental and task situation, and any other interesting observations.

Through focusing specifically on the recording of overt explicit behaviour, the method precludes the direct recording of omissions. For example, certain task situations may warrant a particular action to be performed, and the absence of this action may be a useful indicator of sub optimal performance. In order to spot this omission, the observer must be highly skilled in the task. The detection of omitted behaviour lends itself more to scripted test environments where the occurrence (or not) may be anticipated by the observer. Observations made during this research have been undertaken within the lightly scripted simulation and un-scripted live operational environments focusing purely on recording increases (or decreases) in skill level and various indications of proficiency over time.

4.3.2 Identifying Behavioural Markers: Source 5

In addition to the structured observation, unstructured observation was undertaken in parallel. This unstructured observation had the specific purpose of recording further potential behavioural indicators as they presented themselves. Potential markers were recorded that were considered by the observer (the author) as affording insight into the psychological or physical state of the controller, and their engagement with the ATC task.

4.4 RESULTS

In total, 25 observations were made during July 2009, observing the behaviour of 22 individual ATCOs. The first 13 Observations were made on the 8^{th} , 9^{th} , and 10^{th} of July 2009 using a simple observation sheet (Appendix C01). The average length of observation was 29 minutes (the longest 61 minutes, shortest 6 minutes). A number of changes were made to the observation sheet (Appendix C02) in order to consolidate markers and enhance utility, as well as to incorporate a number of new markers identified through the first 13 observations. A comparison of the two observation sheets is presented in Appendix C03. The most significant change between the two sheets was the removal of the Decision Making category (this was considered unclear and difficult to judge by the author using the markers provided). This second sheet was used for a further 12 observations on the 22nd and 23rd of July 2009, with an average observation lasting 32.5 minutes (longest 82 minutes, shortest 5 minutes).

The 25 observations collected a large amount of baseline data, against the behavioural markers (Appendix D01). A great many recordings for behaviours concerning routine task activities were recorded (e.g. responding to communications, and managing flight strips); other behaviours such as those associated with team dynamics appeared far more infrequently, if at all (Appendix D01).

4.4.1 Behavioural Markers identified through observation (Source 5)

Through unstructured observation, 25 potential new behaviours were recorded during the observations (Appendix D01). There were several observations regarding body movement and posture. These include users adjusting their working environment to make it more comfortable, and adopting a relaxed laid-back posture. There were several additional behaviours associated with the movement, and management of strips. These behaviours ranged from tactile interaction and play with strips, through to their maintenance and disposal.

With particular regards to voice communications and instructions issued by the ATCO several potential behaviours were identified. These aspects of communications were associated with alert and attentive controlling and communication standards conformance, or indications of lower levels of workload.

During the observations, three potential new behaviours regarding teamwork were identified. Two of these behaviours (chatting socially during quiet periods, and general good team humour) suggest a team that has a positive team mood. The final behaviour, a reliance on a team member for information due to being busy with another activity, was considered.

4.5 DISCUSSION

4.5.1 Markers identified during the knowledge elicitation workshops

Many of the behaviours suggested by the various experts consulted in the knowledge elicitation workshops were not explicit overt behaviours; instead they were desirable or undesirable character traits. In addition several of the markers proposed would arguably be rarely seen as they are highly situation and environment specific. The workshop output generated more raw material for ATCOs not performing, than performing; which suggests that it is harder to articulate the visual appearance and behaviour of someone who is performing well and just getting on with the job.

4.5.2 Utility of the observation sheet

Unquestionably, the design and content of the first observational sheet significantly affected the amount of head-down time experienced by the author whilst observing. It took considerable time to locate the correct box to make a recording, due to layout and large amount of text. In order to compensate for this, the second variant of the observation sheet employed a short high level descriptor, and a separate more detailed explanation of the behaviour in another box. This aided utility however even with the changes, at times the process of recording seemed a distraction from the primary purpose of the study, that being to observe ATCO behaviour.

In terms of the use of a frequency tally, this method did provide fine granularity of data. However several behaviours occurred at an extremely high frequency rate (i.e. once every 30 seconds), as such these more continuous and high frequency behaviours are either inappropriate to record, or some method of cap to limit the number of recordings once attained may be beneficial.

Several observation sessions lasted for approximately one hour which proved tiring on behalf of the observer, and may have impacted the accuracy and quality of the observation towards the end of the time period. There were also occasions, due to operational staff movements, where an observation period was very short (less than 10 minutes).

4.5.3 Changes in behaviour observed

There were two very important and interesting findings made during these observations. The first finding was in regard to one particular ATCO's behaviour. This ATCO was observed on two occasions. The specific circumstances which made this ATCO's behaviour somewhat different to his/her peers was the fact they were a controller in the final stages of training, working towards sector validation (a process which can take several months of supervised live activity). The ATCO was working in live operations, whilst being monitored by an instructor seated to one side of the trainee (the author seated by the other shoulder).

Whilst observing the trainee, they demonstrated a noticeable lack of confidence and authority in his/her voice communications and general task delivery, which was of striking contrast to the other fully validated ATCOs observed. In terms of quantitative data substantiating the differences in behaviour observed, in several areas when examining the data recorded (Figure 4.1), certain behaviours observed against his/her peers was markedly lower (circa 20%).

However in the specific area of scanning strips the trainee was significantly higher in frequency (over 100%). It is conceivable that a trainee ATCO may lack confidence in certain aspects of the task which could lead to an increased scanning of the strips, and that this scanning compensated against a reduced capacity to interact and manage other elements of the task.

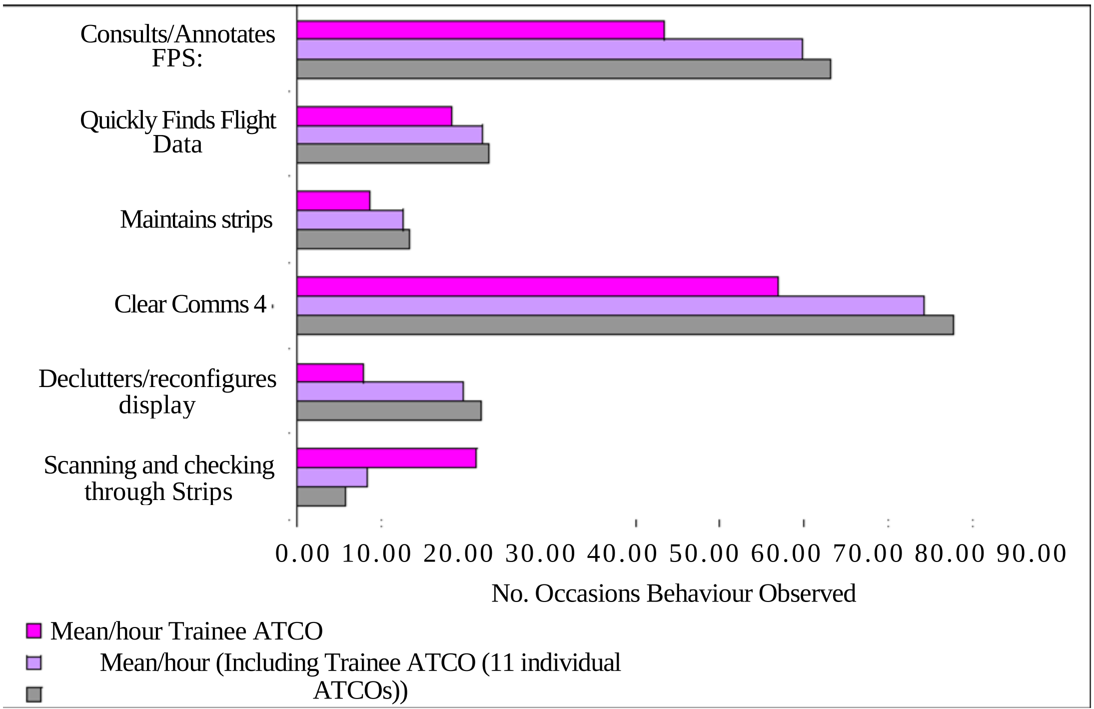


Figure 4.1 - Selected Sector Validating ATCO behaviours indicating large differences to his/her fully validated ATCO peers

In addition to the observations made within this study, as part of the author's project work delivered at NATS he had the opportunity to observe ATCOs within the simulation and training environment on other occasions. In this context, the ATCOs were learning to use an electronic flight strip system called *interim Future Area Controller Tool Set (iFACTS).* The ATCOs demonstrated classic learner behaviours (verbal confusion, unfamiliarity, frustration etc.), associated with the '*Cognitive Phase*' of Fitts & Posner's (1962) three phase learning theory (Patrick, 1992).

The second finding was as the result of observing a complex medical emergency enfolding, where an aircraft with a sick passenger on-board was given priority transfer back to London Heathrow, from where it had departed. During this situation, the mood of the ATCO team changed significantly, becoming sharper and highly alert. The ATCOs posture stiffened, and the casual atmosphere become very task centred and focused. This behaviour was very different to all other behaviour observed in the rather benign conditions which existed.

4.5.4 Study learning Points

- The inclusion of behavioural markers that were ‘continuous’ in nature was a mistake, and resulted in a tedious amount of regular recording. It may be possible to include certain regular behaviours; however the use of a cap should be considered to limit the number recorded.

- Lots of additional potential behavioural markers of NTS performance were identified, through the process of observation, Flin et al’s (2008) fifth source. Identification through observation affords greater levels of contextual relevance and richness than the review of documentation and the consultation of experts alone.

- The unstructured observations allowed a great deal of freedom and flexibility. This open approach was considered to provide real opportunity to approach the observational exercise with an open mind, and consider things from first principals.

- There were very few instances where a frequency recording was made against a negative behavioural marker. In those rare instances, additional comment was recorded on the observation sheet, in order to provide justification and detail as to what it was concerning.

4.6 DERIVING THREE FURTHER RESEARCH QUESTIONS

A significant finding of this preliminary study has been the display of learner or ‘novice type’ behaviours (e.g. under confidence, hesitance) by the ATCO receiving sector validation training. Several existing behavioural observation systems include behavioural markers of bad/poor NTS (Chapter 3, Table 3.31).

Some of the behaviours displayed by the ATCO receiving training provisionally suggest under-developed NTS proficiency. Another way to express this observation is that the ATCO in question has not fully developed certain skills, and is in fact at a different point of learning. Klampfer et al (2001, p13) states that the “*Application of the behavioural marker system must be sensitive to the stage of professional development of the individual*".

Existing behavioural marker systems focus upon ascertaining NTS at the end of training; with the observational focus upon behaviours demonstrating NTS competency. It is argued that there is potentially further value, if the behavioural marker system were to directly consider these differences in the stage of professional development of an individual, through the inclusion of markers indicative of different levels of learning and development.

In order to explore this interesting aspect further, two further research questions have been developed. These research questions are focused upon exploring what learning and development models are in the literature, and whether they afford insight into types of associated behaviour indicative of different phases of learning and development. Such insight could prove a valuable tool to help understand how ATCOs are engaging and developing new skills as they transition to further systemised operations.

4.6.1 Research Question 2

- What phases of behavioural development are there, including transient stages?

4.6.2 Research Question 3

- How might the presence or prevalence of certain non-technical behaviours be
  used to indicate how well a user is engaging and developing with a system?

A final research question also emerges as a consequence of this preliminary study. The overwhelming majority of observations made were during routine daily controlling activities. Permission was given to observe with the live operation, however the proviso from the Local Area Supervisor (LAS), was for the observations to be made only of persons who were not too busy. In a few instances however, the tempo of activity did increase during the time at which an observation was being made, and on one occasion a medical emergency necessitated a great deal of co-ordination and additional effort made on behalf of the ATC team. These situation factors clearly had an impact on the presence of certain behaviours. Task load and unusual circumstances are clearly two situation factors which are confounding variables within the observation of NTS within ATC, but are there any other factors which are significant?

4.6.3 Research Question 4

- What situational factors may impact the presence and prevalence of certain behaviours?

SUMMARY OF CHAPTER 4

35 behavioural markers indicative of both good and bad performance were identified from a number of sources. These were placed within an observational framework and used to collect baseline data for the frequency of specific behaviours observed in ATCOs engaged in normal day-to-day operations. This data would serve as a frequency baseline for which certain task behaviours occur during routine operations and enable a comparison when examining data of ATCO behaviour gathered in other situations (using electronic strips, training, adverse conditions, emergencies etc). Through this process, 23 potential additional markers were identified, and a number of changes were made to the observational marker set in order to improve utility, remove repetition, and poor observability.

The study was a very useful experience for the author, however many limitations of the observational method deleteriously impacted the quality of the observations undertaken. Most notably the limitations experienced extensively concern i) the amount of time spent looking at and navigating around the recording sheet (rather than observing the ATCO), ii) the repetition and monotony of an uncapped frequency count recording high frequency behaviour, iii) the lack of independent intrinsic value of a data set that essentially records ATCO behaviour in day-to-day operations.

This study concentrated on initial work undertaken to develop a practical tool to assess NTS performance within ATC. A key finding of this preliminary study is that a fundamental research question existed, and this has been the overwhelming focus of this thesis. The fundamental question amounts to what are the differences in NTS behaviour between a novice and expert ATCO, and can changes in these behaviours indicate developing competency. The subsequent chapters represent the main focus of this research, which has been to explore NTS behavioural change over periods of learning and development. During this research (Chapters 549) the primary focus has not been towards the development of a practical observation tool; although in order to undertake this more fundamental research a simple observational tool has been developed in support of the activity.

CHAPTER 05 – BEHAVIOURAL MARKERS TO DETERMINE THE STATE OF ATCO DEVELOPMENT

"Human beings, who are almost unique in having the ability to learn from the experience of others, are also remarkable for their apparent disinclination to do so."
(Douglas Adams)

In the previous chapter, observational research revealed differences in overt NTS behaviour between the fully valid ATCOs observed, and a trainee ATCO undergoing sector training. This chapter details further works undertaken in order to identify specific NTS behavioural markers reflective of differing levels of knowledge and proficiency. This was done in order to answer Research Questions 2 and 3; what phases of development are there, including transient stages? How might the presence or prevalence of certain non-technical behaviours be used to indicate how well a user is engaging and developing with a system?

5.1 AIMS OF THE STUDY

1. To observe and record behaviour from ATCOs with varying levels of system experience.
2. Review, identify, and select behaviours which are considered useful in evaluating NTS proficiency from the data recorded; using predefined assessment criteria.
3. Structure the resultant behavioural markers within an observation sheet and schema.

5.2 METHOD

A series of observation activities were undertaken using an open approach, involving detailed note taking to record behaviours of interest (Creswell, 2003). The observations were semi structured, with their purpose focused upon identifying behaviours which may indicate the state of the ATCO, and how proficient they were at engaging with the task. In addition, situation specific factors that may have led to specific behavioural responses were also recorded (e.g. equipment failure, levels of traffic, weather). The following research questions structured the observation process:

- What behaviours may indicate ‘ATCO state’?
- What reactions and interactions do ATCOs have with the flight strip system they are using?
- What working style behaviours do ATCOs demonstrate?
- What behaviours suggest the use of task strategies?
- What situation specific factors exist which generate specific behavioural responses?

These notes recorded potential indicators for ATCO state, reactions and interactions with the interface whilst engaging in the task, task strategies, quotes from the ATCOs, and any specific situation factors which existed to contextualise behaviours observed. These observations took place within the live operational environment and during real-time simulator based testing and training.

In addition to the observations taken, there was also on certain occasions an opportunity for limited questions at the end of the observation period to discuss some of the behaviours observed, and elements of the task and controlling sessions they had just experienced.

The observations were made by the observer (author) sitting by the shoulder of the ATCO who was seated in front of the workstation. The method of observation was similar to that used during the preliminary observational study (Chapter 4). The observer used a pen and paper to record notes, unfortunately photography, voice and video recording was not permitted.

Notes were made on a sheet with a single column used to record the time of specific behaviours and events; such as the ATCO working position (arrivals, departure, ground etc.) the time the observation started and ended, along with the location of the observation. The length of observation varied according to the length of the training or testing simulation run, or when a natural break period occurred in live operations, and the ATCO was replaced by another (an ATCO handover).

5.3 OBSERVATIONAL ACTIVITIES

In order to identify different behaviours manifested by ATCOs with differing levels of exposure and experience to new, electronic, flight strip systems, three separate observational activities were undertaken.

Observational Activity (Group 1):	This explored Tower ATCO behaviour in a group with a small amount (less than 3 hours) of simulator training exposure to an electronic replacement to their paper flight strip system (EFPS)
Observational Activity (Group 2):	This explored en-route ATCO behaviour in a group with a moderate amount (less than 25 hours) of simulator test and evaluation exposure to iFACTS, an electronic replacement of their paper flight strip system.
Observational Activity (Group 3):	This explored Tower ATCO behaviour in a group with a large amount (several years) of experience using an electronic replacement to their paper flight strip system (EFPS).

These observational activities were embedded within i) scheduled simulator training exercises, ii) simulator testing of the new system, or iii) in live operations, piggybacking this study upon a lessons learning activity undertaken by a NATS project in the process of introducing Electronic Flight Progress Strips (EFPS) to several NATS Tower operations[4]. As a result, there were a number of factors outside the control of the author which included timescales and amount of access

[4] This lesson learning activity was undertaken by the author, as part of the work to deliver EFPS to Edinburgh, Glasgow, and Aberdeen ATC towers.

to ATCOs. These limitations extended to the number of ATCO participants available to observe. On all occasions the maximum amount of data was collected within the permitted time-scales available.

5.4 RESULTS

In total, 45 individual observations were made across 28 individual ATCOs, with the length of each observation lasting between 30 and 50 minutes.

5.4.1 Observation of Group 1 (Novice ATCOs)

Two days of observation were undertaken on 1st and 2nd December 2009 at Edinburgh Tower during initial training for EFPS system (explained in Chapter 2). 14 Observations were undertaken across a total of 6 ATCOs, during several hours of simulator based training (each observation session lasting the length of the training run – circa 45 min). Observations were made during the training sessions for EFPS ground and air ATCOs positions; approach ATCO training activities had yet to be undertaken.

5.4.2 Observation of Group 2 (Intermediate ATCOs)

During the 17th to 19th August 2009, 11 individual observations were undertaken across a total of 4 ATCOs; during test and evaluation simulations for a replacement electronic flight strip and decision support system in the London Area Control Centre (LACC). This system (explained in Chapter 2) is known as interim Future Area Controls Tool Support (iFACTS). Each observation lasted the length of the simulation exercise (circa 50 minutes).

5.4.3 Observation of Group 3 (Expert ATCOs)

Over August and September 2009 20 individual observations were undertaken across a total of 18 ATCOs, at 4 NATS Tower units. Each of these observations lasted approximately 30 minutes. ATCO’s individual amount of experience using EFPS in live operations ranged from approximately 6 months – through to 5 years. In total, 20 observations were made.

Observations were undertaken at 4 London airports (Heathrow, Gatwick, Luton, and Stansted) who have replaced their paper flight strip systems with electronic flight progress strips. For the purposes of this study, observations were made of the ground ATCOs and air ATCOs. At Heathrow the air ATCO position was split across an arrivals and departures ATCO; observations were made of these two positions individually. Appendix D02 contains the behaviours captured through the observations, including an indication as to what the behaviours observed allude to. Table 5.1 provides additional detail on the number of ATCOs involved in the three observational activities.

ATCO Experience	Background
Novice	14 observations, 6 individual ATCOs, each observation circa 40 min. Edinburgh tower (during preliminary EFPS training).
Intermediate	11 observations, 4 ATCOs, each observation circa 50 min, during Test and Evaluation of a new Electronic Flight Data System (iFACTS). London Area Control Centre
Expert	20 observations, 18 ATCOs, approximately 30460 min per observation. At 4 ATC units: Gatwick (4), Heathrow (5), Luton (6), Stansted (3)

Table 5.1 - Detail of Various ATCO Observation Activities

5.4.4 Overview of the data collected

The 45 separate observations made across 28 individual ATCOs produced over 25 pages of observational notes. A wide variety of behaviours were exhibited, with many differences across the various environments. A preliminary filtering exercise to identify behaviours from these observational notes generated a list of 75 potential indicators of the state of the ATCO, and how proficient they were at engaging with the task. Not all the behaviours identified are unique to this study, and several of the markers are similar to those contained in other NTS observational systems.

Within this list of potential behaviours several were found to be mutually exclusive between the three ATCO groups, providing supportive evidence that the presence and prevalence of certain markers may change over time, depending on the level of system exposure and experience (Appendix D02). However several behaviours were observed within two of the groups (or even across all three ATCO groups). Within the novice group (those at early stages of training with EFPS), there were two who had worked at previous units. These two ATCOs had prior exposure to using EFPS which resulted in several behaviours indicative of system experience. Overall, many of the behaviours were displayed by all three ATCO groups. This suggests that ATCOs may exhibit a mixture of behaviours depending on levels of expertise, and acquired skills.

During the observations there was an extreme weather event which resulted in a lightning strike close to or possibly on top of the Gatwick control tower. The EFPS and other screens flickered, and the mood and ambience changed instantly as the ATCOs responded to the situation, performing system checks, and assessing the situation. However, for the overwhelming majority of observations, the environmental and task situation was benign, with ATCO behaviour relaxed, well-paced, and generally indicative of ATCOs very comfortable with the system and their performance using it.

Several behaviours were displayed in two or more of the three ATCO groups observed (Appendix D02). The lack of mutual exclusivity is considered to reflect the broad range of system exposure between the individual ATCOs within the three groups, and that certain behaviours may be those displayed by ATCOs irrespective of the system (and amount of system exposure) they have experienced. Prevalence of these behaviours was not recorded during the observational activities.

5.5 DISCUSSION

The observation process was liberating when contrasted with the research undertaken during the preliminary study (Chapter 4). The flexible method of observation permitted a large amount of ‘heads up’ time, with the flexibility to record contextual details using natural narrative. However, the observations made during preliminary study were an important foundation for this study.

This study experienced several limitations, which are primarily due the nature of field research. The number of persons observed, the scenario and environmental situation both in live ops and in simulation, and the amount of system exposure between ATCOs are all uncontrolled variables with this study. Conversely, the fidelity of the collected data is extremely strong as opposed to that collected during laboratory controlled conditions.

A significant limitation and uncontrolled variable within this study is the environment in which the various data collection activities have taken place. The three different environments and situations are considered to have impacted upon the behaviours displayed. For the novice group this was in the training environment with instructors and other simulation staff present and observing. The novice ATCO group had limited prior experience and demonstrated vulnerabilities and insecurities regarding their performance. For the Intermediate group the simulation exercises were toward the test and evaluation of the system, for most of the simulation time these ATCOs were comfortable exploring the strengths and weaknesses of the system, and determining how best to use it. The expert ATCO group were observed during live operations, in a situation where they had operated the ‘new’ system in live operations for a considerable (> 6 months) time.

5.6 DESIGNING THE MARKER SHEET

This section discusses the various activities undertaken in order to filter, consolidate and refine a set of behavioural markers, which reflect different levels of NTS proficiency within ATC.

5.6.1 Consolidation and Selection

Table 5.2 presents a number of rules, which were used in order to consolidate the large amount of data captured into the fewest number of mutually exclusive markers (Klampfer et al, 2001; Yule et al, 2006b). Through a process of consolidation and selection, 41 individual behavioural markers were developed.

Behavioural Marker Design Rules	Source & Explanation	
The skills must be applicable to an ATCO behaviour during Real Time Simulation or Live Operations		x
The observable behaviours must be specific, and well defined	x	x
The skills and behavioural markers should either be directly observable in the case of social skills or inferred from observing communication or other behaviours, in the case of the cognitive skills.	x	x
The system should be parsimonious and encompass the most important behaviours in the least number of categories and elements possible.		x
The categories and elements should have the maximum mutual exclusivity possible		x
The terminology used should be simple, and where possible use domain-specific language for ATCOs' behaviour	x	x
Observed behaviours do not have to be present in all situations, appropriateness depends on context.	x	
There is a suggested relationship to performance and level of ATCO expertise with the system	x	x

Table 5.2 - Rules used to extract behavioural markers from the observational data

5.6.2 Categorisation

The next step in the process was to assign categories to the 41 identified behavioural markers. A simple sorting exercise was undertaken in which the following six categories emerged, as the 41 markers were grouped into as few categories as possible:

1. Undertaking the task (task processes)
2. Attitude and Mood
3. Communications & Verbal Commentary
4. Physical Posture & Body Language
5. Interaction with others
6. Inputs & Interaction with the HMI and workstation

Tables 5.3 to 5.8 provide a condensed and structured list of these behavioural markers, they also indicate in which group these behaviours were observed, and provide detail on the value and benefit in observing each individual marker. In order to have an observation sheet with as little surrounding text as possible, each behavioural marker has been coded into a one or two word codec which is included in the tables.

There is an unequal distribution of the 41 markers across the six categories, with the largest number of markers concerning i) interaction with the system, and ii) verbal communications.

Undertaking the Task - Behaviour observed	Marker Codec
Demonstrates best practice (e.g. corrects incorrect RT, performs handover checks etc).	Best practice
Performs regular scans of the interface(s), searching and checking information	Interface Scan

Table 5.3 - Undertaking the task

Attitude & Mood - Behaviour observed	Marker Codec
Negative comments	Negative comments
Positive Comments	Positive comments
Apologetic of own performance	Apologetic
Talks socially in quieter periods	Social
Laughs with frustration	Frustrated

Table 5.4 - Attitude & Mood

Communications & Verbal Commentary - Behaviour observed	Marker Codec
(Verbal queries – why, what, where, when etc)	Confusion
Gets tongue tied in RT comms	Tongue-tied
Nods head, ok I understand, I'm getting there, talks to themselves, talks through next steps in task	Self Affirm
Er's um's, oh, um, alright um, pauses, delays, "standby", "say again, I missed that"	Delays & Repeats
Relaxed, calm	Cool/Calm
Frustrated, angry, irritated, edgy	Irritated
Confident, Decisive, self assured	Decisive
Swearing, huffing, Apologetic for mistakes & wrong actions	Verb. Frustration

Table 5.5 - Communications & Verbal Commentary

Physical Posture & Body Language - Behaviour observed	Marker Codec
Relaxes when quiet, sits back, crosses arms	Pace Fast
Sitting up alert and attentive	Pace Slow
Gestures of waving hands about, blowing air out	Phys. Frustration
Rubbing face, yawns, rubs eyes, "I'm tired", looks at watch	Fatigue
Adjusts MMI to the needs and requirements of the task	Adjust MMI
Taps pen, wriggles about, taps leg	Fidgets

Table 5.6 - Physical Posture & Body Language

Interaction with Others - Behaviour observed	Marker Codec
Looks for affirmation from colleagues or instructor, before an action	Affirm Before
Looks for affirmation from colleagues or instructor, after an action	Affirm After
Prompted by others 4 Oh right, yes, I see, ah ok	Reactionary
Displays the ability to undertake dichotic listening	Team Aware
Offers suggestions and discusses options with team members	Team Contribute
When busy, and when been put under pressure	Team Short/Snappy

Table 5.7 - Interaction with others

Input & Interaction with HMI & Workstation - Behaviour observed	Marker Codec
Demonstrates spatial/muscle memory of interface and layout functions	**Muscle memory**
Overconfident, fast, but makes mistakes and select wrong functions, performs unnecessary tasks	**Overconfident**
Slow and Hesitant, indecisive, unsure of actions, moving to control a function then moving back, requires multiple attempts to drive the HMI	**Slow/Hesitant**
Plays and experiments with system to see how it responds and behaves, and to rehearse actions	**Play/Sandpit**
Dual Tasking/Multi Tasking 4 whilst engaged in RT will also drive HMI and input data	**Dual Tasking**
Confident smooth flowing control. Deft command of HMI	**Confident control**
Using HMI incorrectly – wrong clicks, taps but no action on interface, can't find right function etc	**Incorrect actions**
Surprised by behaviour of HMI	**Surprise**
Quickly located required functionality and information when required	**Quickly locates**
Displays 'automaton' type actions when progressing through a sequence of steps	**Automatic**
Picks up activities (e.g. changes in HMI) on the periphery	**Periphery**
Plans task ahead 4 Opens up windows in advance, leaves cursor in the position needed for the next action or an action that they need to return back to, highlights all strips requiring QNH update prior to departure	**Plans Ahead**
keeps information windows active and open with applicable info during read back	**Keeps Info open**
Hovers pen/cursor over info on read back to confirm/tick off	**Input device tick off**
of looking out of the tower, checking arrivals and departures on radar, checking surface radar, checking the EFPS display(s), and other ancillary displays as appropriate (e.g. lighting panels, ATIS)	**Cyclic Scan**
when driving the HMI (e.g. tapping across taxi way windows in a 3, 2, 1, motion, moves the cursor round in circles especially on-screen objects, adjusts the windows so they line up absolutely perfectly	**Quirks**

Table 5.8 - Inputs and Interaction with HMI and Workstation

5.6.3 Classification into development classes

Having identified a series of behaviours, and categorised them in terms of types of behaviour, a further classification was warranted; namely organisation into categories of expertise. In order to do this, a review of the learning and development literature was undertaken.

Learning & Development frameworks Literature

A review of learning and development literature was conducted, in order to identify a suitable framework to classify and structure these different ATCO NTS behaviours. Although the literature contains several frameworks regarding teaching styles, and individual learning styles there are very few regarding distinct phases of learning. One reason to account for this may be that these phases can be considered instinctual and implicit, as is the case when we employ terms such as novice, and expert. In total, four learning and development frameworks have been identified (Table 5.9).

'Shu Ha Ri' is a concept taught within Japanese martial arts (Furuya, 1996). '*Shu'* governs the learning of conventional method and the basic forms. '*Ha'* is the

detachment from these conventional methods and indicates development of one’s own style. Finally ‘*Ri*^{0} is the transcendence phase of development where all forms and moves are natural and sub-conscious; although largely involving physical body movement, this framework elegantly describes three distinct phases of learning (and potentially different behaviours in each phase).

Another early framework is Fitts and Posner’s (1967) *three stages of learning* model. Once again this model suggests that an individual moves through sequential phases as they learn a new skill. The three stages to learning a new skill are as follows; the cognitive phase which involves building a mental comprehension through the study of the component parts of the skill. The Associative phase which manifests as the fluid and smooth development of the skill (achieved through repetition and feedback). The final phase is the Autonomous phase where the learned skill becomes so well developed that it becomes automatic and requires little or no conscious thought or attention whilst performing the skill.

Perhaps the most widely taught is the *'conscious competence*^{0} learning model. This four stage model is often attributed without reference to psychologist Abraham Maslow. However there is evidence to indicate that it was developed by Noel Burch, an employee of Gordon Training International (GTI) back in the 1970s; GTI, (2011). The first stage *'Unconscious Incompetence*^{0} is the expression of an individual who does not understand a skill deficit, and does not desire to improve it. *'Conscious Incompetence*^{0} the second stage assumes that the individual recognises a skill deficit, but does not know how to address it. The *'Conscious Competence*^{0} stage assumes an individual is aware of the correct actions to perform a skill however to demonstrate it requires conscious effort and concentration. The final stage *'Unconscious Competence*^{0} is the stage a person reaches where their experience and proficiency with a skill does not require significant concentration and may be performed easily.

Finally, the most recent development phase framework is Dreyfus’s 5 stage model (Dreyfus & Dreyfus, 1980). This model begins with ‘*novices*’ and a rigid adherence to core training rules, moving through ‘*advanced beginner*’ to ‘*competent*’ where the individual has a growing understanding of the cause and effect of their actions. The final stages of '*proficient*' and '*expert*' indicate where an individual obtains a global and holistic view of the situation, and have an intuitive and instinctive awareness and understanding.

Source	Learning and Development Phases				
Shu Ha Ri Furuya, (1996)	**Shu** • 'protect', 'obey' • traditional wisdom • learning fundamentals, techniques, heuristics, proverbs	**Ha** • 'detach', 'digress' • breaking with tradition • detachment from the illusions of self			
Fitts & Posner (1967)					
Burch (GTI, 2011)		**Conscious Incompetence** Though the individual does not understand or know how to do something, he or she does recognize the deficit, without yet addressing it	**Conscious Competence** The individual understands or knows how to do something. However, demonstrating the skill or knowledge requires a great deal of consciousness or concentration.	**Unconscious Competence** The individual has had so much practice with a skill that it becomes 'second nature' and can be performed easily (often without concentrating too deeply). He or she may or may not be able teach it to others, depending upon how and when it was learned.	
Dreyfus & Dreyfus (1980)	**Cognitive phase** Identification and development of the component parts of the skill 4 involves formation of a mental picture of the skill	**Associative phase** Linking the component parts into a smooth action 4 involves practicing the skill and using feedback to perfect the skill			
	Unconscious Incompetence The individual neither understands nor knows how to do something, nor recognizes the deficit, nor has a desire to address it.	**Ri** • "'eave', 'separate' • Transcendence • there are no techniques or proverbs, all moves are natural, becoming one with spirit **Autonomous phase** Developing the learned skill so that it becomes automatic 4 involves little or no conscious thought or attention whilst performing the skill 4 not all performers reach this stage			
	Novice • rigid to taught rules or plans • no exercise of 'discretionary judgment'	**Advanced beginner** • limited 'situational perception' • all aspects of work treated separately with equal importance	**Competent** • 'coping with crowdedness' (multiple activities, accumulation of information) • some perception of actions in relation to goals	**Proficient** • holistic view of situation • prioritizes importance of aspects • 'perceives deviations from the normal pattern' • employs	**Expert** • transcends reliance on rules, guidelines, and maxims • 'intuitive grasp of situations based on deep, tacit understanding'

		• deliberate planning • formulates routines	maxims for guidance, with meanings that adapt to the situation at hand	• has 'vision of what is possible' • uses 'analytical approaches' in new situations or in case of problems

Table 5.9 - Various Learning and Development Phases

There is a dichotomy of views in the content and application of the various learning and development frameworks which have been created. These frameworks are extremely theoretical, represent an inexact science, with particular weakness regarding the boundaries between each phase, and difficulty in segmenting skills and knowledge accordingly (Patrick, 1992).

5.6.4 The selection of a framework

Reviewing the existing learning and development frameworks, Fitts & Posner's (1967) three level framework fitted well against the three groups of ATCOs observed, and their differing levels of experience with the replacement flight progress strip system. However Fitts & Posner's (1967) framework is very cognitive process orientated, and uses '*psychological jargon*' not recommended by van Avermaete & Kruijsen (1998) for behavioural marker systems.

Given a number of mutually exclusive markers were presented across the three groups observed, it was decided to use a novel three level framework in which to classify the identified behaviours. This framework was based on Fitts & Postner's (1967) framework however it employs the three category terms of Beginner, Intermediate, and Expert with the following behaviourally orientated definitions:

- The Beginner stage – covers skill behaviours indicating the person is absorbing system knowledge, looking to acquire and develop task skills, and interaction skills, including supporting NTS.
- The Intermediate stage is an exploratory stage, where behaviours demonstrating developing skills are presented, and wider aspects of the systems in use are tested.
- The Expert stage concerns skills which are autonomous and instinctual, and highly practiced and skilled.

5.6.5 Classification of the behavioural markers

Each individual behavioural marker was classified according to Beginner, Intermediate, or Expert using the author's judgement and where possible the contextual location information of which ATCO group the behaviour was observed. Table 5.10 presents the outcome of this classification process.

Classified Behavioural Markers		
	Beginner Behaviours Intermed	
	• Negative comments	
	• Confusion	
	• Tongue-tied	
	• Delays/Repeats	
	• Overconfident	
	• Slow/Hesitant	
	• Incorrect actions	
	• Surprise	
	• Reactionary	
• Team Short/Snappy		
• Phys. Frustration		
• Serial Tasking		
• Verbal Frustration		
• Irritated		
• Affirm Before		
• Affirm After	Fidgets	

Table 5.10 - Behavioural Markers Classified into Learning Phases

5.6.6 Designing the marker sheet

With a number of suitable markers identified, the next step required was to contain them within a structured observation sheet; the purpose of the sheet being to allow an observer to record the frequency of occurrence of certain behaviours, whilst observing ATCOs at work. A number of design requirements developed from the preliminary study experience (Chapter 4) were used to inform the development of this observation sheet:

1. It should be a single page design
2. The design should be uncluttered and efficient, in order to minimise the amount of heads down time (and maximise the amount of observer heads up time)
3. It needs to be well structured so that it is easy to locate the correct marker
4. There should be space to write notes as required
5. There should be space to record additional behaviours

Following these design principles, an observation sheet has been produced. The observation sheet is a single page design, and is structured around the categories of behaviour identified through the observational study. There is space to record the frequency of occurrence (frequency tally) to the right of each short behavioural marker descriptor. In addition to the behavioural markers literature, and experiences with the design of the observation sheet during study 1, this observation sheet takes design cues from the other 19 Behavioural observation Systems reviewed (Chapter 3, Table 3.31-3.32). The ATCO behavioural observation sheet is supported with a 2-sided reference guide, which contains additional detail to aid the observer in identifying the correct behavioural marker. Examples of the observation sheet and reference guide are provided in Appendix C04 - C07, and Appendix C08 respectively.

5.6.7 Observation Sheet Versions

This thesis details the use of this observation sheet to collect data on 129 occasions, across four separate studies (Chapters 6 through 9). Over this timeframe, it has undergone iterative refinement in order to improve usability, to improve clarity and reduce ambiguity, and to eliminate markers with minimal to no presence or significant change in prevalence. The observation sheet which has been developed through the work detailed within this chapter, is version one (Appendix C04). In total four versions of the observation sheet have been produced (the history for each observation sheet and the associated markers is presented in Appendix C09).

SUMMARY OF CHAPTER 5

This chapter provides detail of an observational study undertaken to identify behavioural markers for future use in the assessment of Non-Technical Skill (NTS) proficiency of Air Traffic Controllers (ATCOs) engaging with their primary flight strip system. Using an open approach, a number of observations were made in the simulator and live operational environment; with ATCOs who had differing levels of exposure and experience with replacement electronic flight strips. A large number of notes were recorded including potential indicators of ATCO state, their reactions and interactions with the interface whilst engaging in the task, task strategies, quotes from the ATCOs, and any specific situation factors which existed to contextualise behaviours observed.

Through an iterative categorisation and classification process, a series of behavioural markers have been identified. A number of mutually exclusive behaviours were found across the three ATCO groups; several behaviours were observed across all three groups. Through the consolidation process 41 behaviours emerged, contained within six categories: Undertaking the task, Input and interaction with the HMI; Interaction with others; Physical Posture and Body Language; Attitude and Mood; Communications and Verbal Commentary. These markers have been structured using a three level development framework (Beginner, Intermediate and Expert).

CHAPTER 06 – CHANGES IN BEHAVIOUR OF QUALIFIED ATCOs

"If you can't measure it you can't improve it"
(Lord Kelvin)

In the previous chapter, a number of behavioural markers were identified for the purpose of determining the levels of Non-Technical Skill (NTS) development an ATCO has achieved. In this chapter a detailed account of a research study is presented which was undertaken to practically apply and evaluate this marker set and to examine changes in these behaviours over time. Specifically, this study examined changes in behaviour with a group of ATCOs in the process of learning to use Electronic Flight Progress Strips (EFPS). The chapter discusses the development of a method used to undertake the observations. Detail of the results is provided, followed by a discussion which explores the findings of the study.

6.1 AIMS OF THE STUDY

1) Explore changes in behaviour using the markers identified in Chapter 5.
2) Identify through observation any further potential markers, and refine the overall set as appropriate.

It is anticipated that the presence and prevalence of certain behaviours will vary according to the amount of system exposure the ATCOs experience with the new electronic flight progress strip system. Firstly behaviours within the beginner category are anticipated to reduce over time, as proficiency with the replacement flight strip system increases. Further into training, it is anticipated that behaviours from the intermediate group will become present, and gradually increase in prevalence. Finally, markers contained within the expert group are considered likely to manifest themselves towards the end of training and during live operational usage (post system implementation).

6.2 METHOD

As explained in Chapter 2, Flight strips are the primary source of information for ATCOs; therefore to change the modality of strips from paper to electronic medium represents a significant system change for users (Chapter 2). Working as part of the EFPS project team, the author visited Glasgow, Edinburgh, Aberdeen, and London City Towers to support their transition from paper to electronic strips. During this process, permission was granted to observe and record ATCOs undergoing electronic flight strip training. This situation has allowed a rare opportunity to observe users with a wide variation in exposure to this electronic flight strip system over a shortened time frame, compared to linearly tracking a single set of system users over an extended time period.

6.2.1 Data collection (Method of Observation)

Agreement to undertake the ATCO observation was obtained from the respective unit's operations manager, and each individual ATCO was asked for verbal consent to the observations. In all instances, consent was given. The observations were made by the observer (author) sitting close to the ATCO, to allow observation of the side of the face, and their interaction with their electronic strips. The observer used a pen and paper to record notes, and to note the occurrence of a specific behaviour using a tally count against the appropriate marker on each instance it was displayed.

A frequency limit of 5 instances for each marker within the observation period was employed for this study, in order to prevent excessive recording of frequent behaviours, whilst ensuring that their occurrence was still adequately captured. It is important to recognise that the introduction of a frequency cap introduces limitations. The removal of a maximum normalises the data, reducing the potential extremes in frequency that might exist. This normalisation can remove the subtleties of individuals' behavioural difference, which is of value when assessing performance across a group. These limitations must therefore be considered for any potential application of a frequency cap.

As per previous observations within NATS photographic voice and video recording was not permitted. Observations were made either during simulator training or during live operations (when the EFPS system had been implemented into service).

The observations undertaken during preliminary research (Chapter 4) and during the identification of potential markers (Chapter 5) varied greatly in length due to these observations being undertaken largely in the live operational environment. The length of observation was governed by the length of working period of the person observed. On reflection, this resulted in instances where the observation period was too long, resulting in observer fatigue. In view of this, a 30 min target length of observation was set. This was considered an optimal length (30 min providing sufficient time to observer a variety of task conditions and different behaviours, whilst preventing fatigue on the part of the observer).

6.2.2 Things to consider as an observer

Following experiences of undertaking earlier observations at various NATS locations, a number of considerations emerge:

The physical position selected for the observations should:

a. Provide a view of the side of the face, and body to monitor facial expression and body language.

b. Afford a view of the flight strip display (paper or electronic) in order to observe what the ATCO is interacting with, although it is unlikely that strips and other elements will be fully legible. The key is to be able to note hand and arm movements and gross interaction with the HMI.

c. Allow the ATCO’s communications to be listened to (both radio and face-to-face communications).

d. Where possible, through the use of available headphones, allow the communications which the ATCO receives via the radio to be listened to (i.e. communication with the flight deck, communication with ground services, and co-ordination with other ATCOs at different units).

When undertaking the observation:

a. Permission should be obtained from unit management prior to any study; in addition to individual ATCOs prior to observation.

b. The observer should be quiet and still in order to minimise themselves as a potential source of distraction.

c. When recording data, recording should be discrete, and covert, delaying recording for a few seconds may help to reduce association between ATCO's action and the recording of behaviour.

d. The ATCO may be interested in the observation sheet, and observations taken, and request to see what has been recorded. Therefore it is important that notes taken are discrete, and in a format which would cause minimal 'offence' if read by the person observed.

e. The observer should be mindful of team interaction, both to ensure that the ability to observe this is maximal, but also that their position minimises team interaction interference.

f. Finally as an observer it is important to be ready and prepared to move, be moved, or remove one's self from the observational situation so as not to interfere with training or live operations.

6.2.3 Schedule of data collection

A series of data collection activities captured observational data from ATCOs with differing levels of exposure to an electronic replacement of their existing paper flight progress strips. A total of 52 observations were undertaken across three NATS units, with a total of 32 individual controllers observed (London City 5 ATCOs, Edinburgh 17 ATCOs, Glasgow 10 ATCOs) Behaviours were recorded using versions 1 and 2 of the observation sheet (Appendix C04 and C05).

The observations were made in both the simulation training and live operational environments; following the training prior to implementation, and the subsequent live operational usage. The ATCO positions observed were both approach and tower operations (Table 6.1)

NATS Unit	Approach Radar Simulation	Tower Simulation	Approach Radar Live Operations	Tower Live Operations
Glasgow Airport		11		
Edinburgh Airport	17	6	5	6
London City Airport				7

Table 6.1 - The location and number of observations undertaken

The unit training for the EFPS system was delivered in a format so that the ATCOs would be at a similar level of experience over the training timeline. The format of the training was a minimum of 8 tower and 8 approach simulation sessions. Additional refresher training was available at the end of the course, in addition to the competency assessment simulation.

The practical limitations of access resulted in data which is somewhat clustered around certain phases in the implementation time line (Figure 6.1). The first cluster of data was collected during training, within the first 5 hours, and the last 5 hours of training. The first data collected post-implementation was from ATCOs with approximately 35 hours of system exposure (across both training and live usage). The final observations were made with ATCOs who had approximately 75 to 80 hours of total system exposure.

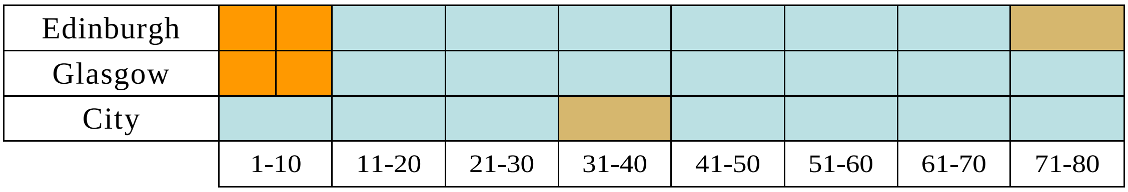


Figure 6.1 - Timeline of data collection activities (data collected in periods shaded orange)

6.2.4 Observation length

The aim was to achieve a 25-30 minutes observation period for each individual observation session. However for the observations made in the live operational environment there were individual instances where an ATCO went on a shift break or handover, and the observation had to be terminated early. The shortest live observation was 9 minutes, the longest 37 minutes, and the mean 27.5 minutes. Regarding the observations made during EFPS training, each session lasted approximately 50 minutes and included a 5 minute briefing and 5 minute debrief. The actual observations lasted around 25-35 minutes (depending how quickly the ATCO completed the core elements contained within the training session).

Across the three NATS units, there was considerable range of exposure practically using the electronic flight strip system (either in the simulation environment or live operations). This exposure ranged from 30 minutes, to approximately 80 hours. These values are based upon unit training records, individual ATCO's estimation; and for the observations made in live operations an estimate of hours was made based upon the unit's duty rota. Therefore the examination of differences with the presence and prevalence of behaviour has been undertaken across a group of users with 0.5 to 80 hours of system experience.

6.2.5 Time weighting

Given the variance in observation length, a time weighting was applied to the behavioural markers frequency data. Where the adjusted frequency (F_a) is sum of the recorded frequency (F) dividend by observation length (t), multiplied by a standard time constant (T); which in this instance is 30 minutes.

$$F_a = \left(\frac{F}{t}\right)T$$

A fundamental assumption of the time weighting formula is that the relationship between time and behaviour is linear. However it is possible that many variables (many of which have been considered by this research) may have a non-linear relationship between behaviour and time. Factors such as the physical and psychological state of the person observed, in addition to changing task and environmental conditions are potential candidates. With these examples the time of day, and the saccadic rhythm may have an influence to the distribution of behaviour displayed over time. In order to preserve the integrity of the data and minimise the impact of non-linearity, it is important to adhere to the desired observational duration. In addition, the time period observed should not be of such lengthy duration to introduce greater variably of behaviour across the timeframe in question. In such circumstances it can be assumed that this model is invalid.

6.3 RESULTS

The primary goal of this research study has been to establish how a number of non-technical behaviours change in presence and prevalence within a population of ATCOs learning to use an electronic replacement to their paper flight strip system. Table 6.2 provides a Spearman's Rho correlation of the normalized total number of instances beginner, intermediate, and expert behaviours were displayed calculated against the amount of system exposure (time). Figures 6.2 – 6.4 present the charts for this data.

As anticipated, the findings reveal reliable negative correlation for beginner behaviours (decreasing moderately over time). However, reliable weak negative correlations were also observed for intermediate and expert behaviours. This represents an unexpected outcome in the results.

Correlations calculated on 52 Observations				
Categorised Markers	**Spearman's Rho**	**Sig.**	**N**	**Means**
Beginner	-.560	.000*	52	29.3
Intermediate	-.375	.006*	52	13.7
Expert	-.292	.036*	52	13.1

Table 6.2 . Marker class correlated against system exposure (* significance at $p \leq .05$)

An examination of the descriptive statistics indicates that the total number of intermediate and expert markers observed in the series of observations decreased over time. With regards to the final 11 individual observations, these were made of ATCOs with approximately 80 hours of system exposure. When these final 11 observations were made (during live operations), the traffic and environmental situation was benign, with very little task activity occurring. As a consequence, any increase in system proficiency (and resulting intermediate and expert behaviours) was not observed; leading to a lack of change (behaviour correlated against time). To test this, the data has been considered as noise, and a second correlation run against the 41 observations collected (excluding the final series of live data), the results are presented in Table 6.3. In this second series of correlations, only the beginner group of behaviours revealed a moderate decrease over time (as expected), and the intermediate and expert behaviours were unchanged, with no reliable correlations.

Correlations calculated on 41 Observations				
Categorised Markers	**Spearman's Rho**	**Sig.**	**N**	**Means**
Beginner	4.494	.001*	41	30.8
Intermediate	4.013	.938	41	15.3
Expert	4.020	.900	41	14.0

Table 6.3 - Marker class correlated against system exposure for first 41 observations only (* significance at p≤.05)

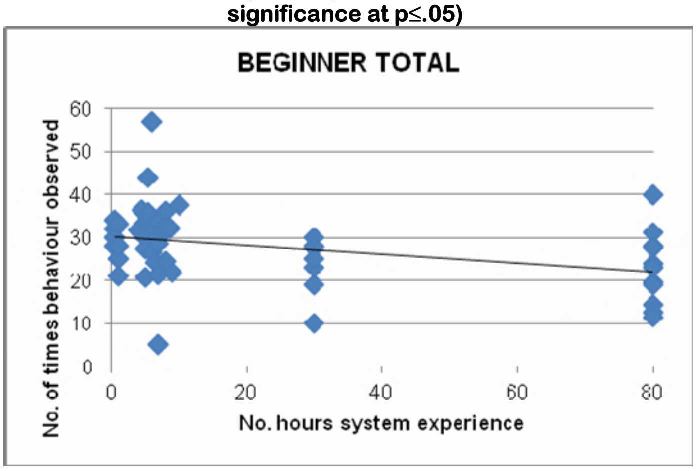
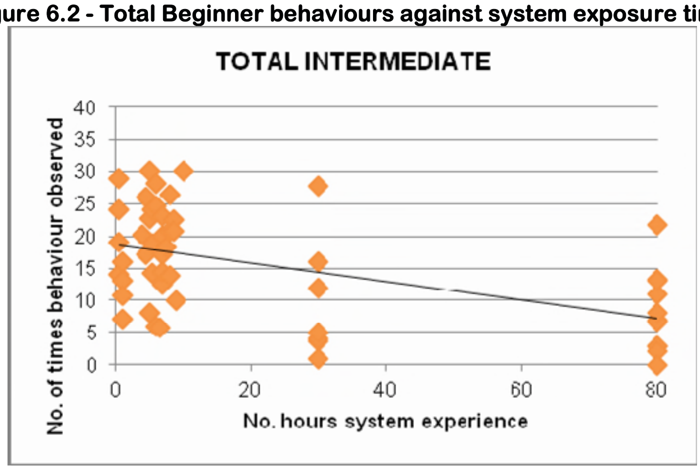


Figure 6.2 - Total Beginner behaviours against system exposure time



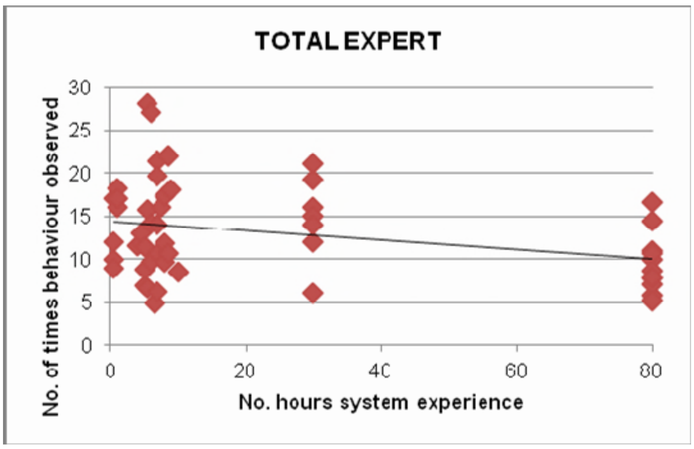


Figure 6.3 - Total Intermediate behaviours against system exposure time

Figure 6.4 - Total Expert behaviours against system exposure time

6.3.1 Individual Beginner Behaviours over time

The following sections present the results for individual behavioural markers, with Spearman's Rho correlations against the amount of system exposure time. Where a reliable correlation is demonstrated ($p \leq .05$), these are indicated by '*'.

There were a number of changes made to the set of markers across the 52 observations, which is reflected in the table. These modest changes were made half way through the study (from observation 24 onwards). Four new markers were identified:

Standbys –	Uses the technique to generate time to respond and react, possibly as a delaying tactic.
General Queries –	Rather than specific technical questions, the ATCO asks for general information regarding the implementation of EFPS.
Technical Discussion –	Engages in technical discussion regarding the design of the system, in order to understand its behaviour and build a strong mental model of the design.
Serial Tasking 4	This is the opposite marker to '*Dual Tasking*' indicating an ATCO performing tasks serially (communication with aircraft, then find and updating the appropriate flight data). This represents a lack of multi-tasking ability.

In addition to the inclusion of the four new markers, there was a consolidation of several markers from observation 24 onwards. '*Verbal Frustration*' and '*Irritated*' were combined, as was the case for '*Interface Scan*' and '*Cyclic Scan*'. Both these pairs of markers were combined due to similarity.

The marker, '*Best Practice*' was removed from the marker set from observation 24 onwards, and has not been included in these results. '*Best Practice*' was removed because the author was aware of some, but not all of the best practice techniques used by the ATCOs observed. Such a marker therefore was considered only observable by a technical task expert, and therefore beyond the scope of this work.

A timeline for each individual behavioural marker, indicating these and other changes made, is presented in Appendix C09.

6.3.2 Beginner behaviours

Table 6.4 presents the Spearman's Rho correlation of 17 individual beginner behaviours' frequency against system exposure time. Correlations for all 52 observations, as well as for the first 41 observations are presented.

It has been anticipated that these markers will decrease prevalence over time; as proficiency with their replacement Flight Progress System increases. Across the 52 observations, 12 out of the 17 beginner behaviours demonstrated a weak to strong negative correlation against system exposure time. When examining the individual correlations for the first 41 observations only, a total of six markers demonstrate a reliable weak to strong negative correlation. Individual charts are presented in Appendix D04.

Correlations calculated on 52 Observations			
Marker	**Spearman's Rho**	**Sig.**	**N**
Negative comments	.048	.733	52
Confusion	-.544	.000*	52
Tongue-tied	-.433	.001*	52
Delays/Repeats	-.327	.018*	52
Overconfident	.032	.823	52
Slow/Hesitant	-.320	.021*	52
Incorrect actions	-.395	.004*	52
Surprise	-.283	.042*	52
Reactionary	-.545	.000*	52
Irritated / Frustrated	.005	.971	52
Phys. Frustration	-.044	.759	52
Affirm Before	-.480	.000*	52
Affirm After	-.307	.027*	52
Team Short/Snappy	Not Observed		52
Standbys (from observation 26)	-.526	.003*	29
General Queries (from observation 26)	-.798	.000*	29
Tech Discussion (from observation 26)	-.451	.014*	29
Serial Tasking (from observation 26)	-.051	.793	29

Correlations calculated on 41 Observations			
Marker	**Spearman's Rho**	**Sig.**	**N**
Negative comments	.097	.547	41
Confusion	-.360	.021*	41
Tongue-tied	-.388	.012*	41
Delays/Repeats	-.335	.032*	41
Overconfident	.012	.939	41
Slow/Hesitant	-.283	.073	41
Incorrect actions	-.242	.128	41
Surprise	-.220	.168	41
Reactionary	-.380	.014*	41
Irritated	-.020	.903	41
Phys. Frustration	.034	.834	41
Affirm Before	-.258	.104	41
Affirm After	-.189	.238	41
Team Short/Snappy	Not Observed		41
Standbys	-.479	.044*	18
General Queries	-.877	.000*	18
Tech Discussion	-.363	.139	18
Serial Tasking	-.078	.759	18

Table 6.4 . Frequency of Beginner Behaviours correlated over system exposure time (*significance at p≤.05)

6.3.3 Intermediate behaviours

Table 6.5 presents the Spearman's Rho correlation frequency of individual intermediate behaviour frequency against system exposure time. It has been anticipated that these markers will increase in prevalence over time; as proficiency with the replacement flight progress system increases.

Across the 52 observations, only '*Pace Slow*' revealed a reliable positive correlation against system exposure time. Conversely, four out of the 17 behaviours reveal a weak to moderate negative correlation with system exposure time, across the 52 observations; '*Apologetic*', '*Pace Fast*', '*Play/Sandpit*', and '*Self affirm*'. When examining only the first 41 observations, '*Self affirm*' is the only marker to show a reliable negative correlation.

The negative correlations for these markers may indicate that their learning and development class should be within the beginner category. This topic is expanded in the discussion section. Individual charts are presented in Appendix D04.

Correlations calculated on 52 Observations			
Marker	**Spearman's Rho**	**Sig.**	**N**
Positive comments	4.181	.198	52
Apologetic	4.283	.042*	52
Social	4.067	.635	52
Self Affirm	4.638	.000*	52
Cool/Calm	4.062	.663	52
Decisive	4.084	.553	52
Pace Fast	4.475	.000*	52
Pace Slow	.394	.004*	52
Fatigue	4.095	.505	52
Adjust MMI	4.055	.697	52
Fidgets	4.043	.761	52
Muscle memory	.104	.464	52
Play/Sandpit	4.280	.044*	52
Periphery	4.161	.254	52
Input device tick off	4.235	.094	52
Team Aware	4.083	.560	52
Team Contribute	.219	.120	52

Correlations calculated on 41 Observations			
Marker	**Spearman's Rho**	**Sig.**	**N**
Positive comments	4.177	.267	41
Apologetic	4.187	.241	41
Social	4.036	.823	41
Self Affirm	4.568	.000*	41
Cool/Calm	.047	.770	41
Decisive	.155	.334	41
Pace Fast	4.132	.410	41
Pace Slow	.528	.000*	41
Fatigue	.023	.884	41
Adjust MMI	.044	.783	41
Fidgets	.182	.256	41
Muscle memory	.280	.076	41
Play/Sandpit	4.297	.059	41
Periphery	.128	.426	41
Input device tick off	.028	.861	41
Team Aware	4.134	.405	41
Team Contribute	.244	.124	41

Table 6.5 - Frequency of Intermediate Behaviours correlated over system exposure time (*significance at p≤.05)

6.3.4 Expert behaviours

Table 6.6 presents the Spearman's Rho correlation of individual expert behaviour frequency against system exposure time. It has been anticipated that these markers will increase in prevalence over time; as proficiency with the replacement flight progress system increases.

Across the 52 observations, no behaviours were found to have a positive correlation against system exposure time. Conversely, two markers ('*%ual tasking*', and '*Keeping info open*') reveal weak negative correlations against system exposure time. The remaining 6 markers reveal no change in prevalence over the

52 observations. When examining only the first 41 observations, no expert behaviours reveal reliable negative or positive correlations against system exposure time. Individual charts are presented in Appendix D04.

Correlations calculated on 52 Observations			
Marker	**Spearman's Rho**	**Sig.**	**N**
Dual Tasking	-.362	.008*	52
Confident control	-.024	.866	52
Quickly locates	-.136	.338	52
Automatic	.048	.735	52
Plans Ahead	-.076	.590	52
Keeps Info open	-.277	.047*	52
Cyclic Scan/Interface scan	-.222	.113	52
Quirks	.092	.516	52

Correlations calculated on 41 Observations			
Marker	**Spearman's Rho**	**Sig.**	**N**
Dual Tasking	-.206	.196	41
Confident control	.076	.635	41
Quickly locates	-.011	.946	41
Automatic	.276	.081	41
Plans Ahead	.106	.511	41
Keeps Info open	-.280	.076	41
Cyclic Scan/interface scan	.194	.225	41
Quirks	.041	.798	41

Table 6.6 - Frequency of Expert Behaviours correlated over system exposure time (*significance at $p \leq .05$)

6.4 DISCUSSION

The results are somewhat mixed. Although several beginner behaviours reliably revealed a decrease in prevalence over time, predicted changes to intermediate and expert behaviour did not manifest as anticipated. There are several likely explanations which might account for this:

- The first is that the intermediate and expert markers themselves are poor markers to be associated with learning and development phases; and the null-effect which has been observed is due to this. However these markers were identified through real-world observation, and displayed by ATCOs with differing levels of system experience. It seems unlikely that this is therefore the reason for a lack of anticipated behavioural change.

- The second is that the timeframe of this longitudinal study was insufficient to reveal the changes in behaviour anticipated. However a quarter way into the study, ATCOs had completed their EFPS training (with around 16-20 hours of EFPS exposure), and they were all assessed by Local Competency Examiners (LCEs). All ATCOs observed using EFPS during live operations have been perceived as competent in their EFPS proficiency by the LCE. It is unlikely in these circumstances that the timeframe of the study was too short to demonstrate positive behavioural changes with intermediate and expert NTS behavioural markers.

- The third concerns the fact that the ATCOs in this study were all experts with their existing paper flight strip system. The EFPS system is designed to provide a great deal of consistency and commonality, as a result it is conceivable that many of the ATCOs existing technical and non-technical skills transferred to EFPS. A lack of change in the intermediate and expert behavioural markers may reflect a level of skill transference at work. Chapter 8 details a study undertaken to explore the behavioural change of ab-initio (novice) trainee ATCOs, over the duration of an aerodrome course; with particular focus on the emergence of these expert NTS behaviours over time.

- The fourth is that the observations made within the live operational (post implementation) environment were frequently shorter than those made in the simulator. Perhaps there was not enough opportunity for behaviours to manifest in this shorter time period. However the data analysis has weighted the results to balance out the effect of varying observational length, it is unlikely therefore that this factor is responsible for the lack of behavioural change observed.

- The fifth concerns the method of analysis used to establish changes in relationship between frequency of behaviours observed, and time. The Spearman Rho correlation expresses a linear relationship, however many studies have revealed that non-linear relationships exist between skill acquisition and practice (Patrick, 1992). Examining the descriptive statistics, there does not appear to be any strongly non-linear relationships exhibited, however the clustered nature of the data limits any conclusions which may be reached. There is a clear limitation identified here, regarding this methodology; which is reliant on large amounts of evenly distributed data in order to study any subtle changes in behaviour which might be exhibited.

- The sixth concerns the frequency cap, and that it may not have provided sufficient range in order to identify subtle change in prevalence over time (the study in Chapter 8 raises the frequency cap from five to ten and evaluates the impact). Notably, however, as an observer, once the cap for a specific behaviour is reached it is permissible to 'lose interest' upon that behaviour, allowing greater concentration upon the other remaining behaviours. It was not considered that a particular pattern was made in completing the observations; however this is a potentially interesting area to explore in future studies.

- The final explanation concerns the somewhat benign conditions within the live environment during which the observations were made. There is clearly a relationship between task demand and the resultant behaviours that will be displayed. For example in a low workload, low traffic environment, the need to dual task is significantly reduced. With the observations in the live environment, permission from the units to observe was only given when the traffic and task was quiet. Indeed there was resistance from the unit for the author to observe when busy – which is understandable given the safety critical nature of the environment and task. It is suggested that the somewhat benign conditions within the live environment impacted the presence and prevalence of behaviours displayed. Chapter 9 provides an account of a small observational study which collects observational data within the demanding environment of an Olympics airspace simulation.

6.4.1 Phases of learning

Research question 2 looks to determine whether there exist separate distinct phases of behavioural change which provide indications of separate stages in learning and development. The results are inconclusive regarding this research goal. There does not appear to be clearly distinct phases – rather an individual will display abilities indicative of a beginner 4 and in the next moment those indicative of an expert. This may reflect the complex multi layered nature of learning, and that proficiency must be acquired across a number of cognitive and motor skills over time.

6.4.2 Evaluating ATCO proficiency

Research Question 3 concerns whether certain non-technical behaviours may be used to evaluate levels of ATCO proficiency with future ATC systems. A question which is useful to answer both for the design of new technology and the subsequent training support that would be provided. Based on the evidence collected during this study, several behaviours did change their prevalence over time, in the direction anticipated. However, the vast majority of the behaviours which exhibited these changes were from the beginner category only. Notwithstanding this fact, an initial conclusion which may be draw from this work is that the absence of these largely beginner behaviours may be an indirect indicator of system proficiency.

6.4.3 Review of individual markers

Four markers assigned to the intermediate category demonstrated moderate negative correlations against exposure time. The first three of these markers concern behaviours which demonstrate a degree of understanding, but reflect a desire to test out this knowledge. The markers '*Self Affirm'*, '*Play/Sandpit*', and 'Apologetic' ultimately are anticipated to reduce following the exploratory learning phase. Although these behaviours may be transient, they may better be assigned to the beginner category. The fourth marker demonstrating a moderate negative correlation is that of '*Pace Fast*'. Conversely, '*Pace Slow*' revealed a moderate positive correlation. Collectively these two markers express the ability of an ATCO to modify their approach to the task (e.g. make adjustments to their posture and interaction style according to lowered task demands), in response to changing demand. A combined marker '*Adapts to the pace of the task*' should be considered.

The behaviour 'team short/snappy' was not observed during this study, as a result its merits cannot be fully considered. A marker which occurs only rarely has limited utility in monitoring subtle change in frequency over time.

Four new potentially valuable beginner behavioural markers were identified within the first 25 observations in this study ('*Standbys*', '*General Queries*', '*Technical Discussion*', '*Serial Tasking*'). The first three of these new markers show reliable negative correlations across the 29 observations in which they were taken. However all four markers have been added to the marker set, for application and evaluation in other settings (Appendix D04 and C09)

6.4.4 Observation method

The observation sheet was easy to use, and the reference sheet was valuable. The 'less is more' principle of sheet design did help ensure maximum 'heads up' observation time. Having space to note down specific things observed was also a very useful aid (not least to record additional potential behaviours for future inclusion).

The target observation duration of 30 minutes seemed like a reasonable length of time to provide opportunity for the fullest range of behaviours to manifest. It did not seem too short, resulting in a number of behaviours yet to be observed or counted on very few occasions. Conversely it did not seem too long with behaviours still being recorded at the end of the time period. It was also considered an acceptable length on behalf of the observer (the author) without suffering 'observation fatigue'.

There was substantial contextual information to record when observing in the operational environment such as the ambient conditions in the tower/control room, the level of traffic, equipment outages, number of staff on the roster etc. This information has provided insight as to why certain behaviours were or were not displayed, and are of benefit to capture as part of the observation process.

SUMMARY OF CHAPTER 6

This chapter details a study undertaken in order to record potential changes in ATCO behaviour, during the process of transitioning from paper to electronic flight progress strips. A number of markers identified through the work covered by Chapter 5 were used to record the frequency data for these behaviours. A mixed set of results was collected. 12 out of 17 beginner behaviours revealed reliable negative correlations; however the results for intermediate and expert behaviours demonstrated very little significant change in behaviour over time. Several new markers were added over the course of the observations, and further revisions concerning the re-classification of four intermediate behaviours as beginner behaviours are discussed. A number of arguments are put forward as to the lack of change regarding the intermediate and expert behavioural markers. The two considered most likely i) skill transference, and ii) the benign conditions are explored within Chapters 8 and 9. In the next chapter, an inter-rater reliability study is presented, which explores the design and content of the behavioural observation sheet, in addition to the method used to record changing prevalence of behaviour over time.

CHAPTER 07 – INTER-RATER VALIDATION OF BEHAVIOURAL MARKERS AND OBSERVATION METHOD

"Judge your success by what you had to give up in order to get it."
(His Holiness the 14th Dali Lama).

An important activity in the design and testing of behavioural marker systems is the assessment of inter-rater reliability. This chapter provides details of a series of dual observations undertaken to assess the inter-rater reliability of the behavioural markers identified through observations in Chapter 5, and used during the study reported in Chapter 6. Although this study actually took place during the study that is in Chapter 6, it has been presented as a stand-alone chapter in order to preserve the iterative process of development undertaken (and reflected across Chapters 5-9). This chapter is structured into four parts. The first part contains details regarding the dual observation method used. Parts two and three are sections outlining the qualitative results, and the quantitative results. Finally part four provides a discussion section which explores the combined quantitative and qualitative results, including implication for changes to the behavioural marker set as an outcome of the qualitative and quantitative results.

7.1 AIMS OF THE STUDY

1) To test the inter rater reliability of the behavioural marker set.
2) Review the observational method through post-observation interview.
3) Make evidence based revisions to the marker set following qualitative and quantitative reliability and usability results.

7.2 METHOD

In order to determine the reliability of the behavioural markers identified through earlier observational research (Chapter 5), and evaluate construct validity, a number of dual observations were undertaken. The dual observations involved both the author and a second Human Factors (HF) observer watching an ATCO at the same time, whilst they performed their controlling task using Electronic Flight Progress Strips (EFPS). Two HF experts participated in this study as observers. One of the HF experts was engaged with the implementation of EFPS at London City Airport, the other with the implementation of EFPS at Aberdeen Airport.

The HF observers were recruited ahead of the planned dual observations. Agreement to undertake the HF observation of the ATCOs was obtained from the operations manager at each unit (Glasgow, London City) – and permission to observe each individual ATCO was verbally obtained prior to the commencement of each dual observation.

7.2.1 Observer Training

The HF observers received individual one-on-one training prior to undertaking the dual observations. Training for each HF observer was conducted by the author over a 90 minute period, where the following syllabus was covered:

- Background on behavioural markers theory and design
- Stages of learning and development
- The purpose of the observation activity
- The design of the observation sheet
- The observation method
- How to record data using the observation sheet
- How to use the supporting information sheet
- Tips and recommendations for successful

observation **Instructions to Observers**

The HF observers were asked to record a frequency tally each time a specific behaviour was exhibited by the ATCO observed using the observation sheet provided (Appendix C06). A frequency cap was employed, whereby there was a maximum limit of five observations to be recorded against an individual behaviour (See Chapter 6). To support this activity, the HF observers were provided with the double sided reference sheet that contained additional detail on each individual marker (Appendix C08).

Observers were encouraged to note down any particular comments queries or other observations surrounding the markers themselves. Observers were asked to consider the merits of each marker as they made their observation and note down any comments regarding their individual utility. Observers were asked to note down any situational elements (environment, task etc) that may have had an influence on events. Finally the observers were asked to note down any addition behaviours they may have spotted which provide insight as to the relationship and level of user development they have achieved with the EFPS system.

7.3 RESULTS

Five dual observations were undertaken with the author and a HF observer at London City. The dual observations at London City were made with ATCOs using EFPS during live operations. The EFPS system had been in service for approximately two months at this point. Each dual observation lasted approximately 30 minutes. Four separate ATCOs were observed on one occasion each; with one ATCO observed on two occasions.

Unfortunately due to work constraints a more limited series of dual observations was undertaken at Glasgow airport. Two dual observations were undertaken during EFPS training in the simulator by the author and the HF observer. Two individual controllers were observed, one for the standard 30 minute observation period, the other for a little over 10 minutes. Unlike the data captured in the observational study presented in Chapter 6, no time weighting has been applied to the data, as it was not required for the purposes of inter-rater reliability analysis.

Qualitative feedback from the two HF observers was obtained through a post observation interview, in order to i) discuss the observations made, ii) the design and use of the observational sheet, iii) the observational method used, iv) finally each observer was asked for their view on the strength and utility of each marker contained within the behavioural marker set.

7.4 QUALITATIVE RESULTS (INTERVIEW)

Following the dual observations, a separate interview was undertaken with each NATS HF observer. The purpose of these interviews was to i) critically review the effectiveness of the observational method, and to ii) gain broader feedback on the overall experience, and evaluate construct validity of the behavioural marker set used.

A structured interview approach was taken (with comments recorded on voice recorder for later analysis).The interview was structured into three parts, the first involved discussion on the method of observation, the second involved discussion on the design of the marker sheet, the third concerned the utility of each marker and potential amendments which could be made.

The next section provides the results of the information collected within these areas of interview. The full interview transcripts are presented in Appendix D05. Overall, both interviewees found the dual observation exercises "*rewarding*", and found it "*insightful*"' sitting and watching the ATCOs work for an extended period of time.

7.4.1 Method of observation

The HF observers were asked if they found the observation process useful. One HF observer questioned in the interview if the observations were measuring the "*impact of the system change, or people's individual's behaviours and individual difference?*" The second HF observer was cautious about "*over-drawing conclusions*" from the observations made, and found the observation process "*resource heavy*", they did however find it useful in "*gaining insight*" as to how ATCOs were performing with the new EFPS system during training, and that certain behaviours exhibited "*helped back up questionnaire feedback*" received during system validation.

The second HF observer commented that "*they were not fully confident using the behavioural observation system*", and "*identifying correct behaviours from the categories*" provided. They acknowledged that "*further exposure and experience* using the system *would likely improve [their] understanding of the technique*." As one HF observer stated "*some behaviours were harder to observe than others (the behaviours were subtle, and discrete)"*.

Both HF observers were happy with the length of observation period, and "*felt like it was the right length of time"*; although one HF observer did make the comment that they "*would not like to make too many observations in one day"* as it would lead to fatigue.

One comment which was made concerned whether the 30 minute observation period was "*sufficient in order for certain behaviours to manifest themselves?"*; for example whether fatigue could "*manifest itself by an ATCO who had just begun their shift"* (and therefore with less than half an hour of working completed). The HF observer noted that it is important that guidance material provided to observers makes reference to the "*limitations of a single observation period"*, and that not all behaviours contained within the marker system "*may manifest themselves"*.

Another comment made by a HF observer concerned the impact of at what time of the day (or night) the observations themselves should take place. The observer felt it important that the timing of observations, where possible, were scheduled to "*accommodate potential differences across watches"*.

7.4.2 Design of the marker sheet

Both HF observers stated that they found making notes important for many reasons. The first being that it "*allowed queries"* regarding the *"definition of certain markers"* to be discussed following the observations. Secondly the notes were made in order to record the "*context of a particular behaviour"* (how and when it was displayed). Thirdly notes were made for the inclusion of "*new markers, or the re-design of specific individual markers"* (e.g. to combine or split certain markers).

Both HF observers liked the fact that the sheet was not too complicated and was "*flexible"* enough to allow notes and notations to be made on the sheet; therefore it was suggested that "*no redesign"* of the observation sheet was needed, specifically for the purposes of note taking. A final point with regards to note taking, one HF observer found it useful to "*scribble notes on the back"* of the observation sheet at the end of the observation; therefore the fact that the back side of the observation sheet was blank proved a useful additional resource.

The first HF observer was very positive regarding making an overall judgement. They felt that they were "*able to do so"*, and that it "*would provide a useful summary of the observed person's behaviour"*. The second HF observer was somewhat more reticent, concerned that the ratings would be interpreted as a direct measure of task performance; which they "*did not feel qualified to provide"*.

The two HF observers were largely satisfied with the specific contextual data already captured by the design of the sheet. One of the HF observers did consider that the ATC watch was of importance to record. Both HF observers considered the recording of environmental and situational conditions very important, appreciating the potential impact these factors may play upon controller behaviour. One HF observer questioned whether the presence of the HF observers themselves might have *"impacted upon controller behaviour"*. This observer then went on to postulate that ATCOs are in-fact regularly observed whilst working, and that they are largely unaffected by the presence of an observer.

Overall, the location of markers on the sheet was felt to be fine, One HF observer questioned the balance of markers across the six categories; in particular that the "*HMI interaction category contains many more markers"*. This raised questions concerning "*potential duplicates?"* The observer acknowledged that the imbalance of markers is "*potentially due to the limitations of what can actually be observed"*. One HF observer felt that familiarity would develop with the layout of the behavioural marker observation sheet, aiding utility of use. The two HF observers also reported that they frequently used the two page reference sheet, using it to aid clarity in appropriate marker selection.

7.4.3 Behavioural markers and categories

Both HF observers commented that the categories worked quite well. Each marker was reviewed individually and a great number received specific comment (Appendix D07). Their comments ranged from queries and clarifications of definitions through to recommendations to delete a marker, or group two or more markers into a single category. No instances during the discussion occurred, concerning the splitting of a single marker into more than one.

Neither HF observer stated that they found this process of delineation difficult. Both referred to the reference sheet when uncertain as to the definition of a particular marker and found this sheet to be useful. However in the later stages of discussion, where each marker is individually considered for its clarity and merit, there were instances where observers found delineation between markers difficult. The comments which were made during this discussion suggested they experienced moments of uncertainty using the marker sheet (and determining definitions), and that the observers felt less confident about the process and the observations that they were making due to difficulties in delineating between certain markers.

Neither HF observer expressed strong views as to whether the five count limit was acceptable, but did comment that it seemed fine. One HF observer felt a limit was appropriate to prevent *"unintentionally tunnelling upon a single behaviour"* and focusing upon it to the detriment of observing other behaviour (a finding experienced with continuous behaviours observed in Chapter 4). Additionally they commented that the frequency is *"not the most important element"* in the observational process, but rather that the behaviour is *"identified and recorded"*. This question did lead onto discussion with a HF observer concerning whether they found any specific pattern for observations (e.g. focusing on a certain number of behaviours before moving onto others). They commented that they did feel there was a pattern to their observation, though they could not *"explicitly state what"* that pattern was. A similar discussion was undertaken with the second HF observer who also found their focus would *"move on from certain markers once they had been observed"*.

There was only one instance where a HF observer noted a new behaviour not previously identified; it concerned a form of body language which conveys expertise and great awareness in the task. An ATCO was observed to sit back in their chair when not interacting with EFPS, but was able to provide RT to aircraft (i.e. aircraft read backs) without interrogating the EFPS screen. This suggested the ATCO had a strong picture as to the traffic situation, and the aircraft operating around the airport. The HF observer felt this contrasted significantly with another controller who was constantly leaning forward toward the EFPS screen (and generally looked under confident). The HF observer wondered if this *'laid back'* working style was reflective of ATCOs who have been controlling for many years; and indicated a high level of spatial and cognitive skill. There is however potential ambiguity and misinterpretation concerning this marker, given that it concerns aspects of internal cognitive processes which must be judged, it has therefore not been taken forward for inclusion as a behavioural marker.

7.4.4 Qualitative marker review

The final part of the observer interview involved the critical discussion and review of each behavioural marker on an individual basis. Several makers received significant comment, this included clarifications regarding definition or discrimination from other markers, and positive and negative comment where warranted. Appendix D05 presents the interview transcripts in full.

Tables 7.1 – 7.6 presents a summary of the comments received for each individual marker which was commented upon during the interviews. Not all markers received comment as the observers on occasion had nothing positive or negative to say about them. The comments received are categorised into different themes, the first of which (Table 7.1) being markers that received positive comments (easy to spot, saw on several occasions, useful and informative).

Markers which received favourable comment	
Irritated	Cool/calm
Fatigue	Negative comments
Confusion	listening to RT communications
Social	

Table 7.1 . Markers which received favourable comment

Table 7.2 presents a number of pairs of markers which received comments that they were difficult to differentiate from one another, and that they either require additional clarification within the guidance notes, or that they could be combined together (increasing mutual exclusivity).

Pairs of markers which were difficult to differentiate from one another	
· Play/sandpit · planning ahead	· Team aware · Team contribute
· Confident control · Cool/calm (verbal)	· Verbal Frustration · Irritated
· Quickly locates · Muscle memory	· Decisive · Cool calm
· Confident control · Automatic	

Table 7.2 - Markers which were difficult to differentiate

A number of markers were considered to require additional clarification, due to i) ambiguity or uncertainty, ii) the scope of coverage, iii) interpretation which is required, or iv) the implication that the presence of a marker may be interpreted to mean different things, depending on the situation at hand (Table 7.3).

Markers considered in need of additional clarification:	
· Team Contribute	· Cyclic scan
· Confusion	· Planning Ahead
· Physical Frustration	· Standbys
· Fidget	

Table 7.3 - Markers in need of clarification

The markers presented in Table 7.4 were considered by the HF observers to be applicable only within the confines of the training environment. These behaviours relate to the ATCOs reaction and response to the task, through their interaction with colleagues, training and support staff. It is conceivable that other behaviours are also situation or scenario dependent which highlights the importance of recording the conditions and environment in which the observations take place.

Markers were highlighted to be only applicable to the training environment:	
• Affirm before	• Reactionary
• Affirm After	• General queries

Table 7.4 - Markers for the training environment only

Table 7.5 presents markers which the observers found difficult to spot. This was due to the behaviour being too subtle, or that it was not seen frequently during the observations. With regards to the markers being difficult to spot, the reason sighted to explain these comments concerns the distance away from the person being observed. This is a limitation where two observers are observing at the same time as in order to not overcrowd the ATCO, the HF observers naturally are positioned slightly further away than if it were a single person making the observations.

Markers considered difficult to spot, or even not spotted (behaviour too subtle, un-observable, or not observed)	
· Play/sandpit	· Apologetic
· Periphery	· Pace slow
· Quirks	· Adjust MMI
· Input device tick off	

Table 7.5 - Markers hard to see, or unseen

Finally, Table 7.6 presents a number of markers the HF observers considered of low value or insight as they are behaviours demonstrated on a regular basis, and their presence is ubiquitous with valid ATCOs.

Markers considered to have limited value, and limited utility	
Pace fast	Decisive
Cool/calm	Dual tasking

Table 7.6 - Low value markers

Following qualitative (and quantitative) feedback, a number of changes were made to the behavioural marker set. Further detail of these changes is provided later in the chapter. Appendix C09 illustrates changes made to the marker set as a result of this study.

7.5 QUANTITATIVE RESULTS

This section summarises the quantitative analysis of the dual HF observations taken in order to evaluate observer agreement, and levels of inter-rater reliability. In total seven dual observations were undertaken, five with the first HF observer and two with the second. This small amount of data does unfortunately limit the depth of analysis permissible, and the strength of conclusions which may be reached. However, the data has high real world validity; and presents a limited opportunity to explore this area; it has therefore been included in this work with those caveats in mind.

7.5.1 Inter Rater Reliability – Overall Marker Set

Inter-Rater Reliability (IRR) has been calculated for each pair of dual HF observations. Table 7.7 presents the Spearman's Rho co-efficient for the seven paired observations. With the exception of observation Pair E whose correlation was not found to be significant, all other inter rater correlations are shown to be reliable. Interpretation of the results indicates weak to moderate agreement (0.460 – 0.680) for these remaining observations. Table 7.8 presents the guidelines used to determine levels of strength for these correlations (LeBreton et al, 2003).

Spearman's Rho Coefficient	Inter-Rater Dual HF Observations							Mean of Correlation Co-efficient
	Pair A	Pair B	Pair C	Pair D	Pair E	Pair F	Pair G	
	.680	.460	.662	.491	.674	.219	.484	0.501
Significance ($p \leq .05$)	.000*	.002*	.000*	.001*	.000*	.154	.001*	
N	44/44	44/44	44/44	44/44	44/44	44/44	44/44	

Table 7.7 - Inter-Rater Reliability

Dual observations with the first observer (Pairs A-E) demonstrate weak to moderate levels of agreement. The dual observations with the second HF observer (Pairs F-G) reveal a lack of agreement with the first observation. The lack of agreement may be a result of a learning effect, with the HF observer uncertain in their first observational experience using the marker checklist. Unfortunately the training did not include an opportunity to undertake practice observations ahead of this activity. Such practice observations are intended to remove learning effects and other reliability issues associated with an observer's earliest exposure to a behavioural marker system (Mitchell et al, 2012).

Level of Inter-rater Agreement	Substantive Interpretation
0 to 0.30	Lack of agreement
0.31 to 0.50	Weak agreement
0.51 to 0.70	Moderate agreement
0.71 to 0.90	Strong agreement
0.91 to 1	Very strong agreement

Table 7.8 - Interpretation of Inter-rater agreement (LeBreton et al, 2003)

7.5.2 Inter Rater Reliability - Individual Markers

The small number of dual HF observations undertaken fundamentally limits complex inter-rater data reliability assessment. However notwithstanding the limitations of the data, and depth of analysis that may be undertaken, there are valuable insights that may be explored through simpler analysis.

A simple analysis has been undertaken following principles from Gatfield's (2008) inter-rater reliability assessment of maritime crisis management behavioural markers. Each of the seven pairs of observations was examined and the difference between the first observer (the researcher) and the second observer was calculated. A threshold was set prior to the data review to interpret the results of this analysis (Table 8.3). The total set of results is presented in Appendix D06. In simple terms, this analysis permits the review of individual markers in order to determine levels of convergence (Lawler, 1967).

Using this scoring threshold (Table 7.3), out of the 44 markers contained within the study, 11 markers receive a low level of agreement between observers, 14 markers receive a moderate level of agreement, and 19 markers receive a high level of agreement (Table 7.10).

It is important to state that a high inter-rater correlation in itself is not necessarily an indication of a strong marker. This is because the marker may be difficult to observe, and both observers may have found it difficult to use the marker; therefore demonstrating high levels of agreement with a weak marker. Neither *overconfident* or *team short/snappy* were recorded by either observer, during any of the observations (as denoted by '*' in Table 7.10).

In the next section, the quantitative results have been contrasted with the qualitative results, in order to produce a comprehensive understanding of the value of each individual behavioural marker.

Levels of agreement	Difference in scores
High Agreement	· No difference in scores between pairs of observations · A difference in scores of only 1 between pairs of observations
Moderate Agreement	· A difference in scores of only 2 between pairs of observations
Low Agreement	· A difference in scores of 3 or more between pairs of observations

Table 7.9 - Score threshold used to determine levels of agreement between observers

High Agreement	Moderate Agreement	Low Agreement
· Team Short/Snappy* · Team Contribute · Reactionary · Affirm After · Affirm Before · Quirks · Periphery · Serial Tasking · Play/Sandpit · Overconfident* · Fatigue · Pace Fast · Irritated · Standbys · Self Affirm · Tongue Tied · Social · Positive comments · Negative comments	· Tech Discussion · Team Aware · General Queries · Keeps Info Open · Quickly Locates · Surprise · Incorrect Actions · Confident Control · Fidgets · Adjust MMI · Physical Frustration · Delays · Confusion · Apologetic	· Cyclic Scan · Input Device Tick off · Plans Ahead · Automatic · Dual Tasking · Slow Hesitant · Muscle Memory · Pace Slow · Verbal Frustration · Decisive · Cool/Calm

Table 7.10 - Distribution of Behavioural Markers according to inter-rater agreement

7.6 COMBINED QUALITATIVE AND QUANTITATIVE REVIEW OF MARKERS

Both qualitative and quantitative assessment has been undertaken in order to determine levels of inter-rater agreement regarding individual markers, and general

HF observer opinion on the method. The results of the qualitative and quantitative data have been combined, in order to determine what changes to the marker set may be warranted. This combined analysis reveals a high degree of linkage between the qualitative and quantitative results. The more favourable and positive the comments were for individual markers, the better the level of quantitative inter-rater agreement shown. The complete detail of which is presented in Appendix D07.

As a result of this analysis, a number of changes were made to the behavioural marker set. These changes were made in order to provide clarity on individual markers and their definitions. Appendix C09 presents every change made to the marker set as a result of the qualitative and quantitative analysis, however a summary of these changes is presented here:

- Of the 19 high agreement markers, six have had clarifications made to the observer notes, one marker has been combined with another in order to improve mutual exclusivity, and 12 unchanged from this review.
- With regards to the 14 moderate agreement markers, four have had clarifications made to the observer guidance notes; two have been combined together with another marker contained within the low agreement category, leaving the remaining eight markers unchanged from this review.
- Finally, of the 11 low agreement markers, two have had clarifications made to the observer notes, three have been combined with other markers, two have been fundamentally changed (‘*Scanning*’ changed to ‘*Maintaining SA*’, and ‘*Decisive*’ changed to ‘*Indecisive*’), three have been deleted leaving only one marker within this category which is completely unchanged from this review.

SUMMARY OF CHAPTER 7

In this chapter detail is presented regarding an observational study undertaken in order to evaluate inter-rater reliability of the marker set developed in chapter 5, and first used in chapter 6. This study involved a number of dual HF observations watching ATCOs engaged in their controlling task using EFPS. At the end of the observations, a structured interview was undertaken to canvas the views of the HF observers regarding the method and content of the behavioural observation system.

The dual HF observations and interview permitted both qualitative and quantitative review of the behavioural observation system; although the limited number of observations did not permit sophisticated analysis. The results revealed strong links between how favourable and positive the comments were for individual markers, and the level of quantitative inter-rater agreement illustrated.

A number of revisions have been made to the observational marker set and method as a consequence of the quantitative and qualitative combined results. There was positive endorsement received regarding the employment of a frequency cap, the inclusion of which is designed to prevent attention tunnelling. Although what number the cap is set at may warrant further manipulation.

This study provided an early wider critique of the behavioural markers and associated method developed in chapter 5. The results have helped refine the design and content of the marker system, and provide initial construct validity to the approach. A more sophisticated and thorough inter-reliability study is warranted; as the maturity of the system is increased. Such maturity will be developed through repeated application in a wide variety of observational situations and environments.

CHAPTER 08 – CHANGES IN BEHAVIOUR OF TRAINEE ATCOs

"There is no rule on how to write. Sometimes it comes easily and perfectly; sometimes it's like drilling rock and then blasting it out with charges."
(Ernest Hemingway)

The study undertaken in Chapter 6 explored the changes in NTS behaviour of ATCOs transitioning from paper to electronic flight progress strips. The result suggests a number of NTS transferred between the two systems. This chapter provides details of a research study which was undertaken in order to examine changes in trainee ATCO Non-Technical Skill (NTS) behaviour, during a NATS aerodrome course. As this ab-initio group had no previous experience in ATC, this provided an opportunity to explore the emergence of proficiency, and validate the content and classification of the behavioural marker set through further application.

8.1 AIMS OF THE STUDY

The primary aim of this study was to observe trainee ATCOs over the duration of an aerodrome training course in order to record how NTS behaviours associated with expertise in ATC emerge. In addition to this, a number of supplementary aims were;

- To collect data with a regular and even distribution across the training course.
- To record any additional NTS markers indicative of each learning phase; review these and select and incorporate within the remaining observations (as per the study in Chapter 6)
- To increase the individual marker frequency cap, and evaluate the effect.

It was anticipated that the presence and prevalence of novice classified NTS behaviours will be high within this group at the start of training and diminish over time. In addition, intermediate behaviours are unlikely to be present at the start of their training and will emerge over time. Expert behaviours will emerge over time as per intermediate behaviours, but perhaps slower. Finally, it is anticipated that additional ab-initio behaviours will be present in the trainee ATCO group.

8.2 METHOD

Agreement from the NATS training college was obtained to follow from start to finish an aerodrome control training course. The four month course includes a number of classroom and simulator training sessions, in addition to periodic summative practical exams. An email was sent to the training unit, and sent onto the 10 trainee ATCOs at the start of their course. In addition information provided by the author was pinned to the notice board within the main simulation room (Appendix D11). At the beginning of the aerodrome course, the entire group of trainee ATCOs were ab-initio, with no previous ATC experience (a total of 10 trainee ATCOs).

Prior to the start of an observational session, consent to observe was verbally requested from each individual trainee ATCO. Permission to observe was granted on all occasions except on one instance (halfway through the study), where consent was refused by the trainee ATCO and instructor; as they had a difficult and challenging session planned, and did not want any form of potential distraction.

8.2.1 Schedule of Observations

With a copy of the course timetable provided by the unit, dates were selected which contained a high number of simulator training sessions (with a maximum of six 1 hour sessions undertaken in a single day). Permission was obtained to undertake seven days of observation across the training course.

In order to achieve the maximum number of observations from the seven days of agreed data collection activity, the observation days were scheduled upon days in which 5 or 6 training runs were scheduled. There was one specific request from the training management which was that no observations should occur immediately prior to a ‘summative’ exam, therefore this was also taken into consideration when selecting appropriate observation dates.

The study in Chapter 7 aimed for a 30 minute observation period to collect data. For this study, the period was change to 20 minutes. This change was made in order to allow the observation of two trainee controllers within one 1 hour training session. In practice, given slight differences in start and finish times within the training run, the second observation on certain occasions was curtailed, with the shortest lasting only 15 minutes. However the mean was 19 minutes and 23 seconds, close to the desired length of 20 minutes.

8.2.2 The Training Simulator

An aerodrome simulation training session is one hour in duration and begins with the trainee ATCOs printing out all the strips for the aircraft in the training scenario. The training begins with a 5 minute brief from the instructor, followed by 40 minutes of real-time simulation, ending with a 10 minute debrief. Several controllers are trained at the same time within the training room, using several separate workstations (Figures 8.1 and 8.2).

Figure 8.1 . Trainee ATCO with Instructor observing

Figure 8.2 - A trainee ATCO viewing aerodrome simulation screens

Half of the trainee ATCOs on the course would sit in the simulation room and receive training, whilst the other half would relocate to a separate room and help drive the simulation inputs. This simulation support involves making inputs into the simulation so that the air and ground vehicles on the surface and air picture screens respond as per instructed by the trainee ATCO.

8.2.3 Frequency cap

The study of expert ATCOs transitioning from paper to electronic flight progress strips (Chapter 6) employed a frequency cap 4 set to five 4 to limit the number of times a single behaviour was recorded. For this study, the cap was increased to ten. The cap was increased in order to provide greater potential fidelity to the resultant data, and to explore the value of this change upon the overall observation methodology.

8.2.4 Time Weighting

As per the study contained within Chapter 6, a weighting has been applied in order to normalise the data against varying observation length. The same formula has been applied to the data. Where the adjusted frequency (F_a) is sum of the recorded frequency (F) dividend by observation length (t), multiplied by a standard time constant (*T*); which in this instance is 20 minutes. The weighting as revealed within the results section has had no significant bearing on the outcome of the analysis.

$$F_a = \left(\frac{F}{t}\right)T$$

8.2.5 Observational detail

This study used version 3 of the observational sheet (Appendix C06), which was developed through the research detailed in Chapter 5, and modified based upon the results of the first application (Chapter 6) and inter-rater study (Chapter 7).

In a similar format to earlier observations (Chapters 4-7), the observer (author) sat in a chair positioned to allow observations of the side of the face, and of the paper flight strip board, and other equipment. The observation task involved the occurrence of specific behaviours noted down against the appropriate marker each instance it was displayed. A frequency limit of 10 instances for each marker within the observation period was set. Space on the observation sheet allowed any additional behaviours not previously contained, to be recorded (and immediately tracked in terms of frequency).

In addition, a number of other details were recorded such as exercise name and type, start and finish times. Following data collection, any additional behaviours recorded were collated, to await further review. Several new behaviours were identified, and displayed by multiple Trainee ATCOs. An assessment of these behaviours was undertaken to determine what these additional behaviours may indicate, and their overall suitability for inclusion with the existing markers set.

8.2.6 Participants

The course comprised ten male trainee ATCOs, all of whom were observed. None of the ten trainees had received previous ATC experience. Approximately two thirds of the way through the course, a further two trainees (one male, one female) joined the course. Both were repeating the last few aspects, having not passed these in a previous aerodrome course. These trainees were excluded from the study. All of the trainees progressed well through the course, with only one trainee failing the final summative course (and re-sitting and passing on the subsequent course). A good relationship was formed with the trainee ATCOs, who showed enthusiasm and interest in the study.

8.3 RESULTS

Across these seven days of observation, data was recorded with trainees having experienced a minimum of 4 and maximum of 43 hours of real-time simulation exposure (out of a total of approximately 45 hours of simulator based training available (Table 8.1)). In total, 67 individual 20 minute observations were made (the shortest 15 minutes, the longest 23 minutes, and the mean 19.5 minutes).

Observation Day	Date	Number of Observations	Approximate hours of simulator training exposure
1	01/02/2011	8 observations	4-5 hours
2	11/02/2011	9 observations	8-9 hours
3	23/02/2011	12 observations	11-12 hours
4	16/03/2011	8 observations	24-25 hours
5	24/03/2011	10 observations	28-29 hours
6	14/04/2011	10 observations	36-37 hours
7	16/05/2011	12 observations	40-43 hours

Table 8.1 . Observation Schedule

The aim of the observations was to gather equal quantities of data from all trainee ATCOs across the seven days of observation. It was not possible on all occasions to observe every trainee ATCO on each observation day, however given the practical constraints; the observations made were distributed as evenly as possible across the group (Table 8.2). A code name known only by the author was used to disguise the identity of each ATCO (Alpha – Juliet).

Individual ATCO Identity	No. occasions observed	Observation Day (no. hours simulator experience)						
		Day 1 (4-5 hr)	Day 2 (8-9)	Day 3 (11-12)	Day 4 (24-25)	Day 5 (28-29)	Day 6 (36-37)	Day 7 (40-43)
Alpha	7	x	x	x	x	x	x	x
Bravo	6	x		x	x	x	x	x
Charlie	6		x	x	x	x	x	x
Delta	8	x	x	xx	x	x	x	x
Echo	6	x	x	x		x	x	x
Foxtrot	6	x		x	x	x	x	x
Golf	6		x	x	x	x	x	x
Hotel	7	x	x		x	x	x	xx
Indigo	7	x	x	xx		x	x	x
Juliet	8	x	x	xx	x	x	x	x

Table 8.2 . Observations for individual Trainee ATCOs ('x' denotes an observation)

8.3.1 New Behaviours

A total of 23 additional potential behaviours were recorded across the 67 observations (Appendix D03). Upon each occasion a new marker was observed, it was recorded on the observation sheet, and a frequency count was made against it during the observational session. After 15 runs, a review of the 23 markers was undertaken in order to select those which were considered of sufficient merit, mutual exclusivity, of sufficient frequency, and of low susceptibility to individual difference; to be taken forward and monitored for the remaining duration of this observational study. Of the seven identified new behaviours included within the analysis, six relate to body posture and body movement.

Table 8.3 presents the seven new markers identified. All seven of these markers were observed within the first day of observation and were considered to offer useful insight into proficiency. The total number of times each of these seven behaviours was recorded is also presented.

Potential Marker	Indicator of...	Codec	No. times observed
Clumsiness moving and interacting with paper strips	Demonstrates physical slowness at moving strips and interacting with workstation, movements over emphasised	Clumsy	36
Nervous Physically hand shaking	Shaking hands, shoulders, and holding and writing with a pen	Nervous Physical	54
Both hands to move strips (poor MMI motor skills)	Experienced controllers move strips with one hand, quickly and deftly, using two hands to move them is extremely unusual	Low dexterity	216
frantic writing/frantic task/rushed	Unless in exceptional situations ATCOS approach the task calmly, not frantically	Frantic approach	10
Nervous or Extremely Quiet Voice	A strong, positive, and commanding voice is required, nervous quiet voices are undesirable	Nervous voice	70
Points things out to themselves or instructor	Re-enforcing spatial information through physical referencing	Spatial Pointing	41
leans right over strips	Tunnelled into the workstation, rather than sitting back and absorbing 'the big picture'	Leans in	72

Table 8.3 . Additional markers identified, selected, and tracked

Table 8.4 contains a Spearman's Rho correlation analysis for the seven additional markers tracked during this study. Only two markers reveal reliable weak negative correlations; ('*CIums~*' and '*Low Dexterit~*'). No other markers reveal reliable negative (or positive) correlations.

Time Weighted Results			
Marker	**Spearman's Rho**	**Sig.**	**N**
Clumsy	-.292	.016*	67
Nervous Physical	-.145	.243	67
Low dexterity	-.334	.006*	67
Frantic approach	-.136	.272	67
Nervous voice	.036	.772	67
Spatial Pointing	-.214	.082	67
Leans in	.005	.968	67

Table 8.4 - Seven additional markers identified correlated against amount of system exposure (* denotes significance at p≤.05)

8.3.2 Changes in behaviour

The primary goal of this research study has been to establish how a number of non-technical behaviours change in presence and prevalence within a population of trainee ATCOs learning to use paper flight strips as part of their aerodrome training. Table 8.5 provides a correlation of the total number of instances beginner, intermediate, and expert behaviours were displayed calculated against the amount of system exposure (time). Figures 8.1 – 8.3 present the descriptive statistics for the correlations presented in Table 8.5. As anticipated, the findings reveal a reliable negative correlation for beginner behaviours (decreasing moderately over time). However unlike the study in Chapter 6, reliable positive correlations were also observed for intermediate and expert behaviours.

Time Weighted Results				
Categorised Markers	**Spearman's Rho**	**Sig.**	**N**	**Means**
Beginner	-.532	.000*	67	12.4
Intermediate	.427	.000*	67	24.3
Expert	.748	.000*	67	20.1

Table 8.5 . Behaviour class correlated against amount of system exposure (* denotes significance at p≤.05)

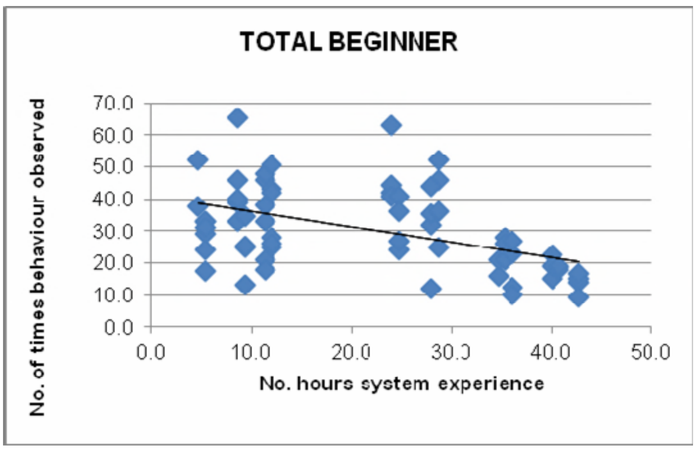


Figure 8.1 . Total Beginner Behaviours against Hours of System Exposure

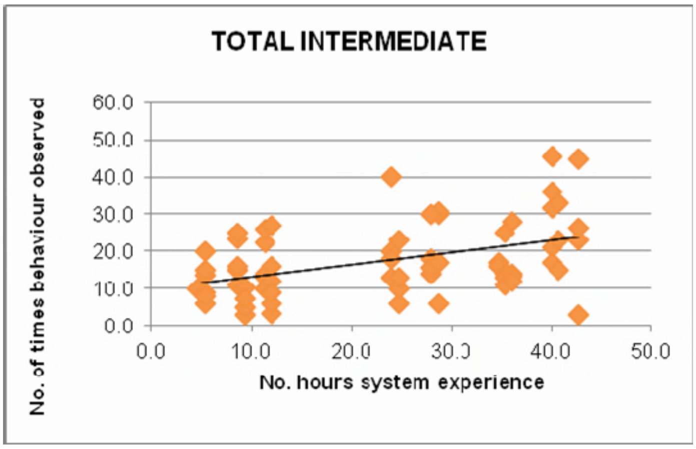


Figure 8.2 - Total Intermediate Behaviours against Hours of System Exposure

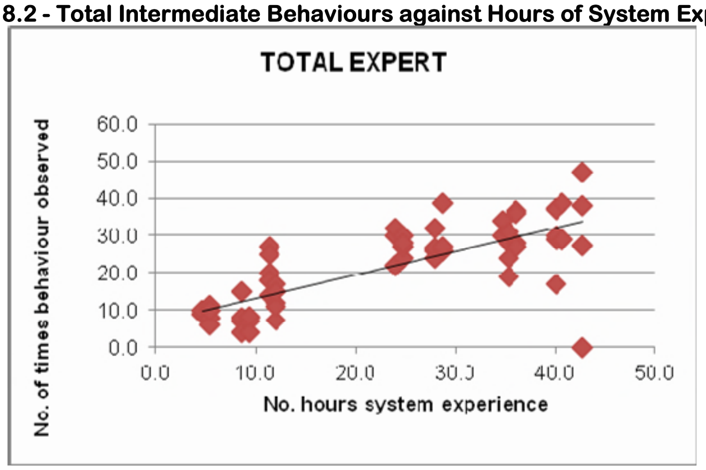


Figure 8.3 - Total Expert Behaviours against Hours of System Exposure

8.3.3 Beginner Behaviours

Table 8.6 presents the Spearman's Rho correlation of 26 individual beginner behaviours against system exposure time. This total includes the seven newly identified markers. It has been anticipated that these markers will decrease in prevalence as proficiency with the replacement flight progress system increases. Across the 67 observations, six out of the 26 beginner behaviours demonstrated a weak to strong negative correlation against system exposure time ('Delays/Repeats', '*Slow Hesitant*', '*Serial Tasking*', '*Surprise*', '*Clums~*', '*ow Dexterit~*'). Charts for all these individual markers may be found in Appendix D08.

Time Weighted Results			
Behavioural Marker	**Spearman's Rho**	**Sig. (2-tailed)**	**N**
Team Short/Snappy	Not observed		67
Negative comments	.061	.621	67
Confusion/Uncertainty	-.185	.135	67
Tongue-tied	-.220	.074	67
Delays/Repeats	-.619	.000*	67
Standbys	-.026	.834	67
Frustrated Irritated	-.029	.818	67
Indecisive / Hesitant	.078	.533	67
General Queries	-.001	.995	67
Reactionary	-.197	.111	67
Tech Discussion	.204	.097	67
Phys. Frustration	-.134	.280	67
Overconfident	.058	.642	67
Slow/Hesitant	-.265	.030*	67
Serial Tasking	-.578	.000*	67
Incorrect actions	-.191	.121	67
Surprise	-.295	.015*	67
Affirm Before	-.129	.300	67
Affirm After	-.072	.562	67
Clumsy	-.292	.016*	67
Nervous Physical	-.145	.243	67
Low dexterity	-.334	.006*	67
Frantic approach	-.136	.272	67
Nervous voice	.036	.772	67
Spatial Pointing	-.214	.082	67
Leans in	.005	.968	67

Table 8.6 - Frequency of Beginner Behaviours correlated over system exposure time (*significance at $p \leq .05$)

Intermediate behaviours

Table 8.7 presents the Spearman's Rho correlations for 13 individual beginner behaviours against system exposure time. It has been anticipated that these markers will increase prevalence; as proficiency with the replacement flight progress system increases. Across the 67 observations, three out of the 13 intermediate behaviours demonstrated a moderate to strong positive correlation against system exposure time (*'Team Aware*', *Team Contribute*', and *'Peripher}*'). Charts for all these individual markers may be found in Appendix D08.

Time Weighted Results			
Behavioural Marker	**Spearman's Rho**	**Sig. (2-tailed)**	**N**
Positive comments	-.118	.341	67
Apologetic	-.096	.440	67
Social	-.027	.827	67
Self Affirm	.060	.632	67
Team Aware	.529	.000*	67
Team Contribute	.589	.000*	67
Pace Fast	-.179	.146	67
Pace Slow	.167	.177	67
Fatigue	-.097	.437	67
Adjust MMI	.044	.726	67
Fidgets	.034	.784	67
Periphery	.689	.000*	67
Play/Sandpit	Not Observed		67

Table 8.7 - Frequency of Intermediate Behaviours correlated over system exposure time (*significance at $p \leq .05$)

8.3.4 Expert behaviours

Table 8.8 presents the Spearman's Rho correlation of five individual expert behaviours against system exposure time. It has been anticipated that these markers will increase in prevalence; as proficiency with the replacement flight progress system increases; although at a more modest rate than those markers classified as intermediate. Across the 67 observations, four out of the five intermediate behaviours demonstrated a moderate to strong positive correlation against system exposure time ('*Automatic Memory*', '*Plans Ahead*', and 'Maintains *Global Awareness*'). Charts for all these individual markers may be found in Appendix D08.

Time Weighted Results			
Behavioural Marker	**Spearman's Rho**	**Sig. (2-tailed)**	**N**
Dual Tasking	.661	.000*	67
Automatic Memory	.546	.000*	67
Plans Ahead	.696	.000*	67
Maintains Global Awareness	.629	.000*	67
Quirks	4004	.975	67

Table 8.8 - Frequency of Expert Behaviours correlated over system exposure time (*significance at $p \leq .05$)

8.3.5 Individual Trainee ATCO Overall Results

In addition to the results for the overall group of trainee ATCOs, it has been possible with the data captured to explore behavioural change for individuals. This has been possible given the small number of ATCOs observed, the large amount of data collected, and the greater degree of control as to how and when the observations would take place. Table 8.9 presents the development class results for each individual ATCO. A large number of the correlations indicate strong correlation, however many of these are unreliable. In total two individual ATCOs showed reliable negative correlations for beginner behaviours reducing in prevalence over time. One ATCO showed a reliable positive correlation for intermediate behaviours increasing over time. Finally seven out of the ten individual trainee ATCOs demonstrated reliable positive correlations for expert behaviours increasing over time.

Trainee ATCO	N	Results	Development Class – Time Weighted Results Spearman's Rho Correlation		
			Beginner	Intermediate	Expert
Alpha	7	Correlation Coefficient	.036	.393	.964
		Significance	.470	.192	.000*
Bravo	6	Correlation Coefficient	-.771	.880	.943
		Significance	.072	.021*	.005*
Charlie	6	Correlation Coefficient	-.429	.771	.829
		Significance	.397	.072	.042*
Delta	8	Correlation Coefficient	-.429	.771	.829
		Significance	.397	.072	.042*
Echo	6	Correlation Coefficient	-.714	-.600	.086
		Significance	.111	.208	.872
Foxtrot	6	Correlation Coefficient	-.543	.638	.829
		Significance	.266	.173	.042*
Golf	6	Correlation Coefficient	-.886	.657	.829
		Significance	.019*	.156	.042*
Hotel	7	Correlation Coefficient	-.847	.721	.464
		Significance	.016*	.068	.294
Indigo	7	Correlation Coefficient	.036	.607	.679
		Significance	.939	.148	.094
Juliet	8	Correlation Coefficient	-.429	.667	.874
		Significance	.289	.071	.005*

Table 8.9 - Individual trainee ATCOs behavioural change correlated over system exposure time (*significance at p≤.05)

8.3.6 Individual Trainee ATCO Single Behavioural Marker Results

In addition to the grouped development category data (beginner, intermediate, expert) for each individual Trainee ATCO, it has been possible to evaluate change to single behavioural markers. Tables 8.10 through 8.12 present the reliable correlations observed for the frequency of single markers against the amount of simulation exposure time the trainee ATCOs had received in the simulator. The data is select, where only the significant affects are presented. The full results may be found in Appendix D09, which includes non-effects. The reliable correlations were all in the direction anticipated, with the exception of 'technical discussion' (Table 8.10) which was found to have a strong positive correlation for trainee ATCOs Alpha and Bravo (as denoted by a '*').

Beginner Markers	Trainee ATCO									
	A	B	C	D	E	F	G	H	I	J
Confusion/Uncertainty								x		
Delays/Repeats		x								x
Frustrated Irritated									x	
Reactionary							x			
Tech Discussion	x*				x*					x*
Serial Tasking			x	x	x	x				
Surprise			x	x						
Low dexterity			x	x						x
Spatial reference							x			
Affirm Before										x
Affirm After										x

Table 8.10 . Individual trainee ATCOs beginner behavioural change correlated over system exposure time (significance at p≤.05), selected results

Intermediate Markers	Trainee ATCO									
	A	B	C	D	E	F	G	H	I	J
Team Aware								x		
Team Contribute								x		
Adjust MMI	x									
Periphery	x		x	x		x	x			x

Table 8.11 - Individual trainee ATCOs intermediate behavioural change correlated over system exposure time (significance at p≤.05), selected results

Expert Markers	Trainee ATCO									
	A	B	C	D	E	F	G	H	I	J
Dual Tasking	x	x							x	x
Automatic/Quick		x	x	x		x	x			x
Plans Ahead	x		x	x						x
Maintains Global Awareness	x	x					x			x

Table 8.12 - Individual trainee ATCOs expert behavioural change correlated over system exposure time (significance at p≤.05), selected results

8.4 DISCUSSION

The results of this study are somewhat mixed, as is the case for the study in Chapter 6. When examining the main effects, for the grouped development class behaviours (Table 8.5), reliable correlations are found for all three categories (beginner, intermediate, expert)[5]. This result is a key finding, as it was anticipated that an ab-initio group would reveal reliable correlations for the intermediate and expert behaviours increasing in prevalence over time. This result does suggest that these intermediate and expert behaviours might transfer across systems, and account for the lack of difference when examining expert ATCOs migrating from paper to electronic flight strips.

When individual trainee ATCO behaviours are examined, there are fewer instances where the group development class data reveals reliable change (Table 8.9), although this may be the result of stretching the small amount of individual data too thinly. This section explores the main results and discusses the main reasons considered to account for the anticipated and unanticipated results.

[5] This result is different to that of Chapter 6 where the intermediate and expert categories revealed no significant change over time (when considering the first 41 observations only).

8.4.1 Individual marker results within the trainee ATCO group

Of the new markers identified in the first few observations, and included during the study as further indications of beginner behaviour, ‘clumsy’ and ‘low dexterity’ have reliable correlations, although the strength of correlation is weak.

With regards to the existing individual markers observed in this study, Tables 8.64 8.8 present six beginner, three intermediate, and four expert behaviours that demonstrate reliable weak to strong correlations against system exposure time (all in the desired direction). In total this represents 13 out of 44 markers demonstrating reliable weak to strong correlations in the anticipated direction. Of particular note are the results for individual expert behaviours (Table 8.8), where four out of the five (80%) demonstrated a weak to strong positive correlation against system exposure time (‘Automatic Memory’, ‘Plans Ahead’, ‘Maintains Global Awareness’).

Three out of the 13 intermediate behaviours demonstrated moderate to strong positive correlation against system exposure time (‘Team Aware’, Team Contribute’, and ‘Periphery’), Four out of the five intermediate behaviours demonstrated a moderate to strong positive correlation against system exposure time (‘Automatic Memory’, ‘Plans Ahead’, ‘Maintains Global Awareness’). These results are as anticipated.

8.4.2 Behaviour Class results for individual trainee ATCOs

The large number of observations undertaken, across a small number of trainee ATCOs, has permitted the evaluation of behavioural change at an individual level. At this level, only one trainee ATCO (Golf) revealed a strong negative correlation for beginner behaviours. For intermediate behaviours, once again only one trainee ATCO (on this occasion Bravo) was found to have a strong positive correlation. Finally seven trainee ATCOS (Alpha, Bravo, Charlie, Delta, Foxtrot, Golf, and Juliet were found to have very strong (0.8.29 – 0.964) positive correlations with training exposure.

Several further individual behaviour class correlations for individual trainee ATCOS did reveal weak to strong correlations, however they were unreliable. In some instances the p value was approaching significance, which suggests that a few additional data points may have resulted in a reliable effect. It is clear that when calculating changes in behaviour on an individual person, a number of observation sessions are needed in order to infer significant correlation results.

8.4.3 Individual behavioural marker results for individual trainee ATCOs

When exploring the results of individual markers for each individual ATCO the total number of markers which display reliable weak to strong correlations increases to 19 (11 from the beginner category, 4 from the intermediate category, and 5 from the expert category).

When examining individual behaviours for individual ATCOs the marker ‘technical discussion’ (Table 8.10) was found to have a strong positive correlation for trainee ATCOs Alpha and Bravo. This correlation is the only occasion where an individual marker, for individual trainee ATCOs has been in the direction anticipated. Given that both reliable negative and positive correlations have been found with this marker, its utility to indicate phased development is therefore unreliable.

8.4.4 Evaluation of ATCO proficiency

Nine out of the ten trainee ATCOs passed the course in the first instance, with one trainee passing in a subsequent course. All ten are now fully operational ATCOs at various NATS tower units. This is very important to note, because there is an indirect link which may be drawn between the reliable weak to strong correlations for changing prevalence of several behavioural markers observed, and emerging trainee ATCO competency.

The results reveal that the trainee ATCOs rapidly developed the basic motor skills to manage the strip board and HMI, and gradually the novice behaviours displayed reduced in prevalence. They adopted a more relaxed posture and body language, were building a rapport with the instructors, and they generally became more comfortable in undertaking the tasks. One note made during the observations by the author concerned how in the first day's training session observed, the trainee ATCOs were all dressed very smart, with pressed shirts and polished shoes. But by the time of the second visit a few days later, all the trainees were in jeans and t-shirts.

8.4.5 Phases of learning

This study supports the findings of the first study (Chapter 6), in that the various behaviours observed provide evidence of incremental learning and development. However the effects are really subtle and con-current, in that a single person may display a mixture of beginner, intermediate, and expert behaviours all at the same time. Therefore it is the overall combination of markers present, and their respective quantities which might afford an overall indication of proficiency attained.

Research question 2 looks to determine whether there exist separate distinct phases of behavioural change which provide indications of separate stages in learning and development. This study suggests that there does not appear to be clearly distinct phases; instead there is complex multi layered approach to the acquisition of cognitive and motor skills.

8.4.6 Observational method

The increased control afforded within this study has provided a stronger foundation to monitor subtle changes in NTS behaviour over time. The observation length was more consistent than all previous studies (Chapters 4 – 7), with the 20 minute period providing sufficient time for a great number of behaviours to be displayed. The increased frequency count to 10 provided a greater range within the data, and the cap was reached on several occasions suggesting appropriate compatibility with the observational duration.

SUMMARY OF CHAPTER 8

In this chapter the details are provided of an observational study which followed a group of Trainee ATCOs during aerodrome controller training. The trainees demonstrated considerable change in the presence and prevalence of NTS as their training progressed, with several reliable strong correlations found. These correlations reveal how aspects of behaviour evolved, as technical competency and proficiency increased. In addition, a number of ab-initio behaviours were identified two of which demonstrated reliable weak correlations and have been integrated within the marker set. The totality of this study has provided greater insights into how certain NTS change over time, a stronger more refined marker set has also been produced, as well as further insights into the utility of the observation method.

CHAPTER 09 – BEHAVIOUR IN NON-BENIGN CONDITIONS

"Life is a series of experiences, each one of which makes us bigger, even though sometimes it is hard to realize this. For the world was built to develop character, and we must learn that the setbacks and grieves which we endure help us in our marching onward."
(Henry Ford)

The study in Chapter 6 explored changes in behaviour over time for qualified ATCOs transitioning from paper to electronic strips. The observational data did not reveal significant changes in behaviour for markers within the intermediate and expert development classification. One possible cause of this unanticipated result was that the live operational conditions were extremely benign. A repeatedly experienced limitation has been that permission was not obtained to observe ATCOs during high workload and high pressure circumstances. This chapter details a study where non-benign observations were undertaken.

9.1 AIMS OF THE STUDY

The study undertaken within this chapter has the following aims:

1) Capture observational data in an environment with non-benign conditions.
2) Compare this data to selected data captured in benign live operational conditions.
3) Explore potential differences in the presence and prevalence of the behavioural markers, across benign and non-benign conditions.

The expectation of this study is that behaviours within the intermediate and expert categories are affected by the level of task load and complexity within the environment. They will therefore occur significantly more often in a non-benign high workload environment, than the benign environment where the final series of live observations were made (Chapter 6). A comparison of the data collected from these two environments will enable the identification of individual markers which may be considered to be intrinsically impacted by task and environmental conditions. Such a finding will provide useful knowledge when considering changes in behaviour in uncontrolled conditions during periods of learning and development.

9.2 METHOD

As part of his wider employment, this researcher held the position of Human Factors lead for a NATS airspace project that developed new procedures and airspace for the 2012 Olympic Games (this was a period of time anticipated to be of high workload and air traffic complexity). A key activity of this NATS project was to develop and test the procedures and airspace design for London Terminal Control and Area Control operations through a series of real-time simulations, ahead of implementation. These simulations afforded an opportunity to undertake observations of ATCOs working at high levels of intensity and complexity, whilst operating in slightly unfamiliar circumstances with new procedures and airspace. No change was made to their existing paper flight strip systems, at this time.

9.2.1 Time Weighting

The desired duration of each observation in this study was 20 minutes. As per previous studies (Chapter 6, Chapter 8) a weighting has been applied in order to normalise the data against varying observation length. Where the adjusted frequency (F_a) is sum of the recorded frequency (F) dividend by observation length (t), multiplied by a standard time constant (T); which in this instance is 20 minutes.

$$F_a = \left(\frac{F}{t}\right)T$$

9.2.2 Frequency Cap

A frequency cap was applied to each individual marker during observation. This study chronologically followed the observations of trainee ATCOs (Chapter 8), the frequency cap was therefore kept consistent during data collection (set to 10). Although for later comparison analysis a retrospective cap was subsequently applied (lowering the cap to 5).

9.2.3 Subjective questionnaire

As part of the NATS Olympics airspace simulation, an end-of-run questionnaire was administered (by the author) asking the ATCO to make a number of assessments regarding the simulation period they has just experienced in addition to a number of technical questions concerning the design of the Olympics Airspace and procedures.

Two useful pieces of information were captured with relevance to this study. ATCOs' subjective assessment of their average and peak workload experienced within the run, as well as an overall rating of Situation Awareness. A simple behaviourally anchored 10-point rating scale was used to capture this data.

9.3 RESULTS (NON-BENIGN ENVIRONMENT)

The observations to collect data in non-benign conditions were made within two separate airspace simulations at NATS, both of which were designed to test new airspace and procedures ahead of the 2012 Olympic Games. Seven were made during a simulation in March 2011 and a further three observations in a simulation in June 2011. Each observation made was of a different experienced TC ATCO, the number of participants in this study was therefore ten. The shortest observation was 18 minutes, the longest 22 minutes, with a mean of 19 minutes 24 seconds.

The group statistics for the behavioural markers is presented in Table 9.1. The following markers were not observed within this non-benign environment and have not been included in Table 9.1:

- *'Affirm Before'*
- *'Affirm After'*
- *'Apologetic'*
- *'Confusion'*
- *'Over confident'*
- *'Incorrect Action'*
- *'Indecisive'*
- *'Standbys'*
- *'Tongue tied'*

The subjective results recorded by the ten ATCOs who were observed during this study indicated that additional compensation was needed, and that their levels of workload were high, and impacted upon their perceived levels of Situation Awareness. Given the sensitivity of this data, the detail has not been presented in this thesis. The data does however confirm that the conditions experienced were not benign. A non-benign (Olympics simulation) versus benign (final 11 observation made in observational study detailed in Chapter 6).

Behavioural marker	Olympics (N = 10) (Non-benign Environnent)	
	Mean	Std. Deviation
Negative Comments	1.0497	1.22179
Positive Comments	0.2164	0.45637
Social	3.0809	1.87542
Self Affirm	0.3222	0.51971
Delays Repeats	0.1111	0.35136
Irritated / Frustrated	1.5877	1.33981
General Queries	0.4386	0.77152
Reactionary	0.3164	0.51008
Team Aware	4.6516	1.13169
Team Contribute	4.2587	1.21122
Team Short / Snappy	0.1053	0.33287
Technical Discussion	3.8839	2.26749
Pace Fast	2.7152	2.23489
Pace Slow	1.8661	1.69052
Phys. Frustration	0.4216	0.73703
Fatigue	3.5394	2.13992
Adjusts MMI	2.9821	1.99031
Fidgets	3.6649	2.11239
Slow Hesitant	0.1111	0.35136
Play Sandpit	0.1111	0.35136
Dual Tasking	5.1738	0.3281
Serial Tasking	0.8198	0.78838
Surprise	0.1	0.31623
Quickly Locates	5.1738	0.3281
Periphery	1.4224	0.92618
Keeps Information Open	4.3142	1.31696
Maintains Global SA	4.9633	0.7135
Quirks	1.6491	2.20825

Table 9.1 - Group statistics (Non-benign condition)

9.4 COMPARISON ANALYSIS

A key purpose of this study has been to compare the results of data collected in non-benign conditions with those captured in earlier research in the benign live operational environment. As a reminder to the reader the final 11 observations made with ATCOs transitioning from paper to electronic flight progress strips (Chapter 6) were made in the live-operational environment. These ATCOs had at this point 75-80 hours interaction experience using EFPS in total, 60-65 hours of which were within the live environment. The desired duration per observation in that study was 30 minutes. In order to allow a direct comparison between this benign environment data, with the non-benign Olympics simulations data, a time weighting of 20 minutes was retrospectively applied.

9.4.1 Frequency cap comparison

The 11 runs of data captured in the live environment were made using a frequency cap of 5. To allow comparison, the data captured in the Olympics simulations was retrospectively capped at 5 (from 10). In simple terms, any frequency score of 6-10 was retrospectively capped at 5.

Observation Sheet Versions

Observation sheets versions 1 and 2 were used to collected data in a benign live environment (Chapter 6). However version 4 of the observation sheet was used to collect the data in this study, within a non-benign simulation environment. In order to manage the disparity resulting from two different observation sheet versions being used in the two studies, a number of changes were made to the data from the two studies:

- Firstly the following markers were removed from the benign conditions (Chapter 6) data as they were markers removed during the previous studies, so no comparison could be made; i) '*Best Practice*', ii) '*Interface Scan*', iii) '*Cool Calm*', iv) '*Decisive*', and v) '*Input Device Tick Off*'.
- Secondly the following markers were combined to the benign conditions data, as these markers were combined during the previous studies, so no direct comparison could be made i) '*Frustration*', '*Verbal Frustration*', and '*Irritable*', ii) '*Automatic*' '*Muscle Memory*', '*Quickly Locates*' '*Confident Control*' (to be referred to as '*Automatic Memory*') and iii) '*Plans Ahead*', and '*Keeps Info Open*' (to be referred to as '*Plans Ahead*').
- Finally the marker '*Indecisive*' was removed from the non-benign (Olympics) data, as this marker was added during the previous studies, so no comparison could be made.

The outcome of the amendments to the data resulted in 37 markers available for direct comparison (Table 9.2)

Behavioural Markers		
Beginner Behaviours	**Intermediate Behaviour**	**Expert Behaviour**
· Negative comments	· Positive comments	· Dual Tasking
· Confusion	· Apologetic	· Plans Ahead
· Tongue-tied	· Social	· Cyclic Scan
· Delays/Repeats	· Self Affirm	· Quirks
· Overconfident	· Play/Sandpit	· Automatic Memory
· Slow/Hesitant	· Periphery	
· Incorrect actions	· Team Aware	
· Surprise	· Team Contribute	
· Reactionary	· Pace Fast	
· Team Short/Snappy	· Pace Slow	
· Phys. Frustration	· Fatigue	
· Serial Tasking	· Adjust MMI	
· General Queries	· Fidgets	
· Standbys		
· Irritated/Frustrated		
· Technical Discussion		
· Serial Tasking		
· Affirm Before		
· Affirm After		

Table 9.2 - Behavioural Markers used in the benign non-benign comparison

The markers; 'Tongue tied', 'standbys', indecisive' 'Affirm before', and 'Affirm after' were not observed in either sets of data, and have been excluded from the analysis, leaving a total of 32 markers for comparison (Table 9.3).

	Olympics (N = 10) (Non-Benign Condition)		Chapter 6 (N = 11) (Benign Condition)	
Behavioural marker	Mean	Std. Deviation	Mean	Std. Deviation
Negative Comments	1.0497	1.22179	0.2039	0.35471
Positive Comments	0.2164	0.45637	0.0826	0.2741
Apologetic	0	0	0.0826	0.2741
Social	3.0809	1.87542	1.6181	1.41258
Confusion	0	0	0.2125	0.48688
Self Affirm	0.3222	0.51971	0.4683	1.09564
Delays Repeats	0.1111	0.35136	0.4704	0.76103
Irritated / Frustrated	1.5877	1.33981	0.5631	1.08173
General Queries	0.4386	0.77152	0.0606	0.20101
Reactionary	0.3164	0.51008	0	0
Team Aware *	4.6516	1.13169	0.4001	0.5797
Team Contribute *	4.2587	1.21122	0.4263	0.63161
Team Short / Snappy	0.1053	0.33287	0	0
Technical Discussion *	3.8839	2.26749	0.7355	1.04982
Pace Fast *	2.7152	2.23489	0	0
Pace Slow	1.8661	1.69052	1.3769	0.66695
Phys. Frustration	0.4216	0.73703	0.2259	0.5647
Fatigue *	3.5394	2.13992	0.2222	0.4969
Adjusts MMI *	2.9821	1.99031	0.1939	0.44767
Fidgets *	3.6649	2.11239	0.8995	1.31596
Over Confident	0	0	0.1212	0.40202
Slow Hesitant	0.1111	0.35136	0.4885	0.86004
Play Sandpit	0.1111	0.35136	0.1212	0.40202
Dual Tasking *	5.1738	0.3281	2.9714	3.14822
Serial Tasking *	0.8198	0.78838	0.1653	0.5482
Incorrect Action	0	0	0.0727	0.24121
Surprise	0.1	0.31623	0.6061	2.01008
Quickly Locates *	5.1738	0.3281	3.8059	0.91772
Periphery *	1.4224	0.92618	0.0606	0.20101
Plans Ahead *	4.3142	1.31696	2.0601	1.72421
Maintains Global SA *	4.9633	0.7135	1.2351	0.98794
Quirks	1.6491	2.20825	0.3708	0.75273

Table 9.3 - Group Statistics (Non-benign and Benign conditions) '*' denotes sig. difference between means

9.4.2 T-Test

A separate independent samples t-test was undertaken for each of the 32 common behavioural markers collected during this study (non-benign) with the live-operational data collected during the Chapter 6 study (benign). Out of the 32 common markers which were observed in both the non-benign and benign studies, Leven's test reveals 21 markers where there is unequal variance between the two conditions, and 11 with equal variance (see Appendix D10 for the full data set). A result of unequal variance from Leven's test indicates a lack of homoscedasticity in the data. This variance is probably the result of a sampling effect (likely caused by a small number of study participants).

Of the 21 markers with unequal variance 9 of these reveal a significant difference ($p \leq .05$) between the benign and non-benign conditions ('*Team Contribute*', '*Technical Discussion*', '*Pace Fast*', '*Fatigue*', '*Adjust MM*', '*Fidgets*', '*Dual Tasking*', '*Quickly Locates*', and '*Periphery*').

Of the 11 markers with equal variance, 4 reveal a significant difference between the benign and non-benign conditions ('*Team Aware*', '*Plans Ahead*', '*Serial Tasking*', and '*Maintains Global SA*'). Table 9.4 presents selected t-test results for all behavioural markers which reveal significant differences between the conditions. Therefore, in total, 13 out of the 32 common markers which were observed in both the non-benign and benign studies demonstrate a significant difference in scores between the benign and non-benign conditions.

Behavioural Markers revealing significant differences between conditions		**Levene's Test for Equality of Variances**		**test for Equality of Means**		
		F	**Sig.**	**t**	**df**	**Sig.**
Team Contribute	Equal variances not assumed	5.017	0.037	8.959	13.27	0.001
Tech. Discussion		10.676	0.004	4.017	12.425	0.002
Pace Fast		44.128	0	3.842	9	0.004
Fatigue		24.154	0	4.786	9.883	0.001
Adjusts MMI		43.051	0	4.331	9.828	0.002
Fidgets		7.213	0.015	3.559	14.812	0.003
Dual Tasking		6.163	0.023	2.306	10.239	0.043
Quickly Locates		5.638	0.028	4.629	12.73	0.001
Periphery		20.467	0	4.553	9.771	0.001
Team Aware	Equal variances assumed	1.729	0.204	10.993	19	0.001
Serial Tasking		3.803	0.066	2.227	19	0.038
Plans Ahead		0.752	0.397	3.34	19	0.003
Maintains Global SA		4.301	0.052	9.821	19	0.001

Table 9.4 - Independent Samples Test – Significant results

Finally, Table 9.5 presents all behavioural markers revealing significant difference across the two conditions (benign and non-benign), along with their corresponding learning and development classification. In all instances, the significant difference revealed greater frequency of prevalence in the non-benign environment.

Behavioural Markers		
Beginner Behaviours	**Intermediate Behaviour**	**Expert Behaviour**
· Serial Tasking	· Pace Fast	· Dual Tasking
· Tech. Discussion	· Fatigue	· Quickly Locates
	· Adjusts MMI	· Keeps Info. Open
	· Fidgets	· Maintains Global SA
	· Periphery	
	· Team Aware	
	· Team Contribute	

Table 9.5 - Significant results according to marker development classification

9.5 DISCUSSION

The purpose of this study has been to capture observational data in a non-benign high workload environment in order to compare this data to selected data captured in benign live ops conditions and explore how behaviours displayed may differ in prevalence due to the task and environment. In particular, the focus has been upon whether intermediate and expert development class markers may differ, and account for a lack of change in these behaviours when examined in the previous study (Chapter 6). The 10 observations made in the non-benign environment have been compared to the data collected during the final 11 live observations made at Edinburgh Airport (Chapter 6). The results reveal several significant differences between these two data sets, the detail of which is now discussed.

9.5.1 Beginner Behaviours

In both the benign and non-benign environments, the ATCOs involved may be considered as experts; the lack of presence for a number of beginner behaviours is therefore explicable *$'Tongue tied', 'standbys', 'Indecisive', 'Affirm before',* and *'Affirm after'*). However there were two beginner behaviours that were significantly more prevalent within the non-benign Olympics environment, namely; *'Technical discussion*', and '*Serial tasking'*. The significant increase in '*Technical Discussion*' is likely to be accounted for by the revised airspace and procedures being assessed within the simulation. The significant increase in '*Serial Tasking*' is considered likely due to increased resilience on behalf of the ATCO whilst undertaking the task in demanding conditions, and 'safeguarding' their actions and activities accordingly.

9.5.2 Intermediate and Expert

Seven intermediate and four expert development class behaviours reveal significantly greater prevalence within the non-benign high workload environment of the Olympics *('pace fast', 'fatigue', 'adjust mmi', 'fidgeting', 'periphery', 'team aware', 'team contribute', 'dual tasking', 'quickly locates', 'Plans Ahead', and 'maintains Global SA'*). Collectively these markers indicate a number of skilled NTS on display, specifically in the areas of task management, and levels of task activity. These results provide supportive evidence that behaviour displayed may be directly impacted by environment and task situation, and as such they are compounding variables to observational research in this domain.

SUMMARY OF CHAPTER 9

In this chapter a small study has been undertaken in order to determine whether the presence and prevalence of certain behavioural markers may be impacted upon by the task and environmental conditions at hand. This study is simple, and has many uncontrolled variables which might account for the results collected. However it affords a rare opportunity to examine behaviour collected in non-benign conditions; and therefore might provide further insights into the behavioural markers identified and studied through earlier work (Chapters 5 to 8). Data collected in high-workload conditions was contrasted against data captured within benign conditions. This comparison indicated that several markers associated with task delivery and awareness was exhibited more frequently in non-benign conditions. These findings suggest that behaviour is impacted by task and environmental factors, and as a consequence must be taken into consideration when undertaking observations. The lack of change for the intermediate and expert behaviours in Chapter 6 is at the very least partially accounted for by the benign conditions which existed when the final observations were made in the live environment.

CHAPTER 10 – CONSOLIDATION OF MARKERS & MAIN DISCUSSION

"Not everything that counts can be counted. Not everything that can be counted counts" (Albert Einstein)

The preceding Chapters have described several research activities concentrating upon the identification, selection, application, and evaluation of Behavioural Markers within the Air Traffic Control domain. Specifically this research has explored changes in behaviour as NTS proficiency increases over time and proposed a set of Behavioural Markers to identify behavioural changes that reflect NTS proficiency development. This chapter comprises two sections. The first section documents the significant findings in order to derive a set of behavioural markers indicative of an ATCO's level of development. The second section is critical review of the work and reflects upon the interesting and useful insights and conclusions which may be reached.

10.1 CONSOLIDATION OF MARKER SET

Real-world application has helped to validate the assignment of markers against the appropriate learning and development classification, and provided reliability against individual markers which demonstrate significant change in presence and prevalence over time. During the research studies detailed in Chapter 6-9 a number of evidence based revisions have been made to the set of markers and observational sheet. Where appropriate, new markers were added, and conversely several markers were removed. Changes made to the set also included the consolidation of markers, and finally minor re-classification isolated individual markers. Appendix C09 contains a full timeline for each behavioural marker studied. A summary of the changes made to the set of markers is presented in Table 10.1, whilst Table 10.2 presents the final learning and development classification of markers from the end of Chapter 9.

Version	Amendments Made
Chapter 6 Changes (Changes made after 26 runs of the study)	The following markers were removed: • 'best practice' • interface scan' The following markers were added: • 'serial tasking' • 'general queries' • 'technical discussion'
Chapter 6 Changes Changes made at the end of the Chapter 6 study:	The following markers were combined: • 'Frustration' and irritation' (codec: 'verbal frustration') The following markers were re-assigned to the beginner category (from the intermediate category): • 'Self affirm' • 'Play/Sandpit' • 'Apologetic'
Chapter 7 Changes Changes made following inter-rater feedback:	The following markers were combined: • 'Muscle Memory', 'Automatic', 'Quickly Locates' (codec: Automatic Memory) • 'Plans Ahead', 'Keeps Info Open' (Codec: 'Plans ahead')
Chapter 8 Changes	The following markers were added during the first few observations made and included for the remainder of the study • 'Low Dexterity' • 'Clumsy' • 'Frantic approach' • 'Leans over' • 'Nervous Voice' • 'Nervous Physical' • 'Spatial reference'

Table 10.1 - A summary of the changes made to the set of behavioural markers over the different studies.

Final Classification of Behavioural Markers		
Beginner Behaviours	**Intermediate Behaviour**	**Expert Behaviour**
· Frustrated/irritated	· Positive comments	· Dual Tasking
· Affirm Before	· Social	· Quirks
· Affirm After	· Play/Sandpit	· Maintains Global SA
· Apologetic	· Periphery	· Automatic Memory
· Technical Discussion	· Input device tick off	· Plans Ahead
· Self Affirm	· Team Aware	
· Indecisive Hesitant	· Team Contribute	
· Low Dexterity	· Pace Fast	
· Clumsy	· Pace Slow	
· Frantic approach	· Fatigue	
· Leans over	· Adjust MMI	
· Nervous Voice	· Fidgets	
· Nervous Physical		
· Spatial reference		
· Negative comments		
· Confusion		
· Tongue-tied		
· Delays/Repeats		
· Overconfident		
· Slow/Hesitant		
· Incorrect actions		
· Surprise		
· Reactionary		
· Team Short/Snappy		
· Phys. Frustration		
· Serial Tasking		
· Verbal Frustration		

Table 10.2 - Final classification of markers

10.1.1 Key Findings

Table 10.3 presents the correlation results from across all of the research work undertaken. Where reliable weak to strong correlations were found in the intended direction, these are denoted by an ‘x’ in Table 10.3, Where reliable weak to strong correlations were found in the unintended direction, there are denoted by an ‘xx’ in Table 10.3. Not all of the behavioural markers observed (Chapters 648) resulted in significant change. In other words, in a number of cases no relationship has been shown between the frequency of prevalence between these behavioural markers and the amount of learning or exposure to a new system the ATCO has been exposed to.

Chapter 6 revealed a total of six individual markers demonstrating reliable significant weak to strong correlations in the unanticipated direction. The first three of these markers *'Apologetic'*, *'Self Affirm'*, and *'Play Sandpit'* were all anticipated to be strong transient behaviours presenting low prevalence at the start and end of the observational study, but peak somewhere in the middle, indicating a certain level of acquired knowledge and skill. All three of these markers were assigned the intermediate classification. The fourth marker, *'Pace Fast'* again was anticipated to increase however following the study in Chapter 9 the t4test results indicated that this marker is impacted by task load as is the case for the fifth and sixth markers ‘Dual Tasking’ and ‘*Plans Ahead*’.

Chapter 8 revealed one correlation in an unexpected direction. This was shown by individual trainee ATCOs where the marker ‘*Technical Discussion*’ demonstrated reliable strong positive correlations against two individual ATCOs (and a strong negative correlation against a third ATCO).

In order to derive a set of credible behavioural markers, a separation exercise has been undertaken. Of the 45 markers presented (Table 10.3), 25 were found to be unreliable, very limited, or unable to demonstrate significant change in presence and prevalence across the individual studies (these markers are shaded in pink). This results in 30 markers (shaded in green) which demonstrate supportive evidence for increased/decreased prevalence as learning and system exposure increases.

Behavioural Category	Behavioural Marker	Significant result				
		Experts (Chapter 6)	Trainees (Chapter 8)	Individual Trainees (Chapter 8)	Non-benign vs. benign (Chapter 9)	inter-rater agreement (Chapter 7, Table 7.10)
Attitude & Mood	Negative comments					High
	Positive comments					High
	Apologetic	x				Moderate
	Social					High
Comms & Verbal Commentary	Frustrated/Irritated			x		High
	Confusion	x		x		Moderate
	Tongue-tied	x				High
	Self Affirm	x				High
	Delays/Repeats	x	x	x		Moderate
	Standbys	x				High
	Indecisive/hesitant					
	Verb. Frustration					Low
Physical Posture & Body Language	Pace Fast	xx			x	High
	Pace Slow	x				Low
	Phys. Frustration					Moderate
	Fatigue				x	High
	Adjust MMI			x	x	Moderate
	Fidgets				x	Moderate
Inputs and Interaction with HMI and Workstation	Surprise	x	x	x		Moderate
	Overconfident					High*
	Slow/Hesitant	x	x			Low
	Play/Sandpit	x				High
	Dual Tasking	xx	x	x	x	Low
	Serial Tasking		x	x	x	High
	Periphery		x	x	x	High
	Maintains Global SA			x	x	
	Quirks					High
	Incorrect actions	x				Moderate
	Automatic Memory		x	x	x	Low-Mod.
	Plans Ahead	xx	x	x	x	Low-Mod.
	Low Dexterity		x	x		
	Clumsy		x			
	Frantic approach					
	Leans over					
	Nervous Voice					
	Nervous Physical					
	Spatial reference			x		
Interaction with others	Affirm Before	x		x		High
	Affirm After	x		x		High
	General Queries	x				blank
	Tech Discussion	x		xx	x	Moderate
	Reactionary			x		High
	Team Aware		x	x	x	Moderate
	Team Contribute		x	x	x	High
	Team Short/Snappy					High*

Table 10.3 - The collective reliable and significant results from Chapters 6-9. An (x) denotes all reliable correlations (weak to strong) from this research, (xx) denotes the direction is not as anticipated

10.1.2 Final set of markers

Table 10.4 presents the final 30 ATCO behavioural markers, whose presence and prevalence reliably and significantly changes through learning and increased exposure to a new flight strip system. These markers are considered to offer the ATC human performance researcher insight into levels of attained ATCO proficiency.

Categories	Beginner Behaviour	Intermediate Behaviour	Expert Behaviour
Attitude & Mood	• Apologetic		
Communications & Verbal Commentary	• Confusion • Tongue-tied • Delays/Repeats • Frustration/Irritation • Self-affirm		
Physical Posture & Body Language		• Adjust MMI* • Fidgets* • Fatigue* • Pace Fast * • Pace Slow*	
Inputs and Interaction with HMI and Workstation	• Surprise • Slow/Hesitant • Serial Tasking* • Incorrect Actions • Two hands to move strips • Clumsy • Low Dexterity • Play/Sandpit	• Periphery*	• Dual Tasking* • Automatic Memory* • Plans Ahead* • Maintains Global SA*
Interaction with others	• General Queries • Reactionary • Affirm Before • Affirm After	• Team Aware* • Team Contribute*	

Table 10.4 - Final set of behavioural markers (* t-test result denotes behaviour impacted by non-benign conditions)

The distribution of reliable and significant markers across the five categories is uneven, as is the distribution across the three development classes. The lack of symmetry is considered to reflect the complex and multi-dimensional nature of learning and development of complex tasks Patrick, (1992).

10.1.3 ATCO Phases of learning

The significant findings in Table 10.4 provide supportive behavioural evidence to the learning and development frameworks reviewed (Chapter 5, Table 5.9), where Fitts & Posner (1967), and Dreyfus (1980) in particular, suggest behaviour is quite rigid as novices acquire and develop core skills; but gradually becomes more free and expressive as they develop a growing understanding of the cause and effect of their actions; leading through to a global and holistic view of the situation, with intuitive and instinctive awareness and understanding.

It is interesting to note that all of the reliable correlations for intermediate and expert behaviours are impacted by the environment and conditions of the task (as shown through the study in chapter 9). This finding suggests that the display of complex and sophisticated NTS necessitates a task environment and situation which warrants their use.

Presenting these findings into a narrative, as per the descriptions of the other key learning and development frameworks reviewed in Chapter 5 (Table 5.9), the classifications are as follows:

Beginner: The initial practical application of knowledge in order to develop an understanding of the ATC system. This may be achieved through asking questions, reacting to prompts from instructors or colleagues, and testing out ideas and concepts to 'see how the system responds' and possibly showing signs of surprise at the result. There may be hesitancy, delay, or other intonation in the verbal communications suggesting confusion, and at times frustration and apology for any errors made. The application of prior knowledge and growing experience may also be displayed through verbal narrative (talking through the task); and looking for feedback from colleagues and instructors either before or after an action. Over time, these behaviours become less prevalent, indicating developing technical and non-technical skills and task strategies, as overall task proficiency increases.

Intermediate: Having developed a rudimentary understanding of the ATC system, the ATCO gains greater awareness of the global task and environment, response to peripheral stimulus is heightened including the awareness and active contribution to team activities. The ATCOs approach to the task becomes more variable, adapting more to the pace of the task both when quiet and busy and undertaking actions such as the adjustment of the MMI in order to benefit task delivery. Signs of fatigue and fidgeting may manifest themselves, as the ATCO looks to continue to push their performance and achieve more refined task delivery.

Expert: Having established good motor skills and the ability to approach the wider task with heightened global awareness and sensitivity, this final stage concerns advanced autonomous and efficient task delivery, including the use of multiple modalities (voice, writing, HMI inputs) concurrently in order to multi task, planning ahead, and ensure they remain one step ahead with the setup of the workstation and task situation in order to deliver optimal performance.

10.2 MAIN DISCUSSION

This section reviews the important research findings and comparisons are drawn with previous key research within the behavioural markers domain. This section has been structured following the identification, development and testing phases of the research which were undertaken in order to provide answers to the research questions 1-4[6].

10.2.1 The initial identification process

The preliminary study (Chapter 4) identified a number of behavioural markers using the five principal sources recommended by Flin et al (2008). Of the five sources employed, expert consultation, review of material from other domains, and direct observation were the most fruitful in identifying NTS, and associate markers. Least productive was the use of incident data, as the incident reports reviewed had very little detail on individual ATCO behaviour; a scenario experienced by other researchers (Fletcher et al, 2004).

Many of the behaviours suggested by the various experts consulted through the marker identification process (Source 4) did not generate explicit behaviours; instead they reflected desirable or undesirable character traits (Flin et al, 2008). In addition several of the markers were rarely seen as they were highly situation or environment specific.

Through practical application of the observation sheet developed, several further potential behavioural markers were identified (Appendix D02). It is clear from a review of the literature (Chapter 3, Table 3.30) that Flin et al's (2008) fifth source for behavioural markers identification has not been widely used. However the

---

6 **Research Question 1:** What non-technical behavioural markers may be used to evaluate ATCO performance?

Research Question 2: What phases of development are there, including transient stages?

Research Question 3: How might the presence or prevalence of certain non-technical behaviours be used to indicate how well a user is engaging and developing with a system?

Research Question 4: What situational factors may impact the presence and prevalence of certain behaviours?

validity of this source is unquestionable. Clearly correct selection and filtration, as advocated by Klampfer et al (2001) remains paramount, in order to filter out behaviours concerning individual difference, and those where no strong link with performance can be derived.

10.2.2 The scope of the marker set

The preliminary study (Chapter 4) was focused upon the identification of behavioural markers and the practical application of an observation tool base on this content. However, a key finding from this early research concerned the observation of a trainee ATCO, who was working towards sector validation (a process which can take several months of supervised live activity). The trainee, who was observed on two separate occasions, appeared to demonstrate a lack of confidence, and authority. Indeed his/her general approach to the task did not appear to be as skilful as his/her fellow ATCOs. This observation raised the question as to how overt behaviour might change over time, as confidence and expertise with the task increases.

Existing behavioural marker systems (Chapter 3) focus upon ascertaining NTS at the end of training or for the purposes of Continuous Professional Development (CPD). Here the observational focus is upon behaviours which demonstrate NTS competency. However it was theorised that NTS behavioural markers for earlier phases of learning and development might exist. Once identified these markers may be used to i) gauge how proficient a user is with the task they are undertaking as they progress through training, or ii) monitor existing experts transition between new ATC systems and learning new skills. This second use is particularly significant given the complex systemisation that is being undertaken within the ATC domain (Chapter 2).

The scope of this research therefore re-aligned towards the identification of behavioural markers that indicate different levels of learning and development; and to assess how these behaviours evolve, develop, and change over time as knowledge and proficiency increases. Research Questions 2 and 3, were generated as a result of this change in focus. A fourth Research Question was also generated as a result of the observational study in Chapter 4, which considered that specific task scenarios and environmental conditions appear to impact NTS.

A series of observations (Chapter 5) from groups of ATCOs with differing levels of experience with electronic flight strips identified a number of potential behavioural markers of differing levels of learning and development. Observation was chosen as the method partly due to the availability of access to undertake observation, but mainly as it was considered the most appropriate method to capture the subtle differences in behaviour that might exist.

10.2.3 The markers identified through observation

Three Observational activities were undertaken. The first explored Tower ATCO behaviour in a group with less than 3 hours of simulator training exposure to an electronic replacement to their paper flight strip system (EFPS). The second explored en-route ATCO behaviour in a group with less than 25 hours of simulator test and evaluation exposure to an electronic replacement to their paper flight strip system (iFACTS). The third explored Tower ATCO behaviour in a group with several years of experience using an electronic replacement to their paper flight strip system (EFPS). A number of behaviours were recorded and found to be independent of ATCO groups, whilst other behaviours recorded were observed in two or three of the ATCO groups (Appendix D02). Through consolidation and categorisation six NTS categories emerged:

1. Undertaking the task (task processes)
2. Attitude and Mood
3. Communications & Verbal Commentary
4. Physical Posture & Body Language
5. Interaction with others
6. Inputs & Interaction with the HMI and workstation

The category ‘Undertaking the task’, and two markers contained within it were removed after the first application (Chapter 6). They were removed because although they concerned aspects of best practice skill, the observer needed a high level of technical ATC task knowledge in order to identify and record them (e.g. the use of *aide memoire* strip management techniques).

The second category of markers (Attitude and Mood) appears at first inspection to break from the behavioural markers literature best practice convention (Klampfer et al, 2001). The markers within this category (Chapter 5, Table 5.4) relate to the ATCO's willingness to learn, an awareness of their lack of knowledge, and their level of engagement towards the new system they are learning. When considering the Behavioural markers content of other systems from the literature, the markers identified under 'Attitude and Mood' have a high degree of similarity. It is perhaps therefore an aspect of semantics, as to how the marker is coded that is of real significance.

In terms of the remaining categories, there are areas broadly consistent within the themes emerging across the 19 behavioural marker systems reviewed (Chapter 3, Table 3.31), although 'Physical posture and body language', and 'input and interaction with the HMI and workstation' contain markers which would more generally be categorised under the term 'Stress Management, Self-Care, Workload/Task Management'.

10.2.4 The learning and development class assignment

A review of the literature revealed that there are in fact very few learning and development frameworks. These frameworks dichotomise in terms of the number of phases, and the coverage of physical and mental skills they encompass, though they are all broadly similar in concept. A decision was taken to use a three-level framework (Beginner, Intermediate, Expert) as this was compatible with the three ATCO groups observed during the data, and the small number of development phases was considered to maximise opportunity for exclusivity for the process of assigning markers.

A series of observations were undertaken in order to explore and record how the presence and prevalence of these pre-identified behaviours changed over time through learning and increased system exposure. Three separate studies were undertaken in response to the emerging findings. The first study (Chapter 6) explored behavioural change in fully qualified ATCOs transitioning from paper to electronic flight strips. The second study (Chapter 8) explored behavioural change in ab-initio trainee ATCOs learning the controlling task and paper strips for the very first time. The final study (Chapter 9) gathered data within a high workload Olympics airspace simulation in order to explore how certain conditions might impact upon observable behaviours. An evidence based reclassification of the markers was made iteratively across the studies (Table 10.1). The final classification is illustrated in Table 10.2. A review of the key findings is presented later in this chapter.

10.2.5 The observation sheet and method

The initial observation sheet design (Appendix C01 - C02) was utilitarian and did not adopt the style convention of other systems. An error which was corrected for the subsequent observation sheets in order to improve usability. Versions 1-4 of the main observation sheet further simplified the layout, and the design was refined iteratively through repeated applications (Appendix C04 – C07). The observation sheet received positive feedback from the HF observers who used it (Chapter 7). This observation sheet has now been used to collect observations on 129 real-world occasions collecting data in four studies (Chapters 6-9).

As shown in this thesis, existing NTS behavioural observation methods employ the use of Likert scales to record frequency of performance (van Avermaete & Kruijsen, 1998, Fletcher, 2006; Yule, 2008; Oprins, 2008; Gatfield, 2008; Mitchell, 2012a). In a departure from this convention, observations were recorded using a frequency tally. The frequency tally recorded every occasion that a specific behavioural marker was observed.

The frequency count was chosen in order to gather data with a high degree of sensitivity, the greatest possible objectivity and mitigation of observer bias, and provide the maximum opportunity to monitor subtle behavioural changes over time. It is acknowledge however that the use of a frequency count is reliant to a certain extent on the observer being aware and maintaining in working memory markers which have reached saturation, in order to discount their additional inclusion. The frequency data has proved useful for the purposes of correlation analysis. However the use of an uncapped frequency count (used by the study in Chapter 4) was a mistake, and resulted in a large amount of 'heads down' time on behalf of the observer.

The studies in Chapters 6 and 7 employed a frequency cap set at five. This greatly improved the observation experience, although the cap was increased to 10 for the later studies (Chapters 8 and 9) as the initial limit was considered too low to afford sufficient perspective on subtle differences in prevalence within the data. Although the frequency tally has been a useful research tool, it is very labour intensive to make the observational recordings. A conventional Likert scale is considered to provide greater widespread utility, as employed by existing NTS systems.

10.2.6 Inter Rater Reliability (IRR) and Usability

Quantitative analysis using a spearman's Rho co-efficient to assess Inter-rater agreement for each pair of dual HF observations revealed, with the exception of one pair of inter rater correlations, reliable weak to moderate agreement between observers (0.460 – 0.680). Eiβfeldt (2003) and Oprins (2008) assessed the reliability of their behavioural observation systems using Pearson's R; correlation analysis using Pearson's R and Spearman's Rho afford broadly similar results (Swift, 2002; Gwet, 2012). Eiβfeldt (2003) obtained IRR between 0.2864 0.648 for 34 out of 35 individual competency areas, Oprins (2008) obtained IRR between 0.46 – 0.6 for individual competency areas. The Spearman's Rho reliability correlations of this research compare favourably with the results of Eiβfeldt (2003) and Oprins (2008).

A reliability analysis for each individual behavioural marker has not been possible due to the limited number of dual observations. Instead a simple analysis was undertaken following principles from Gatfield's (2008) inter-rater reliability assessment of crisis management behavioural markers, using absolute difference between scores as a means to evaluate levels of agreement.

Using a scoring threshold developed by the author (Chapter 7, Table 7.9), qualitative inter rater analysis was performed, the results indicate 25% of the markers receive a low level of agreement between observers, 32% markers receive a moderate level of agreement, and 43% markers receive a high level of agreement (Table 7.10).

It is anticipated that reliability would increase still further, if the training which preceded the dual observations included a practical opportunity to gain familiarity in the application of the method and to ground ratings and judgements between observers. Fletcher (2006) drew a similar conclusion, indicating that the strength of her inter-rater reliability was impacted by the limited amount of training which took place ahead of the observations. Klampfer et al (2001) advocate two full days of training ahead of practical application, although this is to train observers for the purposes of actual NTS evaluation, rather than to merely test Inter Rater Reliability.

A structured interview with the two HF expert observers recruited for the Inter Rater Reliability study provided feedback regarding the observation method, the design of the sheet, as well specific comment on individual behavioural markers. Both observers found the observation exercises valuable and rewarding, and were complimentary regarding the behavioural markers used, and the observational technique itself. Their feedback enabled a number of changes to the behavioural marker set and supporting guide (Appendix D07) in order to provide clarity on individual markers and their definitions, and consolidate the set as much as possible.

It is clear that a larger number of participant observers would have provided a more extensive evaluation of reliability. In addition, Neilson, (2006) suggest that 5 users are sufficient to provide a broad understanding of a system's utility and usability from a qualitative perspective. The view of two observers is therefore unlikely to have captured all of the significant usability aspects which could be made.

10.2.7 Limitations

There was a lack of control regarding the conditions surrounding nearly all of the observations taken. This included who was observed and their specific task position (radar or tower position), the length of the observation was often curtailed due to operational management (breaks, splitting and band-boxing sectors according to changes in traffic levels).

Permission was not obtained to record video at any time during this research, as unit management indicated the observations needed to be undertaken with the minimum amount of distraction. For a number of reasons this had a detrimental impact. Firstly, video may be used to record behaviour which may be analysed repeatedly on later occasions. This has the potential to provide additional opportunity to identify important behaviours missed through direct observation. Secondly, video has value both in the assessment of inter-rater reliability – where all observers may watch the same footage – but also in the training of observers where their scoring should be grounded ahead of practical observation.

A repeated limitation has been that permission was not obtained to observe ATCOs during high workload and high pressure circumstances. Understandably ATCOs prefer not be observed when working under challenging circumstances. However in order for the fullest range of behaviours to be displayed as possible, observations are needed when the situation is non-benign.

Although a training session was provided to the HF observers who participated in the inter-rater study (Chapter 7), it did not include a practical session to baseline scoring and familiarise the observer with the technique, ahead of the dual-observations themselves. Unfortunately video recording has not been permissible over the course of this research, an alternative therefore would be to use the first few dual-observation sessions as baseline training; and potentially discard any data collected ahead of the main study. However with an already small data set, further reduction would not have enabled the limited analysis which has been possible. The number of HF observers available to participate in the inter-rater study (chapter 7) was very small, this also impacted the depth of analysis and critical review that has been permissible.

10.2.8 Discussing the general results

The results of the first study (Chapter 6), which explored behavioural change in expert ATCOs transitioning from paper to electronic strips, was mixed. Several beginner behaviours revealed reliable correlations for decreasing prevalence over time. However the predicted changes of increasing prevalence for intermediate and expert behaviours did not occur. A number of reasons have been put forward for this (Chapter 6: Discussion). The two reasons considered most likely are that i) the intermediate and expert behaviours identified are flight strip system independent and transferred across to the new system without change, and ii) the task and environmental conditions were benign for the final live observations made in the study, and did not afford the opportunity for certain task skills and abilities to be demonstrated.

The second study explored changes in trainee ATCOs whilst undertaking a NATS aerodrome course with the intention of exploring behavioural development in a group with no previous ATC experience. It was anticipated that trainee ATCOs would not possess certain skills and behavioural traits indicative of an expert ATCO at the start of their course; and that their emergence over time would provide evidence to underpin the classification of markers in Chapter 5 (Table 5.10).

Across the three learning and development classes, beginner behavioural markers demonstrated a reliable reduction in prevalence over time, whilst intermediate and expert markers reliably increased in prevalence. In addition, several new behaviours were recorded, and selected for inclusion across the remainder of the study. Of these ‘clumsiness’ and ‘low dexterity’ revealed reliable weak correlations, indicating improving manual dexterity over time. Further data analysis revealed that several individual beginner behavioural markers demonstrated reliable weak to strong correlations across the training course. In addition, and most significantly, several individual behavioural markers from the intermediate and expert classes also demonstrated reliable moderate to strong correlations.

Chapter 9 presents the results of further observations undertaken during a high intensity TC Real Time Olympics Airspace simulation. This data has been compared to the final eleven observations made in the live environment with EFPS (where the task and environmental conditions were extremely benign). The results of this data comparison indicate that several markers associated with task delivery and situation awareness appear more frequently in non-benign conditions, than benign.

Collectively these findings provide supportive evidence that certain NTS behavioural markers demonstrate a relationship with technical proficiency and development. Furthermore, certain overt behaviour is impacted by task and environmental factors, and as a consequence must be taken into consideration when undertaking observations. Tables 10.3 and 10.4 present the behavioural markers which demonstrate change in prevalence over time.

10.2.9 Discussing the key findings

The intention of this research was to identify behavioural indicators of an ATCOs attained learning and development level. However the observational results paint a more colourful picture.

Firstly, the results show that behavioural change can prove extremely subtle and hard to detect and monitor. The implication of this result, in terms of the development of a simple observational tool to study subtle changes in NTS behaviour, is therefore limited.

Secondly, behavioural change is complex. ATCOs have been shown to display behaviours from all three learning and development categories concurrently. In other words, a single person may display a mixture of beginner, intermediate, and expert behaviours all at the same time. In terms of the practical implications of this research, it is the combination of markers present at one moment in time, and their respective quantities, which afford an overall indication of proficiency attained.

Thirdly, a key differentiator between the behavioural markers contained within the developed classification framework produced, and markers from other systems is the inclusion of 'beginner' behaviours. Reflecting on the 19 beginner behaviours which demonstrate reliable weak to strong Spearman's Rho correlations (Table 10.4), there is a great deal of similarity to a number of behavioural markers contained in several of the behavioural observation systems reviewed (Table 10.5). The markers in Table 10.5 have been classified as 'poor behaviours', 'poor practice', even 'negative behaviour'. It is argued that a number of the markers assigned under this term are in fact indicators of beginner behaviour. It is important to note that this does not change the desire for a person to demonstrate as few of these as possible in a NTS evaluation!

Behavioural Observation System	Behavioural Markers of Poor performance / Bad performance
ANTS:	• Does not alter physical layout of workspace to improve data visibility
NOTSS:	• Asks questions which demonstrate lack of understanding
	• Overconfident manoeuvres with no regard for what may go wrong
	• Selects inappropriate manoeuvre that leads to complication
	• Becomes hasty or rushed due to perceived time constraints
	• Engages in 'tunnel vision' approach to technical aspects of operation
	• 'Freezes' and displays inability to make decisions under pressure
SPLITS:	• Asks questions that indicate a lack of understanding
	• Loses track of surgical activity, i.e. is caught unaware
	• Fails to communicate in a clear and precise manner
	• Shows a lack of understanding of instrument purpose or sequence of usage
	• Raises voice unnecessarily
	• Loses temper/ displays emotional outbursts
	• Waits for instruction when should take action
Maritime Crisis Mgt:	• Communications inaudible and garbled.
	• Is surprised by system interactions.
	• Acts immediately on suggestions from team members without any prior reflection.
	• 'Grasps' at suggestions of others.
	• Movements hesitant.
RSSB:	• Unclear communications e.g. provides ambiguous or vague information
	• Waffles in communication
	• Rude or aggressive in dealing with others e.g. uses offensive language
	• Appears to be unable to do more than one task at a time
	• Does not maintain an appearance of being calm and in control when under pressure e.g. raises voice, becomes aggressive
	• Lacks confidence/ initiative to work without over-checking e.g. regularly checks things when inappropriate

Table 10.5 - Markers identified from the 19 Behavioural Marker Systems reviewed that could be considered as indications of a beginner.

What the research undertaken in this thesis shows, is that the presence and prevalence of certain negative or 'beginner' behaviours decrease in prevalence against system exposure time as NTS generally improve. Therefore certain markers contained within existing NTS behavioural observation systems may provide further indication as to the general stage of learning and development a person has reached, through subtle changes in prevalence.

10.2.10 The utility and practical implications of this work

Tables 3.4 and 3.5 (Chapter 3) present respectively Klamper et al's (2001) and Flin et al's (2008) principles for an effective behavioural marker system. These principles provide a useful means by which the output of this research may be evaluated. Tables 10.6 and 10.7 present a critical assessment of the works undertaken by this research 4 using Klamper et al's (2001) and Flin et al's (2008) principles 4 to develop a set of behavioural markers and underpinning tool that provide insights into the level of learning and development an ATCO has achieved with the task, as well as adapting and updating their skills to work with replacement ATC systems.

The 30 markers contained in Table 10.4 all demonstrate reliable weak to strong correlations for changing prevalence against total system exposure time. Collectively they provide an original means of tracking and monitoring subtle changes in the behaviour of ATCOs, as their levels of proficiency in the task matures with new ATC systems. However, the limited inter-rater reliability assessment, in addition to more wide spread testing in a broader set of operational conditions (high workload, equipment failure, fallback and emergency situations) stops it short as being an 'off the shelf' tool to undertake NTS assessment.

Design Principles Area	**Critical Assessment**
Training	The Human Factors observers received an extensive briefing ahead of the dual observations; however it was not possible to include a practical component to the training ahead of the dual observations. The results suggest a learning effect resulted (Chapter 7, Table 7.7) as both observers became familiar with the structure of the observational tool and method to record behaviours.
Domain Specificity	Both the preliminary markers identified (Chapter 4) and the subsequent markers observed and selected for the main set (Chapter 5) were developed using ATC domain specific sources (Sources 1-5).
Implementation	The marker system developed has been designed specifically as a research tool to better understand changes in NTS behaviour during training and the acquisition of new knowledge and skills. The work stops short of developing an observational tool end-product, due to limited inter-rater evaluation.
Purpose	The observational studies undertaken were principally focused upon the subtle changes of behaviour during ATC college and ATC system transition training. The research employed a capped frequency scale to capture subtle change in NTS behaviour as it changed over increased training and system exposure. The use of a frequency tally enabled subtle changes in prevalence to be recorded, and this value was iteratively refined through progressive applications.
Environment	Research question 4 specifically focused upon the identification of situational factors which may impact the presence and prevalence of certain behaviours. High workload and task complexity were found to impact the following behaviours: 'Adjust MMI' 'Fidgets', 'Fatigue', 'Pace Fast', 'Pace Slow', 'Periphery', 'Dual Tasking', 'Automatic Memory', 'Plans Ahead', ' Maintains Global SA', 'Team Aware', 'Team Contribute', 'Serial Tasking'.
Sensitivity	The preliminary set of markers identified in Chapter 4 was explicitly selected in order to provide indications of good or bad NTS. This selection was as a result of SME marker contributions, and the expertise of the author. For the main studies (Chapters 5-9) a second set of markers was identified through observational activities, however unlike other existing marker systems, markers we selected when considered to provide indications as to the level of development an ATCO has attained during periods of learning and development. Through evidenced based evaluation, a final set of markers was derived which show reliable change in prevalence over time, indicating differing levels of learning and development attained.

Table 10.6 Critical Review against principles of an effective Behavioural Marker System – Part 1 (Klampfer et al, 2001; Flin et al, 2008)

Design Principles Area	Critical Assessment
Reliability	Limited inter-rater reliability testing was undertaken using version 2 of the observation sheet developed in Chapter 5 and applied during Chapter 6. The results of this testing enabled a number of revisions to the marker set to be made in order to improve consistency and clarity. The marker set was subsequently used in a further two studies (Chapters 8, and 9) with further iterative refinement to the set as a result. More widespread inter-rater reliability testing is required before the assessment of individual markers can be definitive. Therefore the output of this work is towards the development of a research tool, rather than that of a fully refined observational tool end product.
Validity	The behavioural markers of learning and development produced by this research have been identified exclusively through direct observation or real-world end users in simulation and live operational environments. Individual markers have been selected through evidence based review, where clear change in prevalence can be shown between behaviour and amount of system exposure time. Qualitative review has been provided during inter-rater reliability testing (Chapter 7). The final set of derived markers has a good fit against the phases of learning and development contained within the literature (Table 10.4). Behavioural marker prevalence and task performance have not been assessed however competency assessment undertaken by unit and ATC college trainers indicated attained proficiency across all qualified ATCOs and all but one of the trainee ATCOs.
Structure	The behavioural marker set has been developed through a process of iterative development and consolidation using qualitative feedback and application experience from several observational studies. This process has enabled evidence based consolidation in order to derive as fewer markers as possible that provide an indication of developing NTS proficiency. A supporting guide provides the observer with a reference to explain any uncertainty or ambiguity over certain markers.
Usability	The marker framework and supporting guidance sheet were reviewed through post-dual observation interview. The qualitative feedback received was generally positive regarding the simple and uncluttered observations sheet. The sheet itself was iteratively developed, and designed for simplicity following experiences with an early observation sheet (Chapter 4) which was found to be difficult to use due to too much text on display. The observational methodology developed from un-capped to capped frequency tallies, in order to maximise the amount of observer heads-up time. The method proved useful for the purposes of recording subtle changes in ATCO behaviour, analysis reveals that considerable amounts of data are required in order to draw clear conclusions. Therefore although useful as a research instrument, the application of this observational tool to gather routine behavioural change data during periods of learning and development is prohibitive.

Table 10.7 Critical Review against principles of an effective Behavioural Marker System – Part 2 (Klampfer et al, 2001; Flin et al, 2008)

SUMMARY OF CHAPTER 10

The first half of this chapter collectively reports all of the significant and important findings regarding behavioural changes observed through the longitudinal studies conducted. The key output of this work is the evidence based collation of markers all of which have been shown to vary in presence and prevalence as a result of increased training and system exposure. Half of this chapter provides a critical review of the research. It begins by exploring the approach undertaken in the preliminary study and the significant findings. Next it discusses the generation of further research questions and the works undertaken through a series of studies in order to provide answers to these questions. The methodological approach has been reviewed, exploring the key strengths and weaknesses, with references back to the literature where appropriate. A final discussion specifically explores the beginner behaviours identified, and draws strong parallels to several ‘poor behaviour’ markers contained within existing behavioural observation systems. The implication of which is that NTS observational systems may benefit from a framework that considers both unlearnt skill and weak behaviour at one end, and learnt skill and strong behaviour at the other.

CHAPTER 11 – CONCLUSION AND FUTURE RESEARCH

"I keep six honest serving-men, They taught me all I knew; Their names are What and Why and When And How and Where and Who."
(Rudyard Kipling)

Over the course of this work, four research questions have been generated, and a number of research activities have been undertaken for the purpose of answering these questions. A concluding answer to each of the questions is now presented.

Research Question 1 - What non-technical behavioural markers may be used to evaluate ATCO performance?

The works undertaken in Chapter 4 identified through conventional means, a set of NTS behavioural markers comparable to other benchmark behavioural observation systems. However, this work was preliminary and further effort is needed to evaluate and validate the set before a practical tool is derived for the works undertaken. The key output of this preliminary research has been the identification of three further research questions. These have been the main focus of this thesis.

Research Question 2 - What phases of development are there, including transient stages?

A review of the unexpectedly limited literature has revealed very few examples of learning and development frameworks. There is a dichotomy of views in the content and application of the various learning and development frameworks which have been created. These frameworks are extremely theoretical; represent an inexact science, and show particular weakness regarding the boundaries between each phase and subsequent difficulty in segmenting skills and knowledge accordingly Patrick (1992). The linkage with associated overt behaviour is also limited, focusing either on cognitive knowledge acquisition, or upon physical dexterity.

A simple and novel three level learning and development framework (beginner, intermediate, expert) was produced in the absence of a suitable system identified in the literature. A set of behavioural markers identified through direct observation was classified using this framework, and a number of observational studies were undertaken in order to evaluate change over time.

The observational results indicate that changes in behaviour for markers assigned to all three categories occurred concurrently. The conclusion reached is that although the presence of certain individual behaviour over time provides indications of learning within the complex multi-dimensional tasks such as Air Traffic Control, an Air Traffic Controller (ATCO) may display a dichotomy of behaviour not conducive to the assignment of an overall attained phase of learning. In other words, this research suggests that there does not appear to be clearly distinct phases, instead there is complex multi layered approach to the acquisition of cognitive and motor skills.

Research Question 3 - How might the presence or prevalence of certain non-technical behaviours be used to indicate how well a user is engaging and developing with a system?

This research has identified a number of behavioural markers which provide indirect indications of emerging system proficiency. Although no direct comparison with task performance metrics was made, all trainee ATCOs and ATCOs undergoing transition training were evaluated for competency at the end of their training. The evidence gathered indicates that the 30 markers presented in Chapter 10 (Table 10.3) show an indirect link with levels of task and system proficiency an ATCO has attained. Most significantly, on a simplistic level, the research reveals that the absence of 'beginner behaviour' is a useful pragmatic indicator for ATCOs gaining system proficiency.

Research Question 4 - What situational factors may impact the presence and prevalence of certain behaviours?

Invariably the observations conducted during this research have been in benign conditions, there has therefore only been limited opportunity to ascertain the impact of environment and situation factors upon the behaviours under investigation. However a small study was undertaken to explore behaviour in the high traffic, complex environment of an Olympics airspace simulation. The results suggest that high workload and task complexity impact the presence and prevalence of several behaviours associated with physical posture and body language, the level and type of interaction with the HMI, and interactions made with other team members. The important implications of these results, are that a NTS observer must consider the situation factors and environmental conditions with an observation made, as they might have a direct impact on the behaviour which is displayed.

CONTRIBUTION TO KNOWLEDGE

1. For the academic behavioural markers researcher, this research has explored changes in NTS within the domain of ATC, and determined that it is overly simplistic to anticipate that an ATCO will step through a number of learning and development phases in a clear and unidimensional way. Instead this research has shown that an individual might display a multidimensional combination of beginner, intermediate, and expert behaviours during the same observation session, it is therefore the quantity and mixture of the behaviours on display that affords an overall understanding of proficiency.

2. As a product of this research, a novel learning and development framework has been created. This research has highlighted that very few learning and development models are available in the academic literature. Across the small numbers which have been produced, there is considerable diversity in depth, structure, and content. One repeated consistency however is the limited amount of material regarding underpinning behaviours. Through an evidence based approach, this framework has been populated with a number of indicative NTS behavioural markers of relevance to the ATC domain.

3. This research has explored the utility of a frequency count in order to record individual occurrences of specific pre-defined behaviours. This approach

represents a key difference to Likert scales used by other comparable benchmark NTS observational systems. The findings reveal that it is an effective means by which to monitor subtle change. However the large amount of observations and extensive analysis needed, particularly when used in uncontrolled conditions, makes it more suitable as research tool.

FUTURE RESEARCH

The depth of conclusions has been constrained at times by the uncontrolled conditions the data has been collected within, and the amount of observation that has been permitted. Further observation in non-benign conditions, additional observations by a larger number of HF observers to facilitate deeper inter-rater reliability analysis, and practical application as an informative framework by training instructors and validation experts would deepen the level of understanding within this area.

There is clearly a research gap to explore behavioural change in other domains, where simpler tasks and greater control over the conditions could afford greater insight into how someone's overt behaviour changes as their skill proficiency increases over time. Such research could provide additional underpinning material for learning and development frameworks which would significantly enhance the literature in this field.

CONCLUSION

Through an iterative observational process, a number of ATC specific behavioural markers indicative of different levels of learning and development have been identified and placed within a novel framework. Their subtle changes in prevalence have been monitored in a number of ATC environments, observing both expert ATCOs transitioning between systems, and Trainee ATCOs undergoing an aerodrome training course. The research findings indicate that learning within the complex multi-dimensional tasks such as ATC results in a dichotomy of behaviour indicative of various phases of learning. In other words, there does not appear to be clearly distinct phases, instead there is complex multi layered approach to the acquisition of cognitive and motor skills. However through the observation of subtle changes in their prevalence, these behavioural markers afford insights into ATCOs levels of emerging proficiency.

REFERENCES

Ahles, T.A., Coombs, D.W., Jensen, L., Stukel, T., Keefe, F.J. (1990). Development of a Behavioral Observation Technique for the Assessment of Pain Behaviors in Cancer Patients. Behavior Therapy Vol. 21, 449-460.

van Avermaete, J. & Kruijsen, E. (1998). NOTECHS. The Evaluation of Non-Technical Skills of Multi-Pilot Aircrew in Relation to the JAR-FCL Requirements. Final Report NLR-CR-98443. Amsterdam: NLR.

Barbarino, M., & Patterson, I. (2007). NOSS Trial Report. Public Final Version. Finavia: Eurocontrol.

Bartunek, J.M., & Murnighan, J.K. (1984). The nominal group technique: Expanding the basic procedure and underlying assumptions. Group and Organization Studies, Vol. 9, 417-432.

Bessone, L., Coffey, E., Inoue, N., Gittens, M., Mukai, C., O'Connor, S., et al. (2008a). International Space Station Human Behavior & Performance Competency Model. Vol. 1. Mission Operations Directorate. NASA/TM-2008-214775 Vol1. ITCB HBP Training Working Group.

Bessone, L., Coffey, E., Filippova, N., Greenberg, E., Inoue, N., Gittens, M., et al. (2008b). International Space Station Human Behavior & Performance Competency Model. Vol. 2. Mission Operations Directorate. NASA/TM-2008-214775 Vol2. ITCB HBP Training Working Group.

Boeing (2001). Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations 1959 – 2001. Seattle: Boeing Commercial Airplane. Retrieved December, 2012, from

www.oaviao.com.br/oaviao_novo/diretorio_aero/seguranca_voo/Boeingacc identstatsum59-01 .pdf

Bonsall, K. (2012). Non-technical skills for rail: A list of skills and behavioural markers for drivers, with guidance notes (2012). Rail Standards Safety Branch.

Bonsall-Clarke, K. (2012). Non-technical skills required in train driver role: Developing an integrated approach to NTS training and investment. Rail Standards Safety Branch.

Brokaw. J.P., Walker. W.C., Cifu. D.X., Gardner. M. (2004) Sitting and standing tolerance in patients with chronic back pain: comparison between physician prediction and covert observation. Archives of Physical Medicine and Rehabilitation, Vol. 85, Issue 5. 837-839

Borman, W.C., Hedge, J.W., Hanson, M.A., Bruskiewicz, K.T. (2001) Development of Criterion Measures of Air Traffic Controller Performance. DOT/FAA/AM-01/6, Atlantic City: FAA. Retrieved, July 2010, from

www.hf.faa.gov/docs/CAMI/0106.pdf
www.hf.faa.gov/docs/CAMI/0106Apen.pdf
www.hf.faa.gov/docs/CAMI/0106Figs.pdf

Brooker, P. (2005). Reducing mid-air collision risk in controlled airspace: Lessons from hazardous incidents. Safety Science. Vol. 44, 715–738.

Brooker, P. (2008).The Überlingen accident: Macro-level safety lessons. Safety Science. Vol. 46, Issue 10, 1483–1508

CAA (2002). Flight crew training: Cockpit Resource Management (CRM) training and Line Oriented Flight Training (LOFT). (Report CAP720). UK: Civil Aviation Authority

CAA (2006a). Crew Resource Management (CRM) Training. Guidance for Flight Crew, CRM Instructors (CRMIs) and CRM Instructor-Examiners (CRMIEs) (CAP 737. Version 2.) Gatwick: Safety Regulation Group, Civil Aviation Authority.

CAA (2006b). Guidance Notes for Accreditation Standards for CRM Instructors and CRM Instructor Examiners. Standards. (Doc. 29 Version 2.) Gatwick: Civil Aviation Authority.

CAA (2012) CAP 493: Manual of Air Traffic Services - Part 1. (CAP493 Part1). Gatwick: Civil Aviation Authority.

Chemers, M. (1997) *An integrative theory of leadership*. Lawrence Erlbaum Associates, Publishers Mahwah, New Jersey.

Creswell, J.W. (2003). *Qualitative, Quantitative, and Mixed Methods Approaches*. Second Edition. Sage Publications Ltd. London.

Cushing, S. (1997). *Fatal Words; Communication Clashes and Aircraft Crashes*. The University of Chicago Press, Ltd. London.

Dreyfus, S.E., Dreyfus, H.L. (1980). A Five-Stage Model Of The Mental Activities Involved In Directed Skill Acquisition. (USAF Contract F496204 794C40063). Operations Research Centre. University of Berkley.

Duke, G. (1997). *Air Traffic Control*. Ian Allan Publishing Ltd.

Eiβfeldt. H (2003). Second Eurocontrol Selection Seminar, 17420 November 2003. Luxembourg: Eurocontrol IANS.

Endsley, M. R. (1995). Toward a theory of situation awareness. Human Factors, Vol. 37, Issue 1. 32464

Endsley, M.R., Farley, T.C., Jones, W.M., Midkiff, A.H., Hansman, R.J. (1998). Situation Awareness Information Requirements for Commercial Airline Pilots. (ICAT-98-1). International Center for Air Transportation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology.

Endsley, M. R. (2000). Direct Measurement of Situation Awareness: In M.R. Endsley, D.J. Garland (Eds.) *Situation awareness analysis and measurement.* (pp147-173). Lawrence Erlbaum Associates: London.

Eurocontrol (1996). Model for Task and Job Descriptions of Air Traffic Controllers. (HUM.ET1.ST01.1000-REP-01). Retrieved, February 2013, from

www.eurocontrol.int/sites/default/files/content/documents/nm/safety/safety-model-for-task-and-job-descriptions-of-air-traffic-controllers.pdf

Eurocontrol (2003). Behaviour Oriented Observation Method (BOOM): Human Factors Business Management Division. Retrieved, July 2008, from

www.eurocontrol.int/humanfactors/public/standard_page/BOOM.html

Eurocontrol (2005). Behaviour Observation Scale (BOS) Form for Air Traffic Controller Trainees.

Eurocontrol (2012). United Kingdom AIP. ENR 1.441. Retrieved, December 2012 from

www.ead.Eurocontrol.int/eadbasic/pamslight-9F14679403A464909296961B7547FA87/7FE5QZZF3FXUS/EN/AI P/ENR/EG_ENR_1_4_en_2012-12-13.pdf

FAA (1993). Advisory Circular 120-51A – Crew Resource Management Training. Retrieved, November 2008, from

http://.rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/0/47b3fc2632397849862573860046 5a 1 d/$FI LE/AC%20 120-51 A. pdf

FAA (2004). Advisory Circular 120-51E – Crew Resource Management Training. Retrieved, November 2008, from http://rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/list/AC%20 120-51 E/$FI LE/AC 120-51 e.pdf

FAA (2006). Advisory Circular 120-90 – Line Operations Safety Audits. Retrieved, November 2008, from

http://homepage.psy.utexas.edu/homepage/group/helmreichlab/

Fitts, P.M., & Posner, M.I. (1967). *Human Performance*. Oxford, England: Brooks-Cole

Flanagan J. (1954). The Critical Incident Technique. Psychological Bulletin. Vol. 51, Issue 4. 327-358.

Fletcher, G., McGeorge, P., Flin, R., Glavin, R., and Maran, N. (2003). Anaesthetists' non-technical skills (ANTS). Evaluation of a behavioural marker system. British Journal of Anaesthesia, Vol. 90, 580-588.

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N. and Patey, R. (2004). Rating non-technical skills. Developing a behavioural marker system for use in anaesthesia. Cognition, Technology and Work. Vol. 6, 165-171.

Fletcher, G.C.L. (2006). Development of a behavioural marker system for anaesthetists' non-technical skills. Thesis (Ph.D.). Aberdeen University.

Flin, R., Goeters, K.M., Hörmann, H.J., and Martin, L. (1998). A Generic Structure of Non-Technical Skills for Training and Assessment (14-18th September 1998). 23rd Conference of the European Association for Aviation Psychology, Vienna.

Flin. R., & Martin. L. (2001). Behavioural markers for Crew Resource Management: A Review of Current Practice. International Journal of Aviation Psychology. Vol. 11, 95-118.

Flin, R., Martin, L., Goeters, K.M., Hörmann, H.J., Amalberti, R., Valot, C. et al. (2003). Development of the NOTECHS (Non-Technical Skills) system for assessing pilots' CRM skills. Human Factors and Aerospace Safety. Vol. 3, Issue 2, 95-117.

Flin, R., Fletcher, G., McGeorge, P., Sutherland, A., Patey, R. (2003). Anaesthetists' attitudes to teamwork and safety. Anaesthesia Vol. 58. p233–242.

Flin. R. (2004). Assessment of Non-technical Skills: Lessons from the flight deck. Scottish Clinical Skills Network, Glasgow (17th September, 2004). Retrieved, March 2008, from

www.scsn.scot.nhs.uk/conferences/2004/docs/rfClinicalskills2004.ppt

Flin. R., O'Connor. P., Crichton. M. (2008). *Safety At The Sharp End. A Guide to Non-Technical Skills*. Ashgate Publishing Ltd.

Flin, R. (2010). CRM (non-technical) skills Applications for and beyond the flight deck. In R. Helmreich, B. Kanki & J. Anca (Eds.) *Crew Resource Management* (2nd Ed). New York: Elsevier.

Flin. R., Patney, R. (2011). Non-technical skills for anaesthetists: developing and applying ANTS. Best Practice & Research Clinical Anaesthesiology. Vol. 25. 215-227.

Fox, N.S., Brennan, J.S., Chasen, S.T., (2008). Clinical Estimation of Fetal Weight and the Hawthorne Effect. European Journal of Obstetrics & Gynaecology and Reproductive Biology. Vol.141, Issue 2. 111–4.

Furuya, K. (1996). *KODO Ancient Ways. Lessons in the Spiritual Life of the Warrior/Martial Artist.* Ohara Publications Inc.

Gaba, D.M., Howard, S.K., Flanagan, B., Smith, B.E., Fish, K.J., Botney, R. (1998). Assessment of Clinical Performance during Simulated Crises Using Both Technical and Behavioral Ratings. Anaesthesiology. Vol. 89. Issue 1, 8–18

Gatfield. D. (2005). Using simulation to determine a framework for the objective assessment of competence in maritime crisis management. SAGSET 2005 Conference. Retrieved, August 2008, from

www.solent.ac.uk/mhfr/resources/SAGSET%202005.pdf

Gatfield D. (2008). Behavioural Markers for the Assessment of Competence in Crisis Management. (PhD Thesis). Nottingham Trent University, and Southampton University.

Gordon Training International (2011). Learning A New Skill is Easier Said Than Done. Retrieved, October 2011, from

www.gordontraining.com/free4workplace4articles/learning4a4new4skill4is4 easier-said-than-done/

Gordon. N.J., Fleisher. W.L (2011) Effective Interviewing and Interrogation Techniques (Third Edition). London: Academic Press

Guttman, J.A., Stein, E., Gromelski, S. (1995). The Influence of Generic Airspace on Air Traffic Controller Performance. (FAA Report DOT/FAA/CT-TN95/38. Retrived, October 2008, from

http://hf.tc.faa.gov/technotes/dot4faa4ct4tn95438.pdf

Gwet KL (2012). *Handbook of Inter-Rater Reliability: The definative guide to measuring agreement among raters*. Third edition. Gaitherburg, Maryland: Advanced Analytics, LLC.

Hedge, J., Borman, W.C., Hanson, M.A., Carter, G.W., Nelson, L.C., (1993). Progress toward development of ATCS performance criterion measure. (Institute Report No 235). Minneapolis: Personnel Decisions Research Institutes, Inc.

Helmreich. R.L., & Wilhelm, J. (1987). Reinforcing and Measuring Flightcrew Resource Management: Training Captain/Check Airman/Instructor Reference Manual. (NASA/University of Texas Technical Manual 8741). Austin: The University of Texas.

Helmreich, R.L., Butler, R.E., Taggart, W.R., Wilhelm J.A. (1994). The NASA/University of Texas/FAA Line/LOS Checklist: A behavioural marker-based checklist for CRM skills assessment. (Aerospace Crew Resource Management Project Technical Report 94404). Revised, 1995. Austin: University of Texas.

Helmreich, R.L., Butler, R.A., Taggart, W.R., Wilhelm. J.A. (1995). Behavioral Markers in Accidents and Incidents: Reference List.. (NASA/University of Texas FAA Aerospace Crew Research Project Technical Report 95-1). Austin: University of Texas.

Helmreich, R.L., Foushee. H.C., (2010) Why CRM? Empirical and Theoretical Bases of Human Factors Training. In Barbara G. Kanki, Robert L. Helmreich, José M. Anca (Eds.). *Crew Resource Management: Second Edition*. London: Academic Press.

Helmreich, R.L., Merritt, A.C., Wilhelm, J.A. (1999). The Evolution of Crew Resource Management Training in Commercial Aviation. International Journal of Aviation Psychology, Vol. 9. Issue 1, 19-32.

Hollnagel, E., and Amalberti, R. (2001). The Emperor's New Clothes or Whatever Happened to "Human Error"? Proceedings of the Fourth International Workshop on Human Error, Safety, and Systems Development (HESSD-OI). June 11-12,2001: Linkoping, Sweden

Industrial Psychology Research Centre, Aberdeen University (2003). Framework for Observing and Rating Anaesthetists' Non-Technical Skills (ANTS) Handbook. Version 1.0. Retrieved, November 2008, from

www.abdn.ac.uk/iprc/ants/

Industrial Psychology Research Centre, Aberdeen University (2006). NOTSS Handbook Version 1.2. Retrieved, November 2008, from

www.abdn.ac.uk/iprc/notss/

Industrial Psychology Research Centre, Aberdeen University (2010). Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS) Handbook. Version 1.0. Retrieved, November 2008, from

www.abdn.ac.uk/iprc/splints

Jeannot, E., Kelly, C., Thompson, D. (2003). The development of Situation Awareness Measures in ATM Systems, (HRS/HSP-005-REP-01), Bretigny-sur-Orge: Eurocontrol Experimental Centre. Retrieved, November 2012, from

www.eurocontrol.int/humanfactors/gallery/content/public/docs/DELIVERABLES/HF35-HRS-HSP-005-REP-01withsig.pdf

Jensen, R.C., & Askren, W.B. (1998). Measuring the Workload of Operators in Automated Plants. Human Factors and Ergonomics Society Annual Meeting Proceedings, Posters. 42nd Annual Meeting. Chicago.

Khurana, A. (2009). *Scientific Management; A Management Idea to Research A Mass Audience*. New Delhi: Global India Publications Pvt Ltd..

Klampfer, B., Flin, R., Helmreich, R.L., Häusler, R., Sexton, B., Fletcher, G., et al. (2001). Enhancing performance in high risk environments: recommendations for the use of behavioural markers. Ladenburg: Daimler-Benz Shiftung. Retrieved, December 2008, from

http://homepage.psy.utexas.edu/homepage/group/HelmreichLAB/Publications/pubfiles/Pub262.pdf

Klinect, J.R., Murray, P., Merritt, A., Helmreich, R. (2003). Line Operations Safety Audit (LOSA): Definition and operating characteristics. *Proceedings of the Twelfth International Symposium on Aviation Psychology* (pp. 663-668). Dayton, OH: The Ohio State University. Retrieved, December 2008, from

http://homepage.psy.utexas.edu/homepage/group/helmreichlab/Publications/pubfiles/Klinect.Operating. Characteristics.2003. pdf

Kontogiannis, T., Malakisis, S. (2013). Strategies in coping with complexity: Development of a behavioural marker system for air traffic controllers. Safety Science. Vol. 57. 27-34.

Kornblum, W., (2011). *Sociology in a Changing World; Ninth Edition*. Belmont CA: Wadsworth Cenage Learning.

Kumashiro. M., (2005). Practical Measurement of Psychophysiological Functions for Determining Workloads. In John. R. Wilson, E. Nigel Corlett (Eds). Evaluation of Human Work: A practical Ergonomics Methodology. Third Edition. (pp605-629). London: Taylor & Francis Ltd.

Hörmann, H.J., & Neb, H. (2004) From NOTECHS to LH Behaviour Markers. An Implementation Case Study. Human Factors Group Seminar on Assessment & Accreditation. (April 30, 2004). London: Royal Aeronautical

Society

International Civil Aviation Organization (ICAO) (2002). Line Operations Safety Audit (LOSA). Retrieved, May 2008, from

www.icao.int/ANB/humanfactors/LUX2005/Info-Note-5-Doc9803alltext.en.pdf

Isaac. A. (2007). Assessment of Non-Technical Skills and Behavioural Markers. (Positive Performance Indicators V.4). Farnham: NATS

Labus, J.S., Keefe, F.J., Jensen M.P. (2003). Self-reports of pain intensity and direct observations of pain behaviour: when are they correlated? Pain. Vol. 102, 109-124

Lawler, E.E. (1967). The multitrait-multirater approach to measuring managerial performance. Journal of Applied Psychology. Vol. 51, 369-381.

LeBreton, J.M., Senter, J.L. (2008). Answers to 20 Questions about Interrater Reliability and Interrater Agreement. Organisational Research Methods. Vol. 11, Issue 4. 815-852

Lindegård, A., Karlberg. C., Wigaeus Tornqvist E., Toomingas, A., Hagberg, M. (2005). Concordance between VDU-users' ratings of comfort and perceived exertion with expert's observations of workplace layout and working postures. Applied Ergonomics. Vol. 36, Issue 3. 319-325.

Liou, Y. I. (1998). Expert System Technology: Knowledge Acquisition. Jay Liebowitz (Ed.) *The Handbook of Applied Expert Systems*. London: CRC Press.

Low, I. (2004). Proposed performance standards for College / Unit boundary in Area Control. (NATS internal document: ATIP Project, WP 2.1. Version 1.1). Hurn: NATS.

Long, M.W. (2010). Development Of The Nontechnical Skills For Officers Of The Deck (NTSODod) Rating Form. (MSc Thesis). Monterey, California: Naval Postgraduate School. Retrieved, November 2012, from

www.dtic.mil/dtic/tr/fulltext/u2/a536533.pdf

McCarney, R., Warner, J, Iliffe, S., van Haselen, R., Griffin, M., Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. BMC Medical

Research Methodology Vol. 7. Issue 30. 1-8.

Malakis, S., Kontogiannis, T., Kirwan, B. (2010). Managing emergencies and abnormal situations in air traffic control (part II): Teamwork strategies. Applied Ergonomics Vol. 41. 628–635.

Manning, C.A., Mills, S.H., Fox, C., Pfleiderer, E. (2001). Investigating the Validity of Performance and Objective Workload Evaluation Research (POWER). (FAA Report DOT/FAA/AM-01/10). Civil Aerospace Medical Institute. Oklahoma City: Federal Aviation Administration. Retrieved, March 2008, from

http://amelia.db.erau.edu/reports/faa/am/AM01-10.pdf

Meister, D. (1985). *Behavioral Analysis and Measurement Methods*. Chichester: John Wiley & Sons.

Meister, D. (1986). *Human Factors Testing and Evaluation. Volume 5*. New York: Elsevier Science Publishing Company Ltd.

Mitchell, L., Flin, R., Yule, S., Mitchell, J., Coutts, K., Youngson, G.G. (2011). Thinking ahead of the surgeon: an interview study to identify scrub nurses' non-technical skills. International Journal of Nursing Studies. Vol. 48, 818–828.

Mitchell. L., Flin. R., Yule. S., Mitchell J., Coutts. K., Youngson. G. (2012a) Evaluation of the Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS) system. International Journal of Nursing Studies, Vol. 49, Issue 2. 201-211.

Mitchell, L., Flin, R., Yule. S., Mitchell, J., Coutts, K., Youngson, G. (2012b) Development of a behavioural marker system for scrub practitioners' non-technical skills (SPLINTS system). Journal of Evaluation in Clinical Practice. Vol. 19. Issue 2. 317-323.

van de Mortel. T., Bourke. R., Fillipi. L., McLoughlin. J., Molihan. C., Nonu. M., Reis. M. (2000). Maximising handwashing rates in the critical care unit through yearly performance feedback. Australian Critical Care, Vol. 13, Issue 3. 91-95

Musson, D. (2000). A proposal for the integration of Behavioural Research into International Space Station Operations. In Mohamed S. El-Genk (Ed.). *Proceedings of the Space Technology and Applications International Forum.* (pp. 1484153) Albuquerque NM.

NATS (2002). Manual of Air Traffic Services Part 2. London Terminal Control Centre. (Edition 4/02). Retrieved, February 2013, from

http://www4methods.ch.cam.ac.uk/meth/nmr/newsite/MATS2.pdf

NATS (2007). Day-to-Day 4 Human Performance observation form. (NATS internal document). Fareham: NATS.

Nock, M.K., Kurtz, S.M.S. (2005). Direct Behavioral Observation in School Settings: Bringing Science to Practice. Cognitive and Behavioral Practice Vol. 12, 3594370.

Neilson, J., (2006). *Quantitative Studies: How Many Users to Test?* Retrieved, January 2013, from

http://www.nngroup.com/articles/quantitative4studies4how4many4users/

Nunes, A., Laursen, T. (2004). Identifying the factors that contributed to the Überlingen mid air collisions. (September 20424) *Proceedings of the 48th Annual Chapter Meeting of the Human Factors and Ergonomic Society*. New Orleans, LA.

O'Connor, P., Hörmann, H.J., Flin, R., Lodge, M. Goeters, K.M., The JARTEL Group (2002). Developing a Method for Evaluating Crew Resource Management Skills: A European Perspective. International Journal of Aviation Psychology. Vol. 12. Issue 3, 2634285.

O'Connor, P., Long M.W. (2011). The development of a prototype behavioral marker system for US Navy officers of the deck. Safety Science. Vol. 49, Issue 10. 138141387.

Oprins, E., Burggraaff, E., van Weerdenburg, H. (2006). Design of a competence-based assessment system for ATC training. The International Journal of Aviation Psychology. Vol. 16, Issue 3, 2974320.

Oprins, E. (2008). Design of A Competence-Based Assessment System for Air

Traffic Control Training. (PhD Thesis). Universitaire Pers Maastricht

Patrick, J. (1992). Training: Research & Practice. London: Academic Press Limited.

Rathje, H., Golany, Z., Eiβfeldt, H. (2004). Pan-European Test Battery for Air Traffic Control Applicants. In Klaus-Martin Goeters (Ed.) *Aviation Psychology: Practice and Research*. (pp171-202). London: Ashgate Publishing Ltd.

Reason, J. (1990). *Human Error*. Cambridge: Cambridge University Press.

Reason, J. (2008). *The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries.* Farnham: Ashgate Publishing Ltd.

Riener, A. (2010). Assessment of Simulator Fidelity and Validity in Simulator and On-the-road Studies. International Journal on Advances in Systems and Measurements. Vol. 3, No. 3-4. 110-124

Roske-Hofstrand, RJ, Murphy, E.D. (1998). Human Information Processing in Air Traffic Control. In M. Smolensky and E. Stein (Eds.) *Human Factors in Air Traffic Control.* (pp65-114). New York: Academic Press.

Salas, E., Burke, S., Bowers, C.A., Wilson, K.A. (2001). Team Training in the Skies: Does Crew Resource Management (CRM) Training Work? *Human Factors*. Vol. 43, Issue 4. 641-674.

Salas, E., Cooke, N.J., Rosen, M.A., (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. Human Factors. Vol. 50, Issue. 3. 540–547

Sanders, M.S., McCormick. E.J. (1992). Human Factors in Engineering and Design. Seventh Edition. McGraw-Hill International Editions.

Santos, J., Merat, N., Mouta, S., Brookhuis, K., de Waard, D. (2005). The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. Transport Research Part F: Traffic Psychology and Behaviour. Vol. 8, Issue 2. 135-146.

Shorrock, S. (2007). Heathrow NVCR Practice Simulation and Shadowing Behavioural Observation Study. (NATS internal Document). Fareham:

NATS.

Silverman, D. (2005). *Doing Qualitative Research. A Practical Handbook. Second Edition*. London: Sage Publications Ltd.

Smith, A.D.N. & Stamp, R.A. (1973). A method for estimating the capacity of air traffic sectors – An interim report. (CAA DORA Research Paper 7301). London: CAA.

Smith, A.D.N. and Stamp, R.A. (1974). A method for estimating the capacity of air traffic sectors – Final report. (CAA DORA Research Paper 7402). London: CAA.

Sollenberger, R., Stein, E., Gromelski, S. (1997). The Development and Evaluation of a Behaviorally Based Rating Form for Assessing Air Traffic Controller Performance. (FAA Technical Report No: DOT/FAA/CT-TN96/16). Retrieved, July 2008, from

http://hf.tc.faa.gov/technotes/dot4faa4ct4tn96416.pdf

Stanton, N. (2005). *Human Factors Methods: A Practical Guide for Engineering and Design*. London: Ashgate Publishing, Ltd.

Swift, R.H., Harrigan, E.P., Cappelleri, J.C., Kramer, D., Chandler, L.P. (2002). Journal of Psychiatric Research Vol. 36, Issue 2. 87–95.

Thatcher, S. (2007). Experiential learning and the classification of team skills in pilot education. World Transactions on Engineering and Technology Education. Vol. 6, Issue 2. 2714274.

Thomas. L., & Robinson. K. (2008) ATCO Competency framework: En route, Aerodrome & Approach Controllers. (NATS internal document). Fareham: NATS.

Unsworth, K.L., & West, M.A. (2000). Teams: The challenges of cooperative work. In: Chmiel, Nik, (ed) *Introduction to Work and Organizational Psychology*. (pp. 3274346) Oxford: Blackwell Publishing.

Vardaman, J.L., Stein. E.S. (1998). The Development and Evaluation of a Behaviourally Based Rating Form for the Assessment of En Route Air Traffic Controller Performance. (FAA Technical Report No: DOT/FAA/CT-TN98/5/). Retrieved, March 2008, from

www.hf.faa.gov/docs/508/docs/wjhtc/tn9805.pdf

Williamson, A. & Feyer, A.M. (1990). Behavioural Epidemiology as a Tool for Accident Research. Journal of Occupational Accidents, Vol. 12. 207-222.

Wickens, C.D., & Hollands. J.G. (2000). *Engineering Psychology and Human Performance. Third Edition*. London: Prentice-Hall International (UK) Ltd.

Wilhelm, J.A. & Helmreich, R.L. (1996). The Line/LOS checklist for checking pilots: A short form for evaluation of crew human factors skills in line flight settings. (FAA Technical Report 96-6). Austin: University of Texas Aerospace Crew Research Project.

Willems, B., Truitt. T.R. (1999). Implications of Reduced Involvement in En Route Air Traffic Control. (FAA Technical Report No: DOT/FAA/CT-TN99/22). Atlantic City: FAA William J. Hughes Technical Center.

Woldring, M., & Patterson. I. (2003). Identifying and Evaluating Non-Technical Skills. (12-13th November). Presentation to CRM – European Air Traffic Management. University of Aberdeen.

Woldring, M., Van Damme, D., Patterson, I., Henriques, P. (2005). Team Resource Management in European Air Traffic Control: Results of a Seven-Year Development and Implementation Program. In B. Kirwan, M. Rodgers, D. Schafer (Eds.) *Human Factors Impacts in Air Traffic Management*. London: Ashgate Publishing Ltd.

Yule, S., Flin, R., Paterson-Brown, S. & Maran, N. (2006a). Non-technical skills for surgeons: A review of the literature. Surgery, Vol. 139, 140-149.

Yule, S., Flin, R., Paterson-Brown, S., Maran, N., Rowley, D. (2006b) Development of a rating system for surgeons' non-technical skills. Medical Education. Vol. 40, Issue 11. 1098–1104.

Yule, S., Flin, R., Maran, N., Rowley, D., Youngson, G., Paterson-Brown, S. (2008). Surgeons' Non-technical Skills in the Operating Room: Reliability Testing of the NOTSS Behavior Rating System. World Journal of Surgery, Vol. 32, Issue 4. 548-556

APPENDIX A – ENGD BACKGROUND INFORMATION

A01 - BACKGROUND OF ENGD RESEACHER

The author graduated from Loughborough University in 2000 with a Bachelor of Science degree in Ergonomics. A Diploma in Professional Studies was also awarded for successful completion of an industrial placement year spent at the Defence Evaluation and Research Agency (DERA). A DERA Student bursary and offer of employment was awarded at the end of placement.

The author returned to DERA following graduation to work within the Centre for Human Sciences Department. He was involved in human performance assessment with new interfaces and technologies for several projects within the defence and security domains. During this time the organisation underwent privatisation and transformed into the public company QinetiQ, whilst the author attained Registered Membership of the Institute of Ergonomics and Human Factors (MIEHF) and advancement to Senior Human Factors Scientist.

In October 2006 the author joined the Human Factors team at NATS as a Senior Human Factors Specialist. Over the next 12 months, experience on various projects highlighted the potential opportunity to undertake doctoral level study within the ATC industrial setting. Following preliminary investigations into various PhD possibilities the author explored a publicised Engineering Doctorate (EngD) position scoped in part by Dr. David Bush (a former UCL Communications EngD student at NATS). The author until this point had been unaware of the philosophy and mechanics of an EngD; however its mixture of industrial placement and grounded research focus coupled with strong academic underpinnings was extremely attractive.

With agreement from the former Head of Human Factors (Liz Skinner) and Deputy Head of Human Factors (Nic Turley), an application for the UCL E&EE Department's Communication EngD was submitted. The application itself included a preliminary proposal of research concerning the development of a behavioural marker set to assess Air Traffic Controller Performance. This research topic stemmed from preliminary investigations in this area by Dr Steve Shorrock (formerly NATS and now EUROCONTROL and the University of New South Wales) and Dr Anne Isaac. This topic strongly appealed to the author, and was found to be highly acceptable as a research topic for NATS. An EPSRC Case studentship was awarded to the author, with Dr Nadia Berthouze and Prof. Ann Blandford (UCL Interaction Centre) as primary and secondary supervisors. Dr Brian Janes, the author's immediate manager at NATS acted as industrial supervisor throughout the research. From 2010 onward, supervisor roles were reversed with Professor Blandford as primary, and Doctor Berthouze as secondary supervisor, due to the slight change in focus from the initial proposal.

A02 - ENGD TAUGHT COMPONENT

Personal & Professional Management Skills (PPMS)

This was a three day residential skills reflection and development course run by the Centre for Applied Learning and Teaching (CALT). A pre-course seminar was undertaken to identify the skills we wished to develop during the residential stage. The author focused on verbal communications development, and took opportunities available to practice skills within this area of personal development. This module was non-assessed; as the assessment is undertaken by the individual participant and is self-reflective.

Usability Evaluation Methods (PSYCGI11)

This eight week modular course involved the study of how to successfully design and undertake a usability assessment. The course included practical computer lab based application of methodologies and theory presented within the lectures. This module was assessed through course work only, and involved designing and undertaking a usability assessment of a commercial website. The author received a distinction for this module.

Project Management (MECHGS11)

A five day block module structured around the Association of Project Managers *Body of Knowledge*. Satisfactory completion of this model was achieved through the successful undertaking of the APMP examination. This examination tests the student on 37 knowledge areas from the *Body of Knowledge* and leads to an internationally recognised qualification in project management.

Applied Cognitive Science (PSYCGI08)

This eight week modular course involved the study of cognitive modelling and analysis. This included classical models of cognition, and cognitive modelling, and modelling human performance and error. This module was assessed through examination, involving questions on experimental design, and cognitive science theory.

Advanced Experimental Design and Analysis (D481 09K)

This is an Open University distance learning course. The course covers a number of statistical methods used in psychological data analysis. The syllabus included analysis of variance, planned and unplanned comparisons, multiple random factors, power analysis, regression (multiple, log-linear and logistic), analysis of covariance, and meta-analysis. This double module was assessed through three pieces of coursework, and one unseen written examination.

Strategy, Marketing and the Business Environment (ELECGB11)

A five day intensive module from the MSc Telecoms with Business, this concerned the economic environment in which ICT and telecommunications companies operate. Business operations and planning were explored, in addition to macro-economic models, regulation and legislation, marketing, and competitor analysis. This module was assessed through a coursework assignment, which involved the production of a future business strategy for a telecommunications company. The author received a distinction for this module.

Finance and Product Management (ELECGB12)

The second five day block module from the MSc Telecoms with Business focused upon the foundation principles of business finance and management accounting. Areas covered included financial accounting, portfolio management and business cases, and the economic infrastructure of networks. This module was assessed through one unseen written examination.

Customer Service, Operations and Planning (ELECGB13)

A further five day MSc Telecoms with Business module, concerned the fundamental principles of organisational design and management, to deliver day-to-day service. This includes the management of the customer experience and customer service principles, ICT project management and network planning, risk management, and organisational culture and climate. This module was assessed through a coursework assignment, which involved the critical review of customer and service management within a telephone network provider. The author received a distinction for this module.

Global Aspects, innovation Management, People Management and Organisational Design (ELECGB14)

The final business module from the MSc Telecoms with Business focused upon the principles of effective organisational design and dynamics, in particular; the transformation of a business through resource and team management, innovation and cultural management, and skills training. This module was assessed through a coursework assignment, which involved the critical review of a business transformation plan, and recommendations for strategic management, organisational structure, and enhancements to cultural dynamics. The author received a distinction for this module.

A03 - ENGD ROBERTS POINTS

A requirement for the EngD student uptake of 200742008 has been to undertake industry transferable skills development and training in the order of 20 points per annum (1 point equates to a 1/2 day). Table A1 presents the various activities, courses, and events undertaken by the author over the course of study, and the value of each activity in Roberts Points.

Roberts Points	Title/Descriptor	Date
4	Experiment Planning for the Life Sciences 4 Year 1	2008
1	Departmental Seminar: Engaging participants in ubicomp design: Lessons from Chawton House Geraldine Fitzpatrick (University of Sussex)	27/02/08
1	Departmental Seminar: Changing the face of the PC Quentin Stafford-Fraser (Cambridge Visual Networks)	21/05/08
1	Departmental Seminar Prioritisation, resources and search terms: A study of decision-making at the virtual reference desk. Simon Attfield (UCLIC)	05/09/08
10	Personal & Professional Management Skills (PPMS)	29/08/2008 (pre course workshop) 11413/08/08 (Residential Course)
1	EngD Conference Event 4 Prepare Presentation	09/2008
2	EngD Conference Event 4 Present Presentation	18/09/2008
4	Experiment Planning for the Life Sciences 4 Year 2	2009
4	LBS 4 Finance & Value Creation Workshop	9410/01/2009
2	Communications EngD Poster Competition 4 Prep	03/2009
1	Communications EngD Poster Competition 4 Present	03/04/2009
10	System Engineering Project Management Module	16420/02/2009
6 (3 x 2)	Presentations to Sponsor (NATS) 4 Presentation Prep	06/2009, 07/2009, 09/2009
3 (3 x 1)	Presentations to Sponsor (NATS) 4 Give Presentation	07/06/2009, 15/07/2009, 12/10/2009
4	Experiment Planning for the Life Sciences 4 Year 3	2010
2	IEHF Doctoral Consortium Paper Preparation Nottingham University	04405/2010
2	IEHF Doctoral Consortium Presentation Preparation	
1	Presentation to Industrial Sponsor (IEHF) paper	04/05/2010
1	IEHF Doctoral Consortium Presentation of paper Nottingham University	19/05/2010
2	Communications EngD Poster Competition 4 Prep	25/06/2010
1	Communications EngD Poster Competition 4 Present	
1	NATS Training 4 Dealing With Difficult People Jason Demagalski (NATS Human Factors)	07/07/2010 4 2 hour introduction 16/07/2010 4 3 hour practical session
2	LCS Paper Preparation	10/09/2010
2	LCS Presentation Preparation	
1	LCS Presentation of Paper	
4	Experiment Planning for the Life Sciences 4 Year 4	2011
1	NATS Training 4 FEAST (3 hour training session) Marc Damitz from Eurocontrol	24/02/2011
2	Communications EngD Poster Competition 4 Prep	25/03/2011
1	Communications EngD Poster Competition 4 Present	
1	Digital Story Workshop 4 Prof Harold Thimbleby	18/03/2011
10	Strategy, Marketing and the Business Environment	10//201 1 4 03/2012 http://www.ee.ucl.ac.uk/cpd/modules
10	Finance and Product Management	
10	Customer Service, Operations and Planning	
10	Global Aspects, innovation Management, People Management and Organisational Design	
4	NATS Course 4 Communications & Influencing Skill	1441 6/11/2011
2	LCS Paper Preparation	09/2011
2	LCS Presentation Preparation	
1	LCS Presentation of Paper	

Table A8 - UCL Roberts Points (Years 1 - 4)

APPENDIX B – BEHAVIOURAL OBSERVATION SYSTEMS

B01 - FAA ADVISORY CIRCULAR AC120-51 E

1. COMMUNICATIONS PROCESSES AND DECISION BEHAVIOR CLUSTER.	
Briefings. An effective briefing is interesting and thorough. It addresses coordination, planning, and problems. Although briefings are primarily a captain's responsibility, other crewmembers may add significantly to planning and should be encouraged to do so.	(1) The captain's briefing establishes an environment for open/interactive communications (e.g., the captain calls for questions or comments, answers questions directly, listens with patience, does not interrupt or "talk over," does not rush through the briefing, and makes eye contact as appropriate).
	(2) The briefing is interactive and emphasizes the importance of questions, critique, and the offering of information.
	(3) The briefing establishes a "team concept" (e.g., the captain uses "we" language, encourages all to participate and to help with the flight).
	(4) The captain's briefing covers pertinent safety and security issues.
	(5) The briefing identifies potential problems such as weather, delays, and abnormal system operations.
	6) The briefing provides guidelines for crew actions centered on standard operating procedures (SOP); division of labor and crew workload is addressed.
	(7) The briefing includes the cabin crew as part of the team.
	(8) The briefing sets expectations for handling deviations from SOPs.
	(9) The briefing establishes guidelines for the operation of automated systems (e.g., when systems will be disabled; which programming actions must be verbalized and acknowledged).
	(10) The briefing specifies duties and responsibilities with regard to automated systems, for the pilot flying (PF) and the pilot monitoring (PM).
Inquiry/Advocacy/Assertion. These behaviors relate to crewmembers promoting the course of action that they feel is best, even when it involves conflict with others.	(1) Crewmembers speak up and state their information with appropriate persistence until there is some clear resolution.
	(2) "Challenge and response" environment is developed.
	(3) Questions are encouraged and are answered openly and non defensively.
	(4) Crewmembers are encouraged to question the actions and decisions of others.
	(5) Crewmembers seek help from others when necessary.
	(6) Crewmembers question status and programming of automated systems to confirm situation awareness.
Crew Self-Critique Regarding Decisions and Actions. These behaviors relate to the effectiveness of a group and/or an individual crewmember in critique and debriefing. Areas covered should include the product, the process, and the people involved. Critique may occur during an activity, and/or after completing it.	(1) Critique occurs at appropriate times, which may be times of low or high workload.
	(2) Critique deals with positive as well as negative aspects of crew performance.
	(3) Critique involves the whole crew interactively.
	(4) Critique makes a positive learning experience. Feedback is specific, objective, usable, and constructively given.
	(5) Critique is accepted objectively and non defensively.
Communications/Decisions. These behaviors relate to free and open communication. They reflect the extent to which crewmembers provide necessary information at the appropriate time (e.g., initiating checklists and alerting others to developing problems). Active participation in the decision-making process is encouraged. Decisions are clearly communicated and acknowledged. Questioning of actions and decisions is considered routine.	(1) Operational decisions are clearly stated to other crewmembers.
	(2) Crewmembers acknowledge their understanding of decisions.
	(3) "Bottom lines" for safety are established and communicated.
	(4) The "big picture" and the game plan are shared within the team, including flight attendants and others as appropriate.
	(5) Crewmembers are encouraged to state their own ideas, opinions, and recommendations.
	(6) Efforts are made to provide an atmosphere that invites open and free communications.
	(7) Initial entries and changed entries to automated systems are verbalized and acknowledged.

Table B1 - FAA CRM Behavioural Markers, Advisory Circular 120-51E - Part 1 (FAA, 2004)

2. TEAM BUILDING AND MAINTENANCE CLUSTER.	
Leadership Followership/Concern for Tasks. These behaviors relate to appropriate leadership and followership. They reflect the extent to which the crew is concerned with the effective accomplishment of tasks.	(1) All available resources are used to accomplish the job at hand.
	(2) Flight deck activities are coordinated to establish an acceptable balance between respect for authority and the appropriate practice of assertiveness.
	(3) Actions are decisive when the situation requires.
	(4) A desire to achieve the most effective operation possible is clearly demonstrated.
	(5) The need to adhere to SOPs is recognized.
	(6) Group climate appropriate to the operational situation is continually monitored and adjusted (e.g., social conversation may occur during low workload, but not high).
	(7) Effects of stress and fatigue on performance are recognized.
	(8) Time available for the task is well managed.
	(9) Demands on resources posed by operation of automated systems are recognized and managed.
	(10) When programming demands could reduce situation awareness or create work overloads, levels of automation are reduced
Interpersonal Relationships/Group Climate. These behaviors relate to the quality of interpersonal relationships and the pervasive climate of the flight deck.	(1) Crewmembers remain calm under stressful conditions.
	(2) Crewmembers show sensitivity and ability to adapt to the personalities of others.
	(3) Crewmembers recognize symptoms of psychological stress and fatigue in self and in others (e.g., recognizes when he/she is experiencing "tunnel vision" and seeks help from the team; or notes when a crewmember is not communicating and draws him/her back into the team).
	(4) "Tone" in the cockpit is friendly, relaxed, and supportive.
	(5) During times of low communication, crewmembers check in with others to see how they are doing.
3. WORKLOAD MANAGEMENT AND SITUATION AWARENESS CLUSTER.	
Preparation/Planning/Vigilance. These behaviors relate to crews anticipating contingencies and the various actions that may be required. Excellent crews are always "ahead of the curve" and generally seem relaxed. They devote appropriate attention to required tasks and respond without undue delay to new developments. (They may engage in casual social conversation during periods of low workload and not necessarily diminish their vigilance.)	(1) Demonstrating and expressing situation awareness (e.g., the "model" of what is happening is shared within the crew).
	(2) Active monitoring of all instruments and communications and sharing relevant information with the rest of the crew.
	(3) Monitoring weather and traffic and sharing relevant information with the rest of the crew.
	(4) Avoiding "tunnel vision" caused by stress (e.g., stating or asking for the "big picture").
	(5) Being aware of factors such as stress that can degrade vigilance, and watching for performance degradation in other crewmembers.
	(6) Staying "ahead of the curve" in preparing for planned situations or contingencies, so that situation awareness and adherence to SOPs is assured.
	(7) Ensuring that cockpit and cabin crewmembers are aware of plans.
	(8) Including all appropriate crewmembers in the planning process.
	(9) Allowing enough time before manoeuvres for programming of the flight management computer.
	(10) Ensuring that all crewmembers are aware of initial entries and changed entries in the flight management system.
Workload Distributed/Distractions Avoided. These behaviors relate to time and workload management. They reflect how well the crew manages to prioritize tasks, share the workload, and avoid being distracted from essential activities.	(1) Crewmembers speak up when they recognize work overloads in themselves or in others.
	(2) Tasks are distributed in ways that maximize efficiency.
	(3) Workload distribution is clearly communicated and acknowledged.
	(4) Non operational factors such as social interaction are not allowed to interfere with duties.
	(5) Task priorities are clearly communicated.
	(6) Secondary operational tasks (e.g., dealing with passenger needs and communications with the company) are prioritized so as to allow sufficient resources for primary flight duties.
	(7) Potential distractions posed by automated systems are anticipated, and appropriate preventive action is taken, including reducing or disengaging automated features as appropriate.

Table B2 - FAA CRM Behavioural Markers, Advisory Circular 120&51E - Part 2 (FAA, 2004)

B02 - NOTECHS

Element	Good practice	Poor practice
Non-Technical Skills Category COOPERATION		
Team building and maintaining	Establishes atmosphere for open communication	Blocks open communication
	Encourages inputs and feedback from others	Keeps barriers between CM
	Does not compete with others	Competes with others
Consideration of others	Takes notice of the suggestions of other CM even if s/he does not agree	Ignores suggestions of other CM
	Takes condition of other CM into account	Does not take account of the condition of other CM
	Gives personal feedback	Shows no reaction to other CM
Support of others	Helps other CM in demanding situations	Hesitates to help other CM in demanding situations
	Offers assistance	Does not offer assistance
Conflict solving	Keeps calm in interpersonal conflicts	Overreacts in interpersonal conflicts
	Suggests conflict solutions	Sticks to own position without considering a compromise
	Concentrates on what is right rather than who is wrong	Accuses other CM of making errors
The Non-Technical Skills Category LEADERSHIP AND MANAGERIAL SKILLS		
Use of authority and assertiveness	takes initiative to ensure crew involvement and task completion	hinders or withholds crew involvement
	takes command if situation requires, advocates own position	passive, does not show initiative for decisions, own position not recognizable
	reflects on suggestions of others	ignores suggestions of others
	motivates crew by appreciation and coaches when necessary	does not show appreciation for the crew, coaches very little or too much
Providing and maintaining standards	subscribes to SOPs, makes sure SOP compliance in crew	does not comply to SOPs, does not monitor crew for SOP compliance
	intervenes if task completion deviates from standards	does not intervene in case of deviations
	with crew being consulted, deviates from standards if necessary	deviation from standards are neither announced nor consulted
	Demonstrates will to achieve top performance	does not care for performance effectively
Planning and coordination	encourages crew participation in planning and task completion	plans only for him/herself, crew not involved
	plan is clearly stated and confirmed	intentions not stated or confirmed
	with crew being consulted, changes plan if necessary	changes plan without informing crew or follows plan blindly
	clearly states goals and boundaries for task completion	goals and boundaries remain unclear
Workload management	distributes tasks among the crew, checks and corrects appropriately	flying "solo" without other crewmembers involved
	secondary operational tasks are prioritized to retain sufficient resources for primary flight duties	secondary operational tasks interfere with primary flight duties
	allots adequate time to complete tasks	workload is increased through inadequate planning
	notifies signs of stress and fatigue	ignores signs of stress and fatigue

Table B3 - NOTECHS (Adapted from Flin et al, 1998)

Element	Good practice	Poor practice
The Non-Technical Skills Category SITUATION AWARENESS:		
System awareness	Monitors and reports changes in systems' states	does not ask for updates
	Acknowledges entries and changes to systems	does not signal awareness of changing systems
Environmental awareness	Collects information about environment (position, weather and traffic)	does not enquire about environmental changes
	Shares key information about environment with crew	does not comment on relevant environmental factors, or is surprised by them
	Contacts outside resources when needed (to maintain situation awareness)	Operates a 'closed shop'
Awareness of time and anticipation of future events	Discusses time constraints with crew	does not set priorities wrt time limits
	Discusses contingency strategies	does not dis cuss relationship between past events and present/future
	Identifies possible future problems	is surprised by outcomes of past events
The Non-Technical Skills Category DECISION MAKING:		
Problem definition/ diagnosis	gathers information to identify problem	Nature of problem not stated or failure to diagnose
	reviews causal factors with other crew members	no discussion of probable causes
Option generation	states alternative options	does not search for information
	asks crew members for options	does not ask crew for alternatives
Risk assessment	considers and shares estimated risk of alternative opt	Inadequate discussion of limiting factors with crew
	talks about possible risks for action in terms of crew limits	no consideration of limiting factors
Option selection	Confirms and states selected option / agreed action	does not inform crew of decision path being taken
Outcome review	Checks outcome against plan	Fails to check selected outcome against goal

Table B4 - NOTECHS continued (Flin et al, 1998)

B03 - NASA ISS BEHAVIOURAL MARKERS

Category	Competency	Behavioural Markers
Self-Care, Self-Management	Refine accuracy of self image	Identifies personal tendencies and their influence on own behaviour.
		Identifies factors for personal successes or failures
		Seeks formal and informal feedback to understand impact of own behaviour on others
		Assesses own skills knowledge and abilities against task requirements
	Manage stress	Identifies symptoms and causes of personal stress
		Takes action to prevent and mitigate stress, negative mood, or low morale
		Uses calm and flexible approach in dealing with unfamiliar situations
	Care for oneself	Uses mistakes as learning opportunities
		Maintains social relationships
		Maintains personal goals for satisfaction and motivation and to maximize performance
		Maintains balance of work, personal time and rest
	Maintain efficiency	Sets challenging and attainable goals
		Uses time efficiently
		Keeps items organized
Communication	Optimize communication	Communicates information clearly and concisely
		Shares information
		Communicates intentions before taking action
		Communicates task status and completion
		Provides constructive feedback
		Adjusts time and/or style of communication to fit the situation
		Communicates concerns; persists until acknowledged
		Establishes atmosphere for open and constructive communication
		Briefs and debriefs behavioral and technical issues with team members
		Shares information
		Communicates intentions before taking action
		Communicates task status and completion
		Provides constructive feedback
		Adjusts time and/or style of communication to fit the situation
		Communicates concerns; persists until acknowledged
		Establishes atmosphere for open and constructive communication
		Briefs and debriefs behavioral and technical issues with team members
	Ensure Understanding	Listens "actively" addresses barriers to communication
		Seeks answers in proactive manner
		Verifies information
		Acknowledges confusion or misunderstanding
		Resolves discrepancies, confusions, and misunderstandings
Cross Cultural	Demonstrate respect towards other cultures (national, organisational, professional)	Demonstrates respect and appreciation for team members' culture[s] and viewpoints
		Respects differences in gender role expectations, behaviours, and attitudes
	Understand culture and cultural differences (national, organizational and professional)	Uses understanding of cultural factors and circumstances to interpret team members' behaviours
		Acknowledges the impact of cultural dominance on crew interaction
		Mitigates the impact of cultural stereotypes and prejudices on group interaction
	Build and maintain social and working relationships	Demonstrates tolerance of cultural differences and ambiguities
		Develops strategies to clarify ambiguities created by own behavior
	Intercultural communication and language skills	Communicates respectfully with people from different cultural and linguistic backgrounds
		Makes an effort to learn and use the languages of colleagues
		Puts a common "space-faring culture" ahead of one's own national organizational and professional cultures

Table B5 - NASA ISS Behavioural Markers (Bessone et al, 2008a)

Category	Competency	Behavioural Markers
Teamwork & Group Living	Active team participation	Acts cooperatively rather than competitively
		Takes responsibility for own actions and mistakes
		Puts common goals above individual needs
		Works with teammates to ensure safety and efficiency
		Respects team member's roles, responsibilities, and task allocation
	Interpersonal relationships	Demonstrates patience, respect and appreciation for crewmembers
		Provides emotional support to crewmembers
		Encourages participation in team activities
		Develops positive relationships with team members
	Group living	Adapts living and working habits to improve team cohesion
		Volunteers for routine and unpleasant tasks
		Offers and provides assistance if accepted
		Balances own needs with those of crewmembers
		Shares attention and credit for achievements with teammates
Leadership	Execution of designated leader's authority	Accepts leadership responsibilities
		Assigns tasks according to capabilities and individual preferences
		Assigns tasks with clearly defined goals
		Adapts leadership styles to situation
		Responds to information, suggestions, and concerns of team members
		Maintains team cohesion in adverse and uncertain circumstances
	Mentoring skills	Provides direction, information, feedback, and encouragement and coaching as needed
		Leads by example
	Followership	Supports leader
		Reacts promptly to situations requiring immediate response
	Workload Management	Plans and prioritizes tasks
		Adapts plans according to progress and changing conditions
		Ensures team members have the appropriate tools and authorization to complete tasks
Conflict Management	Conflict prevention	Addresses potential sources for conflict
		Prevents disagreements from influencing personal and professional relationships
	Conflict resolution	Reviews causal factors of a conflict with all involved team members
		Adapts conflict management strategies to resolve disagreements
		Exchanges views and positions
		Seeks resolution
		Keeps calm in interpersonal conflicts
		Focuses on what is wrong rather than who is wrong
		Mediates between conflicting parties
		Defines agreement and positive closure
Situational Awareness	Maintenance of an accurate perception of the situation	Monitors people, systems, and environment
		Monitors self and others for signs of stress, fatigue, complacency, and task saturation
		Reduces distractions while performing operational tasks
		Maintains awareness of the environment while focusing on a task or problem
		Maintains the required level of vigilance for low and high workloads
		Uses the two-person approach to execution of critical tasks and procedures
	Processing of information	Analyzes information to determine operational relevance
		Assesses impacts of actions, plans, and decisions on others
		Anticipates potential problems
		Verifies team readiness to meet operational demands
		Communicates when situations "feel" wrong
		Identifies and resolves discrepancies between conflicting data or information
Decision Making and Problem Solving	Problem solving and decision making methods	Adopts a problem solving method to meet situational demands
	Preparation of decision	Involves team members in the process as applicable
		Assembles **F**acts
		Considers different **O**ptions
		Evaluates **R**isks and benefits
		Decides on an option
	Execution of decision	**E**xecutes decision
		Checks results of decision, and if necessary reapplies process

Table B6 - NASA ISS Behavioural Markers continued (Bessone et al, 2008a)

B04 - ANTS

Non-Technical Skill - Task Management		
Planning and preparing Developing in advance primary and contingency strategies for managing tasks, reviewing these and updating them if required to ensure goals will be met; making necessary arrangements to ensure plans can be achieved.	**BM for good practice** • Communicates plan for case to relevant staff • Reviews case plan in light of changes • Makes post-operative arrangements for patient • Lays out drugs and equipment needed before starting case	**BM for poor practice** • Does not adapt plan in light of new information • Does not ask for drugs or equipment until the last minute • Does not have emergency/alternative drugs available suitable for patient • Fails to prepare post-op management plan
Prioritising Scheduling tasks, activities, issues, information channels, etc., according to importance (e.g. due to time, seriousness, plans); being able to identify key issues and allocate attention to them accordingly, and avoiding being distracted by less important or irrelevant matters.	**BM for good practice** • Discusses priority issues in case • Negotiates sequence of cases on list with surgeon • Conveys order of actions in critical situations	**BM for poor practice** • Becomes distracted by teaching trainees • Fails to allocate attention to critical areas • Fails to adapt list to changing clinical conditions
Providing and maintaining standards Supporting safety and quality by adhering to accepted principles of anaesthesia; following, where possible, codes of good practice, treatment protocols or guidelines, and mental checklists.	**BM for good practice** • Follows published protocols and guidelines • Cross-checks drug labels • Checks machine at beginning of each session • Maintains accurate anaesthetic records	**BM for poor practice** • Does not check blood with patient and notes • Breaches guidelines such as minimum monitoring standards • Fails to confirm patient identity and consent details • Does not adhere to emergency protocols or guidelines
Identifying and utilising resources Establishing the necessary, and available, requirements for task completion (e.g. people, expertise, equipment, time) and using them to accomplish goals with minimum disruption, stress, work overload or underload (mental and physical) on individuals and the whole team.	**BM for good practice** • Identifies resources that are available • Allocates tasks to appropriate member(s) of the team • Ensures time is free for busy/critical periods • Requests additional resources if needed	**BM for poor practice** • Fails to utilise available resources • Overloads team members with tasks • Does not recognise when task load is unworkable • Does not request necessary resources in advance

Table B7 - ANTS Checklist (Industrial Psychology Research Centre, Aberdeen University, 2003)

Non-Technical Skill - Teamworking		
Co-ordinating activities with team members Working together with others to carry out tasks, for both physical and cognitive activities; understanding the roles and responsibilities of different team members, and ensuring that a collaborative approach is employed.	**BM for good practice** • Confirms roles and responsibilities of team members • Discusses case with surgeons or colleagues • Considers requirements of others before acting • Co-operates with others to achieve goals	**BM for poor practice** • Does not co-ordinate with surgeon(s) and other groups • Relies too much on familiarity of team for getting things done - makes assumptions, takes things for granted • Intervenes without informing/ involving others • Does not involve team in tasks
Exchanging information Giving and receiving the knowledge and data necessary for team co-ordination and task completion.	**BM for good practice** • Gives situation updates/reports key events • Confirms shared understanding • Communicates case plans and other relevant information to appropriate people • Maintains clear case documentation	**BM for poor practice** • Does not inform team of plan or subsequent alterations • Gives inadequate handover briefing • Does not include relevant people in communications • Fails to express concerns in a clear and precise manner
Using authority and assertiveness Leading the team and/or the task (as required), accepting a non-leading role when appropriate; adopting a suitably forceful manner to make a point, and adapting this for the team and/or situation.	**BM for good practice** • Makes requirements known with necessary level of assertiveness • Takes over task leadership as required • Gives clear orders to team members • States case and provides justification	**BM for poor practice** • Does not challenge senior colleagues or consultants • Does not allow others to put forward their case • Fails to attempt to resolve conflicts • Does not advocate position when required
Assessing capabilities Judging different team members' skills, and their ability to deal with a situation; being alert to factors that may limit these and their capacity to perform effectively (e.g. level of expertise, experience, stress, fatigue).	**BM for good practice** • Calls for assistance when it is needed • Asks new team member about their experience • Notices that a team member does not perform a task to the expected standard • Adapts level of monitoring to expertise of other team members • Observes that a member of the team has returned from sick leave and enquires about their general health	**BM for poor practice** • Does not ask if trainee/assistant can cope with task • Allows team to accept case beyond its level of expertise • Does not pay attention to the performance of other members of the team, e.g. scrub nurse • Joins established team without ascertaining their capabilities • Fails to respond to obvious cues of fatigue - person yawning, not remembering simple instructions, etc.
Supporting others Providing physical, cognitive or emotional help to other members of the team.	**BM for good practice** • Acknowledges concerns of others • Provides reassurance/encouragement • Debriefs and thanks staff after a difficult case • Anticipates when colleagues will need equipment/information	**BM for poor practice** • Asks for information at difficult/high workload time for someone else • Does not offer assistance to team members • Fails to recognise needs of others requiring task reallocation • Uses a dismissive tone in response to requests from others
Gathering information Actively and specifically collecting data about the situation by continuously observing the whole environment and monitoring all available data sources and cues and verifying data to confirm their reliability.	**BM for good practice** • Obtains and documents patient information pre-operatively • Conducts frequent scan of the environment • Collects information from team to identify problem • Watches surgical procedure, verifying status when required • Cross-checks information to increase reliability	**BM for poor practice** • Reduces level of monitoring because of distractions • Responds to individual cues without confirmation • Does not alter physical layout of workspace to improve data visibility • Does not ask questions to orient self to situation during hand-over

Table B8 - ANTS Checklist continued (Industrial Psychology Research Centre, Aberdeen University, 2003)

Non-Technical Skill - Situation Awareness		
Recognising and understanding Interpreting information collected from the environment (with respect to existing knowledge) to identify the match or mis-match between the situation and the expected state, and to update one's current mental picture.	**BM for good practice** • Increases frequency of monitoring in response to patient condition • Informs others of seriousness of situation • Describes pattern of cues and their meaning to other team members	**BM for poor practice** • Does not respond to changes in patient state • Carries out inappropriate course of action • Silences alarms without investigation
Anticipating Asking 'what if' questions and thinking ahead about potential outcomes and consequences of actions, intervention, non-intervention, etc.; running projections of current situation to predict what might happen in the near future.	**BM for good practice** • Keeps ahead of the situation by giving fluids/drugs • Reviews the effects of an intervention • Sets and communicates intervention thresholds • Takes action to avoid or mitigate potential problems	**BM for poor practice** • Does not consider potential problems associated with case • Fails to increase level of monitoring in keeping with patient condition • Is caught unaware by surgical actions • Does not foresee undesirable drug interactions
Non-Technical Skill - Decision Making		
Identifying options Generating alternative possibilities or courses of action to be considered in making a decision or solving a problem.	**BM for good practice** • Generates options for decisions • Discusses various anaesthetic techniques with patient • Asks other anaesthetists for suggestions on a difficult case	**BM for poor practice** • Even though time is available jumps straight to one option without considering alternatives • Fails to ask other team members for options, when appropriate • Ignores suggestions from other team members
Balancing risks and selecting options Assessing hazards to weigh up the threats or benefits of a situation, considering the advantages and disadvantages of different courses of action; choosing a solution or course of action based on these processes.	**BM for good practice** • Considers risks of different treatment options • Weighs up factors with respect to patient's condition • Assesses time criticality associated with possible options • Implements chosen option	**BM for poor practice** • Does not find out about the risks associated with an unfamiliar condition/drug • Does not preview courses of action with relevant people to assess their suitability • Fails to review possible options with the team
Re-evaluating Continually reviewing the suitability of the options identified, assessed and selected; and re-assessing the situation following implementation of a given action.	**BM for good practice** • Re-assesses patient after treatment or intervention • Reviews situation, if decision was to wait and see • Continues to list options as patient's condition evolves	**BM for poor practice** • Fails to allow adequate time for intervention to take effect • Fails to include other team members in re-evaluation. • Is unwilling to revise course of action in light of new information

Table B9 - ANTS Checklist continued (Industrial Psychology Research Centre, Aberdeen University, 2003)

B05 - NOTSS

Situation Awareness: Developing and maintaining a dynamic awareness of the situation in theatre based on assembling data from the environment (patient, team, time, displays, equipment); understanding		
Gathering information Seeking information in the operating theatre from the operative findings, theatre environment, equipment, and people.	**Good Behaviours** • Carries out pre-operative checks of patient notes, including investigations and consent • Ensures that all relevant investigations (e.g. imaging) have been reviewed and are available • Liaises with anaesthetist regarding anaesthetic plan for patient • Optimises operating conditions before starting e.g. moves table, lights, AV equipment • Identifies anatomy/ pathology clearly • Monitors ongoing blood loss • Asks anaesthetist for update	**Poor Behaviours** • Arrives in theatre late or has to be repeatedly called • Does not ask for results until the last minute or not at all • Does not consider the views of operating room staff • Fails to listen to anaesthetist • Fails to review information collected by team • Asks for information to be read from patient notes during procedure because has not been read before operation started
Understanding information Updating one's mental picture by interpreting the information gathered, and comparing it with existing knowledge to identify the match or mismatch between the situation and the expected state.	**Good Behaviours** • Acts according to information gathered from previous investigation and operative findings • Looks at CT scan and points out relevant area • Reflects and discusses significance of information	**Poor Behaviours** • Overlooks or ignores important results • Misses clear sign (e.g. on CT scan) • Asks questions which demonstrate lack of understanding • Discards results that don't 'fit the picture'
Projecting and anticipating future state Predicting what may happen in the near future as a result of possible actions, interventions or non-intervention.	**Good Behaviours** • Plans operating list taking into account potential delays due to surgical or anaesthetic challenges • Verbalises what equipment may be required later in operation • Shows evidence of having a contingency plan ('plan B') (e.g. by asking scrub nurse for potentially required equipment to be available in theatre) • Cites contemporary literature on anticipated clinical event	**Poor Behaviours** • Overconfident manoeuvres with no regard for what may go wrong • Does not discuss potential problems • Gets into predictable blood loss, then tells anaesthetist • Waits for a predicted problem to arise before responding • Operates beyond level of experience
Decision Making: Skills for diagnosing the situation and reaching a judgement in order to choose an appropriate course of action.		
Considering options Generating alternative possibilities or courses of action to solve a problem. Assessing the hazards and weighing up the threats and benefits of potential options.	**Good Behaviours** • Recognises and articulates problems • Initiates balanced discussion of options, pros and cons with relevant team members • Asks for opinion of other colleagues • Discusses published guidelines	**Poor Behaviours** • No discussion of options • Does not solicit views of other team members • Ignores published guidelines
Selecting and communicating option Choosing a solution to a problem and letting all relevant personnel know the chosen option.	**Good Behaviours** • Reaches a decision and clearly communicates it • Makes provision for and communicates 'plan B' • Explains why contingency plan has been adopted	**Poor Behaviours** • Fails to inform team of surgical plan • Is aggressive/ unresponsive if plan questioned • Shuts down discussion on other treatment options • Only does what she/he thinks is best or abandons operation • Selects inappropriate manoeuvre that leads to complication
Implementing and reviewing decisions Undertaking the chosen course of action and continually reviewing its suitability in light of changes in the patient's condition. Showing flexibility and changing plans if required to cope with changing circumstances to ensure that goals are met.	**Good Behaviours** • Implements decision • Updates team on progress • Reconsiders plan in light of changes in patient condition or when problem occurs • Realises 'plan A' is not working and changes to 'plan B' • Calls for assistance if required	**Poor Behaviours** • Fails to implement decisions • Makes same error repeatedly • Does not review the impact of actions • Continues with 'plan A' in face of predictably poor outcome or when there is evidence of a better alternative • Becomes hasty or rushed due to perceived time constraints

Table B10 - NOTSS Checklist (Industrial Psychology Research Centre, Aberdeen University, 2006)

Communication and Teamwork: Skills for working in a team context to ensure that the team has an acceptable shared picture of the situation and can complete tasks effectively.		
Exchanging information Giving and receiving knowledge and information in a timely manner to aid establishment of a shared understanding among team members.	**Good Behaviours** • Talks about the progress of the operation • Listens to concerns of team members • Communicates that operation is not going to plan	**Poor Behaviours** • Fails to communicate concerns with others • Attempts to resolve problems alone • Does not listen to team members • Needs help from assistant but does not make it clear what assistant is expected to do
Establishing a shared understanding Ensuring that the team not only has necessary and relevant information to carry out the operation, but that they understand it and that an acceptable shared 'big picture' of the case is held by team members.	**Good Behaviours** • Provides briefing and clarifies objectives and goals before commencing operation • Ensures team understand the operative plan before starting • Encourages input from all members of the team • Ensures relevant members of team are comfortable with decisions • Checks that assistant knows what they are expected to do • Debriefs relevant team members after operation, discussing what went well and problems that occurred	**Poor Behaviours** • Does not articulate operative plan to team • Does not make time for collective discussion and review of progress • Fails to discuss the case beforehand with unfamiliar team members • Makes no attempt to discuss problems and successes at end of operation • Fails to keep anaesthetist informed about procedure (e.g. to expect bleeding) • Appears uncomfortable discussing the operative plan if challenged
Co-ordinating team activities Working together with other team members to carry out cognitive and physical activities in a simultaneous, collaborative manner.	**Good Behaviours** • Checks that other team members are ready to start operation • Stops operating when asked to by anaesthetist or scrub nurse • Ensures that team works efficiently by organising activities in a timely manner	**Poor Behaviours** • Does not ask anaesthetist if it is OK to start operation • Proceeds with operation without ensuring that equipment is ready
Leadership: Leading the team and providing direction, demonstrating high standards of clinical practice and care, and being considerate about the needs of individual team members.		
Setting and maintaining standards Supporting safety and quality by adhering to acceptable principles of surgery, following codes of good clinical practice, and following theatre protocols.	**Good Behaviours** • Introduces self to new or unfamiliar members of theatre team • Clearly follows theatre protocol • Requires all team members to observe standards (e.g. sterile field)	**Poor Behaviours** • Fails to observe standards (e.g. continues even though equipment may be contaminated or inadequate) • Breaks theatre protocol • Shows disrespect to the patient
Supporting others Providing cognitive and emotional help to team members. Judging different team members' abilities and tailoring one's style of leadership accordingly.	**Good Behaviours** • Modifies behaviour according to trainee needs • Provides constructive criticism to team members • Ensures delegation of tasks is appropriate • Establishes rapport with team members • Gives credit for tasks performed well	**Poor Behaviours** • Does not provide recognition for tasks performed well • Fails to recognise needs of others • Engages in 'tunnel vision' approach to technical aspects of operation • Shows hostility to other team members (e.g. makes sarcastic comments to nurses)
Coping with pressure Retaining a calm demeanour when under pressure and emphasising to the team that one is under control of a high-pressure situation. Adopting a suitably forceful manner if appropriate without undermining the role of other team members.	**Good Behaviours** • Remains calm under pressure • Emphasises urgency of situation (i.e. by occasionally raising voice) • Takes responsibility for the patient in emergency/ crisis situation • Makes appropriate decision under pressure • Delegates tasks in order to achieve goals • Continues to lead team through emergency	**Poor Behaviours** • Suppresses concern over clinical problem • 'Freezes' and displays inability to make decisions under pressure • Fails to pass leadership of case when technical challenge requires full attention • Blames everyone else for errors and does not take personal responsibility

Table B11 - NOTSS Checklist continued (Industrial Psychology Research Centre, Aberdeen University, 2006)

B06 - SPLINTS

Situation Awareness: Developing and maintaining overall awareness of relevant aspects of the theatre environment (patient, team, time, instrumentation and equipment) by watching and listening; understanding what the cues mean and anticipating what might happen next.		
Gathering information	**Example Good Behaviours**	**Example Poor Behaviours**
Actively seeking information in the operating theatre environment by observing, listening, questioning and recognising cues from the surgical process, theatre environment, equipment and people.	• Checks patient consent (immediately pre-operatively) • Demonstrates awareness of location of equipment and movement of staff on the floor • Watches surgical procedure • Conducts frequent scan of the environment • Collects information from other team members	• Fixates on one task • Distracted by non case-specific, inappropriate or irrelevant activity in theatre • Fails to listen to instructions • Fails to listen to conversations between other members of the team • Does not ask for information when appropriate
Recognising and understanding information	**Example Good Behaviours**	**Example Poor Behaviours**
Recognising and interpreting the information gathered from the theatre environment and comparing it with existing knowledge to comprehend the current state of events	• Attends to competing priorities appropriately • Recognises urgency if sudden changes in patient condition/ procedure • Switches between tasks efficiently • Provides correct instrument even when not named/ incorrectly described by surgeon • Reacts to conversational cues exchanged between other team members • Responds appropriately to changes in surgeon's body language/ tone of voice	• Does not change own activity level when appropriate • Does not prioritise tasks and/ or requests • Responds late or not at all to change in pace of procedure • Fails to seek clarification when faced with unclear commands or requests from other team members • Asks questions that indicate a lack of understanding
Anticipating	**Example Good Behaviours**	**Example Poor Behaviours**
Thinking ahead to predict what might happen and what could be required in the near future	• Hands appropriate instruments to surgeon in correct order • Predicts when plan of procedure is going to change; e.g. laparoscopy to open • Requests equipment from appropriate person before it is required by the surgeon • Times requests appropriately (e.g. warm saline, suction)	• Fails to respond to evolving surgical progress • Waits for a predictable problem to arise before requesting required instrumentation or equipment • Asks for items late • Loses track of surgical activity, i.e. is caught unaware
Communication and Teamwork: Sharing information, knowledge, goals and understanding among team members, to facilitate smooth progression through the surgical procedure.		
Acting assertively	**Example Good Behaviours**	**Example Poor Behaviours**
Using appropriate level of confidence to seek clarification/ make a point and adapting own manner of communicating to best facilitate effective teamwork.	• Seeks clarification when deviation from plan (e.g. procedure consented for/ position of patient) • Gives clear instructions/ requests to team members • Demonstrates leadership qualities when appropriate • Changes manner or tone of communicating to reflect situation • Demonstrates awareness of own limitations	• Passively accepts surgeon or other colleagues' decisions when challenging is a more appropriate response • Fails to communicate in a clear and precise manner • Adopts a subservient manner when a stronger response is required • Fails or is slow to communicate requirements • Uses off-hand manner of speech towards other team members
Exchanging information	**Example Good Behaviours**	**Example Poor Behaviours**
Seeks and gives enough detailed information to ensure a shared understanding among team members.	• Provides team members with information • Vocalises what is being given to surgeon to confirm request or where an alternative is available • Uses non-verbal signals where appropriate • Communicates that counts are correct/something missing in a timely manner	• Makes requests without specifying for whom communication is intended • Does not pass on/share important information (e.g. sharp blade, short suture) • Fails to articulate problems in a timely manner • Uses non-verbal communication where verbal clarification is more appropriate
Co-ordinating with others	**Example Good Behaviours**	**Example Poor Behaviours**
Interacting and working with other team members by sharing thoughts/ideas and performing physical tasks in a collaborative manner that facilitates the smooth flow of the surgical procedure.	• Communicates to other team members if there is a foreseeable change in plan/requirement to stop • Prioritises multiple concurrent requests from other team members • Suggests alternative options/ equipment • Deals appropriately with interruptions from others • Supports, provides help & assistance • Verbally acknowledges requests from scrub team members	• Fails to share information about evolving surgical plan • Talks to team members who are trying to concentrate • Ignores requests of others • Allows interruptions to disrupt flow of procedure • Fails to maintain awareness of whereabouts of other team members

Table B12 - SPLINTS Checklist (Industrial Psychology Research Centre, Aberdeen University, 2010)

Task Management: Organising resources and required activities to achieve individual and team oriented goals and maintaining standards with minimum stress to the team.		
Planning and preparing Organising requirements and timing them so that tasks can be completed with the minimum disruption to the smooth flow of the procedure/ list.	**Example Good Behaviours** • Demonstrates preparedness 4 does not make team wait unnecessarily • Utilises time during breaks in procedure for other/ preparatory tasks • Displays effective organisation of scrub practitioner workspace • Organises equipment • Prioritises tasks	**Example Poor Behaviours** • Confuses order of tasks which best promote a flowing surgical procedure • Opens sterile equipment/ supplies indiscriminately • Demonstrates difficulty in locating required equipment • Shows a lack of understanding of instrument purpose or sequence of usage
Providing and maintaining standards Ensuring patient and staff safety, adhering to codes of good practice and guidelines.	**Example Good Behaviours** • Protects sterile field and instrumentation • Controls volume of music and inappropriate conversation in theatre • Follows theatre guidelines and encourages others do likewise • Arranges for colleague to enter theatre if it appears surgeon would benefit from assistance	**Example Poor Behaviours** • Does not adhere to or violates approved protocols or guidelines • Distracted by/engages in irrelevant conversation with colleagues • Fails to check equipment settings/relies on others to do so • Does not display effective organisation of own workspace
Coping with pressure Dealing with stressful situations whilst maintaining a calm demeanour and understanding the demands and pressures for other team members.	**Example Good Behaviours** • Maintains an even tone of voice other than to indicate urgency (but without panic) • Does not rise to others' emotional outbursts • Organises and controls instrumentation appropriately • Takes initiative to delegate tasks where possible to ease pressure of situation	**Example Poor Behaviours** • Raises voice unnecessarily • Loses temper/ displays emotional outbursts • Appears disorganised and unable to locate instrumentation in a timely manner • 'Freezes' and unable to function effectively • Waits for instruction when should take action

Table B13 - SPLINTS Checklist continued (Industrial Psychology Research Centre, Aberdeen University, 2010)

B07 - BEHAVIOURAL MARKERS IN MARITIME CRISIS MANAGEMENT

Behavioural Marker	Example Positive Marker	Example Negative Marker
High degree of crewmember integration within the team	• Involves other crewmembers in planning, diagnostic and decision making processes • Accepts feedback and advice from crewmembers	• Isolates a crewmember or crewmembers • Only utilises crewmember or crewmembers by giving them a command order. • Ignores feedback or advice from crewmembers. • Specific case may be cultural isolation.
Regular soliciting of information between crewmembers	• Regularly asks other crew members for information	• Does not ask other crew members for information
Demonstrates awareness of 'big picture' context	• Verbalises task prioritise in anticipation of future events. • i.e. "we need to do xxxx now in yy minutes or zzzz will happen	• No anticipation of future events evident. • Acts surprised as events happen
Sharing of workload	• Delegates individual tasks in order to maintain situation overview	• Gets too involved in individual tasks to the extent that situation overview is lost.
The number of alternative hypothesis and actions communicated to team members	• Proposes alternative hypothesis and actions to team members.	• Does not propose any alternative hypothesis or actions to team members.
Uses complete and coherent sentences	• Communicates to team with complete and coherent sentences	• Uses unfinished sentences in communications with team
Uses measured patterns of movement	• Uses measured movements within a defined area where team leader is able to maintain overview of situation.	• Uses very fast movements. • No defined area of movement, so overview of instrumentation is not possible. • Focuses too much on one items of instrumentation. • Leaves control room for extended periods.
Crisis manager asking their team questions in order to elicit information so that they can improve their SA	• Asks team members questions about the situation in order to improve situational awareness	• Does not ask team members questions about their situation
Providing big picture updates	• Provides team members with updates on the overview of the situation	• Does not provide team members with any updates on the overview of the situation
Stating clear team and individual priorities	• States clear team priorities and individual task member priorities	• Does not state team or individual team member task priorities.
Utilises team briefings for specific threats	• Briefs team about specific threats	• Does not brief team about specific threats

Table B14 - Behavioural Markers in Maritime Crisis Management (Gatfield, 2008)

Behavioural Marker	Example Positive Marker	Example Negative Marker
Keeps crewmembers focused on the accomplishment of the task	• Provides encouragement and motivates team to keep them focussed on the accomplishment of the task. • Tasks completed in good time.	• Fails to monitor task progression. • Team does not complete tasks in good time.
Ensuring communications are audible and not garbled	• Communications clear and easy to understand	• Communications inaudible and garbled. • Communications not understood by team members.
Focuses on the dynamics of the complete system	• Maintains overview of all systems, thereby maintaining awareness of system interactions	• Focuses on one system to the exclusion of all others. • Is surprised by system interactions.
Communicating in a way that reveals ones mental models	• Communicates thoughts on the situation and how it is developing to team members	• Does not communicate thoughts to team members.
Team leader reflects on the suggestions made by other members	• Reflects on suggestions from team members before accepting or rejecting them.	• Acts immediately on suggestions from team members without any prior reflection. • 'Grasps' at suggestions of others.
Measured movements in response to stimuli	• Makes measured and controlled movements in response to stimuli	• Makes rapid and uncontrolled movements in response to stimuli
Moves smoothly and without hesitation	• Moves without hesitation. • Moves smoothly between tasks	• Movement truncated. • Movements hesitant. • Moves between two task locations without, in the short term, reaching either. • Presents an on the spot rocking motion
Team leader focuses on teams' tasks rather than on own individual tasks	• Maintains focus on overall team task objective.	• Focuses too much on own individual tasks that are not directly related to the overall team task objective. i.e. team leader spends a lot of time trying to put another alternator onto the main switchboard.

Table B15 - Behavioural Markers in Maritime Crisis Management continued (Gatfield, 2008)

B08 - NTSOD

Category	Element	Definition	Example Behavioural Marker
Leadership	Managing Watch Team	Effectively setting and maintaining the standards of the watch team.	The OOD utilized the dead time in the schedule to review the emergency procedures with the helmsman.
	Coping with Stress	Retaining a calm demeanour when under pressure and demonstrating to the watch that one is under control.	Despite the added pressure from the XO, the OOD managed the stress and performed proficiently.
Communications	Providing Information	Passing information along to other watch stations throughout the ship, as well as other assets in the area.	The OOD called the other ships in formation to inform them that the passing oiler was dimly lit and difficult to see.
	Issuing Orders	Effectively giving orders to other members of the watch team and other individuals as required.	The OOD ordered the Engineering Officer of the Watch to start another engine.
Situational Awareness	Gathering Awareness	Actively gathering information to keep up with the changing situation.	At two nautical miles out, the OOD visually inspected the contact through binoculars
	Understanding Awareness	Achieving an understanding of what the available information means.	The OOD identified the contact as a fishing vessel by Situation analyzing the lighting configuration.
	Anticipating Future Events	Forward planning in order to anticipate possible future problems.	The OOD had the Conning Officer drive slightly right of the intended track because he knew the wind and current would push the ship to the left.
Decision Making	Analytical Decision Making	Generating and comparing multiple courses of actions to come up with the optimal solution.	Once the oiler was located, the OOD decided to start driving towards her early to ensure that there was plenty of time to set up later.
	Following Orders & Procedures	Following documented procedures or direct orders from superior officers	The OOD used the wind envelope guide to make sure that the winds were sufficient to conduct flight operations.
	Intuitive Decision Making	Making quick decisions based upon prior experience and intuition.	When the mysterious light finally materialized as a sailboat 300 yards off the port bow, the OOD immediately ordered “Hard Right Rudder.”

Table B16 - Nontechnical Skills For Officers Of The Deck (NTSOD) (Adapted from Long, 2010)

B09 - RSSB NON-TECHNICAL SKILLS

Situation Awareness	Attention to detail
Good behaviour	**Bad behaviour**
A. Demonstrates ability to explain why particular details are important	A. Is sloppy in his or her work. Does not appreciate the need to attend to details
B. Pays attention to the details required to carry out a task eg uses the correct form	B. Overlooks important details in carrying out a task
C. Pays attention to the details required to understand the situation eg equipment displays or feedback	C. Overlooks important details that are necessary to carry out a task/ understand a situation
D. Identifies an anomaly in a complex situation eg physically responds to a fault or problem	D. Does not identify inconsistent or unusual information, ie does not respond or acknowledge in any way
Situation Awareness	**Overall Awareness**
A. Displays a good mental model of the situation: accurate understanding of what is happening	A. Does not bring information together to accurately understand what is happening in the situation overall
B. Regularly assesses the current situation, location and environment	B. Does not recognise the need, or have the capacity to regularly assess the situation, location and environment
C. Can balance attention between specific task and overall perspective of the situation	C. Loses awareness of the overall situation eg by getting overly absorbed in the detail of one particular task
Situation Awareness	**Maintain Concentration**
A. Actively controls distractions in a calm and proficient manner, maintaining performance and focus on primary task	A. Struggles to maintain focus on primary task when distractions arise eg performance suffers
B. Able to consistently maintain concentration eg listening or watching something for a period of time	B. Does not maintain attention / vigilance over time, becomes distracted
C. Remains alert during monotonous tasks	C. Does not maintain attention / vigilance on monotonous tasks, may seek distraction or vary task to make less monotonous
D. Demonstrates ability to provide information on the task which required concentration / vigilance	D. Unable to demonstrate that they were concentrating eg not able to describe the situation/task that they were concentrating on
E. Mindful and alert in carrying out tasks - responds appropriately and considers implications of actions	E. Lacks full concentration in carrying out tasks - slips into 'autopilot'
Situation Awareness	**Retain Information (During a shift)**
A. Recalls relevant information from immediately prior to or during shift, applying it as required eg carries out instructions, remembers emergency speed restrictions and stopping patterns	A. Does not apply information when required - has difficulties in remembering information
B. Uses memory aids to avoid forgetting important information eg writes down information for future use, uses own methods of remembering	B. Does not apply any techniques to try to remember information
Situation Awareness	**Anticipation of risk**
A. Anticipates what hazards, risks and errors could occur in a given situation before they happen	A. Does not demonstrate an understanding of the risks in a situation, may not notice until after they have
B. Acknowledges that errors and hazards can occur and constantly 'on the lookout' for dangers	B. Complacent or over-relaxed approach to the tasks and situation, not expecting any dangers to occur
C. Shows heightened alertness and vigilance when approaching known hazards eg changes driving technique	C. May be aware of a hazard but does not raise own levels of attention or alertness
D. Plans for potential future problems, eg is ready to respond and report to anything abnormal or contacting colleagues and making announcements	D. Over-anticipation; becomes overly focused on what may happen (at the expense of other tasks) E. May be aware of a hazard but does not act upon it by putting plans in place or reporting to appropriate people
Conscientiousness	**Systematic and thorough approach**
A. Takes an organised systematic unhurried approach	A. Disorganised approach to tasks, often rushed
B. Performs tasks in a systematic, logical Manner	B. Performs tasks in an unsystematic, illogical order
C. Is thorough in accomplishing a task and does not take shortcuts	C. Some or all of the task is completed without appropriate care and attention, or not completed at all
Conscientiousness	**Checking**
A. Checks information and does not make assumptions eg checks stopping pattern, checks all signals even though usually green	A. Does not check information, assumes it or their interpretation is correct
B. Ensures all necessary equipment is working	B. Does not check basic equipment eg assumes cab is correct on relieving rather than carrying out own check
C. Applies self-checking techniques and strategies eg risk triggered commentary	C. Does not apply self-checking techniques
D. Reviews safety outcome of the actions/ decisions they have taken	D. Does not check that own actions/decisions have resulted in a safe outcome
E. When asked, can explain why checking and reviewing is important	E. When asked, unable to explain the reason for checks/reviews

Table B17 - RSSB Non-technical skills for rail: A list of skills and behavioural markers for drivers (Adapted from Bonsall, 2012)

Conscientiousness	Positive attitude towards rules and procedures
Good behaviour	**Bad behaviour**
A. Correctly applies formal rules and procedures and acknowledged good practice eg completes required documentation	A. Does not comply with formal rules or procedures or acknowledged good practice eg rebels and is unwilling to follow instructions from others or take short-cuts in procedures
B. Self-disciplined attitude towards formal and informal rules and procedures	B. Haphazard application of rules. Only follows rules and procedures if prompted, regards as inappropriate
C. When asked, able to understand rationale for rules and procedures 4 thoroughly demonstrates underpinning knowledge of rules in training exercises	C. When asked, unable to explain the rationale behind a rule 4 can not appropriately justify why it exists
D. Maintains complete and organised Rule Book	D. Rule Book is messy, disorganised, incomplete
E. Takes action if others do not correctly adhere to rules and procedures	E. Takes no action if aware that others are not correctly adhering to rules and procedures
Communication	**Listening**
A. Able to listen to others, and understand information, and respond appropriately eg correctly carries out instructions, correctly repeats back instructions	A. Does not listen to information carefully 4 responds inappropriately eg repeats back instructions incorrectly
B. Follows instructions that have been given verbally	B. Does not follow instructions
Communications	**Clarity**
A. Gives clear instructions	A. Unclear communications eg provides ambiguous or vague information
B. Does not use local terminology or jargon	B. Uses local terminology and jargon
C. Writes legibly	C. Handwriting is difficult to read
D. Communicates concisely	D. Waffles in communication
E. Clearly explains complex issues and situations, and spells out words/ names that are difficult to pronounce	E. Explanations of complex issues and situations are unclear
F. Adheres to communication protocols	F. Does not follow communication protocols.
G. Able to communicate clearly and concisely in unusual and unexpected situations	G. Does not communicate clearly in unusual Situations
Communications	**Assertiveness**
A. Assertively states point of view eg in situations of peer pressure	A. Unclear communications eg provides ambiguous or vague information
B. Stands ground on basis of sound assessment eg continues to carry out actions as intended or does not give in to pressure to carry out a task (such as setting up the cab) before they are ready	B. Backs down unnecessarily when challenged
C. Uses appropriate tone of voice for situation	C. Becomes aggressive or meek when putting across point of view
D. Challenges others if information is conflicting or incorrect or actions are inappropriate eg challenges breeches of quality, safety and standards to ensure they are maintained	D. Reluctant to challenge others if information is conflicting or incorrect or actions are inappropriate eg allows examples of reduced safety and quality standards to go unnoticed/ unchallenged
Communications	**Sharing Information**
A. Shares information that is relevant to customers and colleagues, in appropriate level of detail	A. Does not share relevant information with colleagues or customers, or shares information that is irrelevant
B. Reports hazards to colleagues/ customers as appropriate	B. Does not report hazards to colleagues or customers
C. Completes relevant forms, providing appropriate level of detail	C. Does not complete relevant forms or omits important information
D. Shares information at appropriate time eg informs colleagues as much in advance as possible if unable to make duty on time	D. Does not correctly evaluate how time-critical information is and shares information too late
Decision Making and Action	**Effective Decisions**
A. Collects and analyses relevant information before making decisions eg asks for more information	A. Fails to consider relevant information before making decisions
B. Considers consequences of decisions ie risks and effect on others eg makes decisions that minimise risk or detrimental effect on other	B. Does not consider consequences of actions 4 makes irrational decision
C. Analyses information appropriately, applying knowledge accurately and exercising sound judgment ie makes justifiable decision eg who/ what to adhere to under certain circumstances	C. Decision is not based on a sensible consideration of all relevant factors, or inaccurate assumptions are made
D. Compares available options for action before making a decision eg verbalises options or discusses with	D. Does not consider alternative courses of action
E. Acts with certainty once a decision is made and takes responsibility for their actions	E. Indecisive or unable to justify decision
Decision Making and Action	**Timely Decisions**
A. Makes decisions and takes associated actions at the right time	A. Deliberates for too long, by the time a decision is made the situation has changed or more risk has been introduced
B. Instinctively carries out urgent actions	

Table B18 - RSSB Non-technical skills for rail: A list of skills and behavioural markers for drivers continued (Adapted from Bonsall, 2012)

Decision Making and Action	Diagnosing and solving problems
Good behaviour	**Bad behaviour**
A. Reviews what could have caused a problem, able to identify specific system fault or failure	A. Does not attempt to establish the cause of a problem, or unable to correctly identify the cause
B. Recognises that a situation requires a non-standard solution	B. Does not recognise an unusual situation and misapplies a standard solution
C. Finds an appropriate solution to address a non-routine situation	C. Does not apply an appropriate solution to a problem
D. Uses all resources available in handling a problem situation eg using all available information/ experience to diagnose and other people to mitigate the risks	D. Does not make use of all the sources of information/ help available
Cooperation and working with others	**Considering others' needs**
A. Aware of other's roles and priorities and so takes into account others' point of view	A. Does not seek others views or incorporate them into final action
B. Discusses options and consequences with other personnel	B. Lacks awareness of colleagues' roles and responsibilities
C. Takes into account others' values and beliefs	C. Ignores others' values and beliefs
Cooperation and working with others	**Supporting others**
A. Works cooperatively with others	A. Does not take actions/ share information required to enable colleagues to do their job effectively
B. Will stop to help others when possible	B. Does not help others if it inconveniences themselves (when it was possible and appropriate)
Cooperation and working with others	**Treating others with respect**
A. Polite in verbal communication	A. Rude or aggressive in dealing with others eg uses offensive language
B. Appropriate non-verbal communication	B. Inappropriate non-verbal communication
C. Treats others with respect regardless of their culture, age, background etc	C. Demonstrates sexist, racist or other intolerant behaviours
Cooperation and working with others	**Dealing with conflict/aggressive behaviour**
A. Recognises inappropriate behaviour and plans to take action	A. Does not recognise inappropriate behaviour and does not respond by making plans for action
B. Reduces conflict where possible	B. Does not attempt to reduce conflict, or aggravates conflict situation by overreacting
C. Calls for assistance in a conflict situation when required	C. Does not seek support in conflict situations where it is necessary to do so
Workload management	**Multi-tasking and selective attention**
A. Able to perform different tasks in parallel when safe to do so	A. Appears to be unable to do more than one task at a time
B. Differentiates between different tasks/processes and the attention they demand	B. Does not differentiate between elements/information that demand attention
C. Switches attention between sources of information	C. Appears to be unable to switch attention, or does so at inappropriate times
Workload management	**Prioritising**
A. Attempts to prioritise when under pressure eg responding to emergency call once completed current task	A. Does not prioritise, attempts to multi-task at inappropriate times
B. Can identify the most crucial information/tasks and prioritise accordingly eg stops a task to address another more important priority	B. Does not distinguish between important and less important tasks and information, prioritising ineffectively
C. Changes normal working practices, is flexible, re-prioritising based on changes to situation	C. Has a fixed way of working and is reluctant to re-prioritise tasks based on changes in the situation
Workload management	**Calm under pressure**
A. Able to maintain performance under pressure eg thinks rationally and clearly and able to perform logical actions and follow appropriate procedures, slowing down if necessary so that task can be completed correctly	A. When under pressure does not maintain performance eg becomes confused and acts irrationally with regard to safety - could start manipulating the wrong controls
B. Maintains control of emotions in a stressful situation eg recovers quickly from setbacks, errors and obstacles	B. Does not maintain an appearance of being calm and in control when under pressure eg raises voice, becomes aggressive
Self-management	**Motivation**
A. Willing and motivated to do the basic requirements of the job as well as possible	A. Does not demonstrate motivation to effectively fulfil basic requirements of the job eg reluctantly carries out everyday tasks
B. Does more than the minimum required to get the job done eg seeks out and accepts additional tasks or responsibilities	B. Does not demonstrate motivation to do more than the basic requirements of the job
C. Manages lifestyle to reduce the effects of fatigue	C. Regularly and consistently prioritises other (inappropriate) commitments over work
D. Motivated to perform well beyond situations when being observed - on train data recorder (OTDR) download suggests careful driving strategy	D. Allows standard of own performance to drop when not being observed - OTDR suggests risky driving strategy

Table B19 - RSSB Non-technical skills for rail: A list of skills and behavioural markers for drivers continued (Adapted from Bonsall, 2012)

Self-management	Confidence and initiative
Good behaviour	**Bad behaviour**
A. Works autonomously and makes decisions and takes responsibility as required	A. Lacks confidence/ initiative to work without over-checking eg regularly checks things when inappropriate
B. Confidently contributes to work-related discussions	B. Lacks confidence in own views
C. Has an accurate level of confidence in own ability - works within areas of competence and authority, contacting others when appropriate	C. Overconfident, failing to contact others when required
D. When required to delegate tasks eg in an emergency, does so with appropriate consideration of others' areas of competence	D. When required to delegate tasks in emergency situation does not delegate or does so inappropriately eg fails to take lead when most competent staff member
Self-management	**Maintain and develop skills and knowledge**
A. Keen to learn eg asks questions, seeks out and responds enthusiastically to learning opportunities such as computer based training stations	A. Reluctant to attend training or develop knowledge eg refuses to attend or makes excuses
B. Seeks and acts on feedback, identifying areas for development and responding accordingly	B. Lack of interest in what others think of performance, responds badly to feedback
C. Reflects on own performance and willing to learn from own/ other's mistakes	C. Does not reflect on own performance and does not attempt to learn lessons from mistakes that have been made by self/others
D. Builds and maintains own knowledge of rules/procedures eg gives up to date answers to questions/carries out up to date actions/Rule Book is up to date	D. Allows knowledge to lapse or does not update knowledge eg unaware of Rule Book changes or rule book is out of date
Self-management	**Prepared and organised for duty**
A. Seeks and reads through all relevant information in advance of tasks	A. Does not read through information as required.
B. Is well prepared, eg full uniform, correct equipment	B. Is unprepared, eg not in full uniform, lacking equipment
C. Arrives on time for duty and all portions of work	C. Arrives late for duty or other portions of work

Table B20 - RSSB Non-technical skills for rail: A list of skills and behavioural markers for drivers continued (Adapted from Bonsall, 2012)

B10 - EUROCONTROL BOOM OBSERVATION SHEET

Part A : GENERAL INFORMATION

Figure B4 - Behaviour Oriented Observation Method (BOOM) Observation Sheet (Eurocontrol, 2003)

Date: **Observer:** **Position observed:**

Sector and simulation information:

Give ID or time code from video recorder if available :

Part B: OBSERVATION NOTE BOOK

IDENTIFIED BEHAVIOUR	CONTEXT IN WHICH BEHAVIOUR WAS PERFORMED	TRM DOMAIN(s) COVERED BY THIS BEHAVIOUR
Describe the behaviour (event, action)	(workload, aircraft, team atmosphere, traffic)	

Part C1: INTERPRETATION OF THE BEHAVIOUR (First guess in pencil)

Part C2: QUESTIONS TO ASK THE BOOMEE TO CLARIFY WHAT WAS HAPPENING FROM HIS/HER POINT OF VIEW

Part C3: QUESTIONS TO ENQUIRE ABOUT THE AWARENESS OF THE POTENTIAL EFFECTS OF THIS BEHAVIOUR ON TEAM PERFORMANCE AND SAFETY

Part D: OUTCOME OF DISCUSSION WITH THE BOOMEE

(Identified behaviours that positively or negatively impact performance of the team and/or

Part E: IDENTIFIED POSSIBLE IMPROVEMENT (and how this will be achieved)

Part F: GENERAL COMMENTS

AGREEING THE CONTENT (OPTIONAL) **Yes**

B11 - EUROCONTROL BOS

Inferior Behaviour	Observed behaviour	Superior Behaviour
The student doesn't pre-plan and analyse the traffic, and works in an ad hoc and chaotic manner.	**1. Organises and maintains an efficient traffic flow**	The student solves complex crossing inbound and departing traffic using all possible means, e.g. different flight levels, radar vectoring as well as ROC/RoD in order to achieve continuous climb and continuous descent for all aircraft concerned.
The student while working in a position with many entry conflicts fails in due time to detect them and take the appropriate actions to facilitate the workload of the executive controller.	**2. Detects conflict early**	The student training as co-ordinating controller solves under high workload multiple entry conflicts by early identification of conflicts e.g. climbing A/C to be levelled out to reduce the number of entry conflicts, thereby reducing the workload of the executive controller, so he/she can solve remaining conflicts in an optimal way.
The student fails to take appropriate actions in time to solve complex traffic scenarios and takes impulsive decisions after STCA warnings.	**3. Resolves conflicts effectively**	The student training as co-ordinating controller, with a complex crossing situation with one entry to a sector from different adjacent sectors, applies solutions that are in agreement with other controllers, and quickly and correctly assesses distances, speeds and destinations. He/she will also contribute to the executive controller being able to work in a safe and simple way.
The student fails to detect that an a/c is climbing towards its cleared FL and over-shoots the cleared FL.	**4. Detects deviations**	The student monitors an a/c navigating towards unexpected exit point, not indicated on the FPL. He/she quickly tries to find the root of the deviation and prepares himself/herself to intervene if the a/c indeed is off track.
The student does not react to an a/c deviating from its assigned flight path and heads toward military airspace without prior coordination.	**5. Corrects deviations**	The student monitors an a/c on the radar screen not following the clearance toward a fix, he/she than informs the pilot that he/she is not navigating according to the clearance, and is prepared to radar vector the pilot when confirmed s/he is off track.
The student cannot revert and operate the fall back system in an appropriate way and loses updated traffic situations, which causes increased workload for colleagues.	**6. Operates technical systems**	The student selects the radio back up system quickly and correctly when the regular R/T system fails, so no disturbances occur.
The Student makes mistakes and/or does not know the required ATC Technical terminology.	**7. Masters the required ATC technical terminology**	The student detects an a/c with the wrong MODE C indication, the student using standard phraseology gets a confirmation of the status of the Mode C and instructs "stop squawk C".
The student manages the medium level of traffic, but starts making small mistakes when time elapses.	**8. Maintains attention over the entire shift**	The student keeps his/her attention and focuses on the job during the whole shift.
The student opposes and questions the feedback given by Instructors	**9. Insight into own limitations**	In a coach/instructor debrief, the student openly states what his/her limitations are, and they are also in line with others' observations.

Table B21 - Behavioural Observation Scale (BOS) (Adapted from Eurocontrol, 2005)

Inferior Behaviour	Observed behaviour	Superior Behaviour
In highly complex traffic on the limit of his/her capacity, the student continues to accept traffic and to climb/descend traffic.	**10. Adapts to own limitations**	A student responds to a pilot's requests of a high FL in a busy traffic situation, includes the fact that the a/c must descend again through heavy traffic, by using a FL with as little penalty as possible to the a/c and also with reduced workload.
The student does not inform the executive controller of a wrong read-back from a pilot he/she has heard, and later blames the executive controller for not having heard it.	**11. Teamwork skills**	When the student training as co-ordinating controller realises that the executive controller is becoming overloaded, he/she adjusts working priorities in order to assist the executive controller, e.g. by planning a/c exiting the sector at levels easy to handle for the executive controller.
The student does not react to pilot request to go round CB and instructs the pilot to maintain the heading.	**12. Ability to identify with the pilots and understand their needs**	The student reacts to a pilot's emergency message about explosive decompression, by acknowledging the pilot's transmission. The student gives all necessary traffic information, effectuates the necessary co-ordinations with partners, vectors any other a/c out of the path of the a/c in emergency; when convenient without disturbing the pilot during emergency descent, comes back to the pilot and offers all appropriate assistance.
The student ignores the company rule of requiring release for climb in an adjacent sector because he/she feels he/she has the "whole picture" of the adjacent sector.	**13. Willingness to work according to company rules**	The student willingly takes on the norms, like always being in position 5 minutes before the start of duty as a well as following the company rules and procedures on sick calls, mission requests etc.
The student takes continuous telephone calls from other units without appropriate inputs and communicating the co-ordination to the executive controller.	**14. Works in an orderly way under pressure**	When the student sees that others are trying to call, he /she quickly completes on-going co-ordination in a correct fashion, takes the next call, and remembers and takes actions on the previous calls.
The student after communication with outboard a/c assumes the wrong gate point, doesn't double-check the flight planned route and hands over the a/c to the wrong sector.	**15. Ability to detect and correct own mistakes**	The student either clears an a/c back to its original route or vectors it to avoid active danger restricted areas, after having first sent it towards it.
The student allows a/c to climb through many Flight Levels to the requested high FL, even when it must descend 10 minutes later to the original FL	**16. Does not give in to pilot's demands when they are in conflict with own view**	The a/c in heavy traffic requests to proceed direct to a point because it is behind schedule, the student refuses the request and explains that it would lead to the traffic situation becoming too complex and impact everybody's workload in a negative way.
The student creates conflicts in adjacent sector by routing two a/c at same FL towards the same entry point, when this could be avoided by direct routing or level change.	**17. Controls in a way that does not create problems for other controllers**	The student delays a/c in own sector before moving it to adjacent sector when he/she sees the adjacent sector already has extremely high traffic and workload.
The student gets angry at military colleague, when being refused to send a/c direct through military airspace when there is no apparent traffic in that airspace.	**18. Shows consideration for colleagues**	The student accepts that even his/her colleague can produce less than brilliant solutions and has no problem taking advice or information from him/her.
A Boeing 747 departing on a transatlantic flight. The student issues an unrealistic rate of climb, e.g. 2500' - 3000' per minute, with the effect that the pilot responds that they are unable to comply. The impression of the pilots and the coaches is "unprofessional".	**19. Thorough knowledge of aircraft and their characteristics**	The student rejects an Aircraft type with poor climb performance to climb from FL 300 to 340, when close to a busy sector boarder and fix to another FIR. The student informs the a/c to contact next sector for climb permission and informs that he/she is released.
A restricted area that has been de-activated becomes activated. This information is displayed. The student misses the information and clears the aircraft through the restricted area.	**20. Constantly checks available information on incoming and outgoing data**	The student working as co-ordinating controller frequently checks the OSDL pages on the CCTV for information on e.g. activation of danger areas to be passed on to the executive controller to avoid inadvertently clearances through e.g. danger areas.

Table B22 - Behavioural Observation Scale (BOS) continued (Adapted from Eurocontrol, 2005)

Inferior Behaviour	Observed behaviour	Superior Behaviour
The student keeps the aircraft within his/her upper sector to save fuel (pleasing the pilots) before descending it to a lower sector to avoid a restricted area. The colleague in the lower sector gets the a/c too late to safely descend it further to avoid the restricted area.	**21. Ensures traffic safety while at the same time taking account of economic aspects**	In a high volume of traffic with multiple departures, the student facilitates continuous climbs for the departing traffic by using heading and direct routing`, in order to get them to their cruising level at the earliest opportunity.
The student knows that a/c has to meet certain FL restrictions in order for lower sectors to continue descent. The student waits too long to descend the inbound a/c because of traffic crossing and opposite traffic, so a/c is unable to meet FL restrictions.	**22. Takes decisions quickly and in a responsible manner with due regard for priorities**	An a/c reports a medical emergency. The student diverts the a/c towards the closest major airport, facilitates the descent, gives the appropriate clearances, co-ordinates and informs the airport about the necessary assistance on the ground.
The student refuses an a/c which declares emergency to dump fuel over land before attempting to land at the nearest airport and instead attempts to vector it over sea to dump fuel.	**23. Improvises in situations requiring unconventional approaches or solutions**	An a/c requests diversion due to technical problem and requests fuel dumping. The student immediately checks for conflicting a/c and assures he/she has vertical separations to other a/c before he/she issues permit to dump fuel over land and diverts the a/c to the nearest aerodrome at the earliest opportunity.
A student is told on the phone that "LFPG" (Paris de Gaulle) is closed and no inbound traffic is accepted. The student passes on that "Paris" is closed and no traffic accepted. This creates a lot of confusion among other controllers and supervisors (whether all airports in Paris are closed and if the TMA is closed), before the situation has become clear and correct information passed on.	**24. Communicates in a clear, unambiguous and to-the point manner**	When R/T becomes congested, the student communicates in a manner that prevents the need to repeat clearances.
The student has identified a conflict after giving a clearance, and is instructing the a/c to climb/descend, without emphasising the urgency in carrying out the instruction.	**25. Adjusts tone of voice for messages in special situations**	In military airspace the student detects an unidentified track with a similar mode C in the military area, which is approaching an a/c flying close to the border. The student transmits to all a/c and passes on this traffic information with careful intonation to avoid repetition and clarifications.
The student applies a new procedure in such a way that it clearly demonstrates that he/she has not understood the information given.	**26. Profound knowledge and use of English language**	The student studies a letter of agreement (LOA) to the extent that he/she understands what it encompasses, what is agreed and what is not agreed, and any likely grey areas that could arise.
The student reacts in a mechanical and non coping way to an increase of traffic and complexity.	**27. Shows stability (emotional control) in crisis situations**	After the radar picture denigrates to back-up radar, the student continues to work in efficient ways and keep discussions and opinions on what went wrong to later.
The student fails to give support and to alleviate adjacent student not coping with high workload.	**28. Gives support to others if needed**	The student shares his/her knowledge and experience of a procedure with another student who shows difficulties in understanding and applying it.
The student pays more attention to status and benefits than actually trying to learn the job.	**29. Shows identification with the job**	The student shares his/her view of the controller work, by explaining the technical, service and social side of the job, excluding salaries and myths, when giving presentations to external visitors or being asked questions about the controller job.
The student prepares for examination only by studying the hand outs and seeks neither other available information nor approaches ops. Room staff with questions.	**30. Shows initiative and motivation**	The student requests own copy of a new document on system features and changes without being asked to do it, or having an examination on the subject.
The student demonstrates a pattern of hesitance (or no solutions) and accepts bad advice from the co-ordinating student controller.	**31. Demonstrates leadership**	The student sets an example for the team and adjusts his/her working method by refusing an a/c or/and requesting a/c at conflict free levels, when he/she sees the partner controller (executive or co-ordinating) becoming over loaded.

Table B23 - Behavioural Observation Scale (BOS) continued (Adapted from Eurocontrol, 2005)

Inferior Behaviour	Observed behaviour	Superior Behaviour
The student argues and disputes the instruction and feedback given by the coach.	**32. Accepts and deals constructively with criticism**	The student attentively processes the advice given from the OJT on how a difficult traffic situation could be solved to be able to better solve it him/herself next time.
The student turns to the coach in a questioning way in routine traffic situations when the student should execute familiar solutions.	**33. Demonstrates self-confidence**	The student seeks clarification from head of training or training officer to resolve the issues arising from contradictory instructions and advice from different OJTs.
The student hesitates and pauses before issuing clearances to the pilot	**34. Demonstrates authority and decisiveness**	The student tries to provide the best service possible, even when this means refusing pilots' request for direct routing, in order to maintain safe, orderly and expeditious service.
The student applies solutions which are too rapid for a given situation, e.g. not using Opposite Direction Levels (ODL) when it is accepted to achieve expeditious air traffic.	**35. Demonstrates Flexibility**	The student easily changes an intended solution on a difficult crossing situation, when new information or facts are added to its complexity.
The student loses belief in a successful outcome when going through a difficult phase of learning	**36. The trainee shows ambition to reach training goals**	The student consults his/her training officer on what leave period to have, so his/her learning process and progress are not hindered.
The student's learning curve has flattened out after extended OJT and the student shows no progress in reaching the required skill levels.	**37. The trainee develops ATCO skills in appropriate time**	The student shows faster progress than the average student.
The student gets defensive in briefs, is not willing to recall or discuss anything arising from events.	**38. The trainee is easy to handle for coaches**	The student takes own initiative in the preparation of an oral exam and collects information from others on the relevant subjects, without waiting for scheduled briefing sessions with the training officers.

Table B24 - Behavioural Observation Scale (BOS) continued (Adapted from Eurocontrol, 2005)

B12 - LVNL ATC COMPETENCY ASSESSMENT

Category	Competencies
Safety	Applies separation minima correctly
	Switches from monitoring to vectoring in time
	Builds in safety buffers sufficiently
	Transfers traffic to the adjacent ATC unit correctly and in time
Efficiency	Applies vector technique correctly
	Applies speed control correctly
	Takes into account differences in height between aircraft
	Takes into account differences in aircraft performances
	Minimises his/her own workload as much as possible
	Creates an optimal sequence of descending and climbing traffic streams
Verbal expression	Expresses him/herself concisely, to the point, unambiguously and firmly
	Has a clear, quiet pronunciation and intonation
	Expresses him/herself in the Dutch language to conform with ICAO level 2
Listening	Listens out well, understands messages well
	Interprets messages well and reacts adequately
	Is easily approachable for others at the sector
	Has an understanding of the Dutch language to conform with ICAO level 2
Coordination	Coordinates in time and with the appropriate ATC function/unit
	Communicates his/her plan concisely and to the point
	Makes clear arrangements and acknowledges these correctly
	Gives and received radar handovers correctly
Equipment operation	Applies valid procedures and working agreements correctly
	Makes correct use of equipment
Strip label management	Updates labels correctly
	Works with strips correctly
Mental picture	Keeps a clear overview of the situation by scanning regularly
	Looks, observes and takes action if necessary
	Controls the accuracy of available information
	Guards the identification process and the label presentation
	Anticipates future and variable traffic situations
Attention management	Divides attention between several situations sufficiently
	Performs several actions simultaneously
	Holds information in his/her memory without forgetting it
Planning	Is creative in inventing various solutions
	Plans according to valid procedures and agreements
	Is flexible in adjusting plans
Decision making	Takes initiative and acts when necessary
	Shows confidence in taking the lead
	Acknowledges priorities correctly
Workload management	Adapts work tempo to traffic load optimally
	Stays calm; also during hectic moments
Attitude	Shows responsibility during work
	Takes his/her training seriously
	Is eager to learn
Team orientation	Collaborates with others easily
	Is willing to adapt to common standards and values

Table B25 - Progression report (preOJT) (Adapted from Oprins, 2008)

	Category	Competencies
	Efficiency / Optimum Sequence	Applies vector technique correctly
		Applies speed control correctly
		Takes into account differences in height between aircraft
		Takes into account (differences in) aircraft performances
		Minimises his/her own workload
		Realises an optimum sequence of aircraft
		Applies continuous climb/descent
		Takes into account winds aloft
	Verbal Expression	Expresses him/herself well in '(non)-standard' phraseology
		Expresses him/herself concisely, to the point, unambiguously and firmly
		Has a clear, quite pronunciation and good intonation
	Listening	Listens out well, understands messages well
		Interprets messages well and reacts adequately
		Is easily approachable for others at the sector
	Coordination	Coordinates in time and with the appropriate ATC function/unit
		Communicates his/her plan concisely ant to the point
		Is able to make arrangements and acknowledges these correctly
		Applies valid procedures and working agreements correctly
		Gives and receives handovers correctly
	Application of valid procedures / apparatus	Applies valid procedures and working agreements correctly
		Makes correct use of apparatus
	Strip- and label handling	Updates labels correctly
		Works with strips correctly
	Mental picture / perception	Keeps a good overview of the situation by scanning regularly
		Looks, observes and takes action if necessary
		Checks available information to be correct
		Guards the identification process of the label presentation
		Anticipates on future and variable traffic situations
	Attention Management	Can divide attention between several situations sufficiently
		Can perform several actions simultaneously
		Can park information in his/her memory without forgetting it
	Planning	Is creative in inventing various solutions
		Can plan according to valid procedures and agreements
		Is flexible in adjusting plans
	Decision Making	Takes initiative and acts
		Shows confidence and takes the lead
		Acknowledges priorities correctly
	Workload Management	Adapts work tempo to traffic load
		Stays calm, also during hectic moments

Table B26 - Final Test ACC. (Adapted from Oprins, 2008)

B13 - FAA SACHA

Communicating and Informing - uses clear concise accurate language to get message across unambiguously, talking only when necessary and appropriate; employing proper phraseology to ensure accurate communication; notifying pilots/controllers/other personnel of information that might affect them as appropriate; issuing advisories and alerts to appropriate parties; listening carefully to requests and instructions and ensuring that they are understood; attending to read backs and ensuring they are accurate.		
Low Performance: Is consistently too wordy, imprecise in phraseology, or uses slang inappropriately during transitions to pilots and other controllers Is careless about informing pilots concerning circumstances that affect them such as weather, nearby traffic etc. Often fails to ensure that own instructions are understood; is not very good at picking up on errors in pilot read backs of clearances, course changes etc.	**Middle Performance:** Radio and interphone communications are usually easy to understand; at times, may be somewhat wordy or use unambiguous phraseology on the air Is normally good at informing pilots about situations and conditions that affect them (e.g. safety related items) For the most part checks to be certain that own instructions are understood; only occasionally fails to pick up on inaccurate read backs from pilots	**High Performance**: Always uses clear, concise phraseology when talking to pilots or other controllers; is very easy to understand Consistently provides pilots with the information they need such as timely safety alerts, weather advisories, warnings about unpublished obstructions Always ensures that own instructions are clearly understood; pays careful attention to pilot read backs of clearances
Managing Multiple Tasks - Keeping track of a large number of aircraft/events at one time; conducting two or more tasks simultaneously; remembering and keeping track of aircraft and their positions; remembering what you were doing after an interruption; returning to what you were doing after an interruption and following through; providing pilots with additional services as time allows		
Has difficulty keeping track of several aircraft at the same time; may focus too narrowly on some aircraft while ignoring other Is ineffective at performing multiple tasks simultaneously; prefers to take one thing at a time Interruptions and distractions often cause him/her to forget about some of the immediate air traffic problems; may be slow in recalling what he/she intended to do before the interruption	Keeps on top of movement of several aircraft simultaneously while also dealing with routine communication; when very busy may have to simplify the situation to reduce the number of things to attend to Is good at performing two or sometimes more routine tasks at the same time (e.g. monitoring the screen, talking with pilots and handling strips) After an interruption, can usually handle the air traffic problems remaining from prior to the interruption successfully	Is extremely adept at keeping track of many aircraft while at the same time handling pilot communications, strip work, etc. Is fully capable of performing two or more complex tasks simultaneously After an interruption, always quickly remember where aircraft are or should be, what he or she was doing with the traffic before the interruption, and the intended control strategy
Technical Knowledge - Knowing the equipment and its capabilities and using it effectively; knowing aircraft capabilities and limitations (e.g. speed, wake turbulence requirements) and using that knowledge; keeping up-to-date on letters of agreement, changes in procedures, regulations, etc.; keeping up-to-date on seldom used procedures or skills		
At times, may not remain current on new letters of agreement, revised air traffic procedures, etc. Has basic knowledge of most aircraft's' capabilities, but may make errors related to not knowing aircraft limitations May be unfamiliar with some of his/her equipment and how it works	Is usually knowledgeable about and up-to-date on all information relevant to controlling traffic (e.g., letters of agreement, air traffic procedures, etc.) Has good knowledge of different aircraft capabilities and applies that knowledge to avoid most errors associated with not knowing aircraft limitations Is reasonably familiar with his/her equipment and how it works	Always keeps up-to-date on letters of agreement, all pertinent procedures and policies, any sector-specific changes (e.g., revised boundaries) Has thorough knowledge of different aircraft capabilities and as a result never mates errors such as climbing an aircraft beyond its limits, making an inappropriate speed assignment, or requiring an impossibly tight turn Is extremely knowledgeable about and familiar with his/her equipment and how it functions.
Reacting to Stress - Remaining calm and cool under stressful situations; handling stressful air traffic conditions in a professional manner.		
Becomes shaken and ineffective in emergency situations. Reacts poorly and performance suffers under stressful air traffic conditions. Does not maintain his/her composure when serious problems arise.	Remains calm and cool in most emergency situations. Stays calm, focused, and functional under busy and/or somewhat stressful conditions. Shows professional cool in handling routine problems.	Remains very calm and cool and reacts effectively even in very serious emergency situations such as in-flight emergencies, lost pilots, VFR pilots in IFR conditions, etc. Stays calm, focused, and very functional in busy, and very stressful conditions (e.g., sudden weather problems that severely reduce usable airspace). Handles even serious problems with professional cool.

Table B27 - FAA Separation and Control Hiring Assessment (SACHA) Competency Assessment (Adapted from Borman et al, 2001)

Maintaining Attention and Vigilance - Scanning properly for air traffic events, situations, potential problems etc. keeping track of equipment weather status; identify unusual events, improper positioning of aircraft, recognizing when aircraft have potential for loss of separation verifying visually that control instructions are followed, remaining vigilant during slow periods.		
Has a tendency to focus too narrowly on one air traffic problem and sometimes fails to scan the radar scope for other potential problems with conflictions, traffic flow, weather, etc. Often does not recognize that an action is required; is often lax in watching the radar scope and tends to significantly reduce vigilance during slow periods.	For the most part, properly scans the scope and monitors aircraft to maintain awareness of air traffic events, potential problems, etc. Is attentive to the radar scope and maintains vigilance, especially during rush periods; may occasionally be less attentive when traffic is light.	Consistently recognizes potentially dangerous conditions such as errors made by pilots (e.g., wrong turns, descending or climbing through assigned altitudes, etc.). Always monitors the radar scope to ensure that clearances and other instructions to pilots are followed; remains highly vigilant, even during slow periods.
Prioritising - Taking early or prompt action on air traffic problems rather than waiting or getting behind knowing what to do first and identifying the most important situational recognising that some problems or situations are less important and can wait; preplanning before busy periods; organizing the boards and using flight strips effectively to keep priorities straight for handling air traffic situations quickly and decisively determining appropriate priorities.		
Has difficulty recognizing which air traffic problems are the most pressing; may deal with problems in chronological order, or take the easy ones first. Often fails to prioritize activities, acting on air traffic problems without evaluating the possible consequences of own actions. Puts off decisions and actions that should be taken right away.	Generally recognizes the most important air traffic problems and handles them before the less pressing ones. When prioritizing own actions, normally looks ahead to assess potential air traffic problems that might result from own actions. Usually takes early or prompt action to deal with air traffic problems.	Always recognizes which air traffic problems need immediate attention and handles them before less pressing ones; consistently uses appropriate priorities for control actions. Prioritizes activities with extreme effectiveness, consistently looking ahead and accurately predicting problems that will result from revised clearances, rapidly degrading weather, etc. Invariably takes early or prompt action to resolve air traffic problems.
Maintaining Safe and Efficient Traffic Flow - Reacting to and resolving potential conflictions effectively and efficiently; using proper air traffic separation techniques effectively to ensure safety; sequencing aircraft effectively for arrival or departure; sequencing aircraft to ensure efficient/timely traffic flow; controlling traffic in a manner that ensures efficient traffic flow; controlling traffic in a manner that minimizes traffic problems (e.g. conflictions, traffic flow problems) for other controllers and pilots.		
Sometimes fails to maintain minimum separation or to recognize and resolve potential conflictions. Uses control actions that fail to resolve potential conflictions or that result in excessive workload (e.g., waits until potential conflictions are critical before taking action, fails to take wind into account, etc.) Does not always sequence aircraft adequately or ensure proper spacing between aircraft; may cause excessive and unnecessary delays by choosing poor control actions, waiting too long to provide needed commands, unnecessarily vectoring or rerouting aircraft, etc	Typically uses appropriate control actions to maintain proper separation or to resolve potential conflictions. Resolves simple conflictions and traffic flow problems without causing unnecessary delays. Generally uses correct procedures to sequence and space aircraft safely; maintains smooth traffic flow, but may not use the most efficient control actions (e.g., may not always take aircraft types into account).	Consistently maintains safe, efficient, and orderly traffic flow, even under difficult or unusual circumstances (e.g., extremely heavy traffic, bad weather, etc.) Consistently recognizes potential problems or conflictions well in advance and takes highly effective action to maintain separation and efficient air traffic flow. Sequences and spaces traffic effectively and efficiently, even when extremely busy (e.g., by taking aircraft types into account); always maintains proper separation while minimizing delays (e.g., avoids delaying vectors as appropriate).
Adaptability and Flexibility - Reacting effectively to difficult equipment problems, changes in weather, traffic situations, etc, or to unexpected actions on the part of other controllers or pilots; using contingency or fall-back strategies effectively when unforeseen/unanticipated air traffic problems emerge or if firs plan doesn't work; asking for help when it's needed; developing/executing innovative solutions to air traffic problems; dealing effectively with situations from which there may not be clearly prescribed procedures, situations which require novel thinking; adapting to equipment updates, new procedures etc.		
Does not adjust well to unusual and difficult air traffic situations. Rarely displays good "fall-back" strategies for dealing with unanticipated air traffic problems. Is ineffective at handling air traffic situations with no clearly prescribed procedures.	Is usually able to adapt effectively to most situations such as worsening weather, equipment problems, etc. Frequently, but not always, has effective contingency strategies for unforeseen or unanticipated air traffic problems when they arise. For the most part, is good at handling air traffic situations that have no "textbook answers," but does better with the more routine problems.	Reacts expediently and effectively to even the most complicating events (e.g., quickly devises and executes a complex re-route plan for several aircraft when thunderstorms begin forming). Is very adept at using effective contingency or "fall-back" strategies when unforeseen or unanticipated air traffic problems arise. Deals effectively with even very difficult air traffic situations where there are no clearly prescribed procedures.

Table B28 - FAA Separation and Control Hiring Assessment (SACHA) Competency Assessment continued (Adapted from Borman et al, 2001)

Coordinating - How effective is each controller at coordinating?		
Is often ineffective in receiving or initiating hand-offs (e.g., may often fail to contact controller in adjacent sector even when a hand-off is clearly required). When coordination is required, often fails to contact appropriate persons (e.g., pilot, other controllers, tower, etc.) or does so too slowly, sometimes causing traffic problems, delays, or worse.	Is generally good at hand-offs and pointouts, but may be somewhat slow in using hand-off line when very busy. When the situation calls for coordination, usually contacts all appropriate persons and coordinates properly with others.	Always coordinates hand-offs and pointouts appropriately, both initiating and receiving them very effectively and efficiently, even when very busy. Even in a tight time frame or difficult circumstances, always contacts and works with other controllers and pilots, as appropriate; effectively and efficiently coordinates to correct and avoid traffic problems or to reduce confusion and workload.
Teamwork - How effective is each controller in the area of teamwork?		
Ignores traffic flow in adjacent sectors and the impact own traffic flow may have on co-workers; avoids pitching in to help fellow controllers, even in high load situations such as loss of radar or poor weather conditions. Often waits until the last minute to take hand-offs; frequently dumps air traffic in adjacent sectors so as to reduce own workload; rarely volunteers to take on additional responsibility to help co-workers. Becomes extremely defensive, even belligerent, if constructive feedback is offered by supervisors or co-workers; may belittle co-workers, sometimes in front of others; rarely works well with others.	Is usually willing to assist co-workers who become extremely busy (e.g., by assuming hand-off and coordination duties). Is generally considerate of coworkers; adjusts own traffic flow to ease workload of adjacent sector when there are obvious problems. For the most part accepts constructive criticism from supervisors or coworkers; is usually able to refrain from criticizing other ATCSs; generally works well with other controllers.	Is always alert to traffic in other sectors and pitches in to help coworkers (e.g., by accepting additional airspace or assuming hand-off and coordination duties). Is always considerate of co-workers, working to ensure smooth and timely traffic flow between adjacent sectors; whenever possible, adjusts own traffic flow to ease workload of next sector (e.g., when traffic in adjacent sectors becomes heavy). Is always open to feedback from supervisors or co-workers, accepting criticism in a positive, constructive, and professional manner; never belittles co-workers; always works harmoniously with other controllers.
Overall Effectiveness - The scales you have just made ratings on represent 10 different areas important for air traffic controller effectiveness. This scale asks you to rate the *overall effectiveness* of each controller, taking into account behavior related to all 10 of the previous categories		
Performs poorly in important effectiveness areas; does not meet standards and expectations for adequate controller performance.	Adequately performs in important effectiveness areas; meets standards and expectations for adequate controller performance.	Performs excellently in all or almost all effectiveness areas; exceeds standards and expectations for controller performance.

Table B29 - FAA Separation and Control Hiring Assessment (SACHA) Competency Assessment continued (Adapted from Borman et al, 2001)

B14 - FAA MODIFIED SACHA

Maintaining Safe and Efficient Traffic Flow	
Maintaining Separation and Resolving Potential Conflicts: • Using control instructions that maintain safe aircraft separation • Detecting and resolving impending conflicts early	1, 2, 3, 4, 5, 6, 7, 8, NA
Sequencing Arrival and Departure Aircraft Efficiently: • Using efficient and orderly spacing techniques for arrival and departure aircraft • Maintaining safe arrival and departure intervals that minimize delays	1, 2, 3, 4, 5, 6, 7, 8, NA
Using control instructions effectively: • Providing accurate navigational assistance to pilots • Avoiding clearances that result in the need for additional instructions to handle aircraft completely • Avoiding excessive vectoring or over-controlling	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall safe and efficient traffic flow scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA
Maintaining Attention and Situation Awareness	
Maintaining Awareness of Aircraft Positions: • Avoiding fixation on one area of the radar scope when other areas need attention • Using scanning patterns that monitor all aircraft on the radar scope	1, 2, 3, 4, 5, 6, 7, 8, NA
Ensuring Positive Control:	1, 2, 3, 4, 5, 6, 7, 8, NA
Detecting pilot deviations from control instructions • Ensuring that pilot follows assigned clearance correctly • Correcting pilot deviations in a timely manner	1, 2, 3, 4, 5, 6, 7, 8, NA
Correcting own errors in a timely manner	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall attention and situation awareness scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA
Prioritising	
Taking actions in an appropriate order of importance: • Resolving situations that need immediate attention before handling low priority tasks	1, 2, 3, 4, 5, 6, 7, 8, NA
Preplanning control actions: • Scanning adjacent sectors to plan for inbound traffic • Study pending flight strips in bay	1, 2, 3, 4, 5, 6, 7, 8, NA
Handling control tasks for several aircraft: • Shifting control task between several aircraft when necessary • Avoiding delays in communications while thinking or planning control actions	1, 2, 3, 4, 5, 6, 7, 8, NA
Marking flight strips while performing other tasks: • Marking flight strips accurately while taking or performing other tasks • Keeping flight strips current	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall prioritising scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA
Providing control information	
Providing essential air traffic control information: • Providing mandatory services and advisories to pilots in a timely manner exchanging essential information	1, 2, 3, 4, 5, 6, 7, 8, NA
Providing additional air traffic control information: • Providing additional services when workload is not a factor exchanging additional information	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall providing control information scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA
Technical Knowledge	
Showing knowledge of LOAs and SOPs: • controlling traffic as depicted in current LOAs and SOPs • performing handoff procedures correctly	1, 2, 3, 4, 5, 6, 7, 8, NA
Showing knowledge of aircraft capabilities and limitations: • Avoiding clearances that are beyond aircraft performance parameters • Recognizing the need for speed restrictions and wake turbulence separation	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall technical knowledge scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA
Communicating	
Using proper phraseology: • Using words and phrases specified in ATP7110.65 • Using ATP phraseology that is appropriate for the situation • Avoiding the use of excessive verbiage	1, 2, 3, 4, 5, 6, 7, 8, NA
communicating clearly and efficiently: • speaking at the proper volume and rate for pilots to understand • speaking fluently while scanning or performing other tasks • clearance delivery is complete, correct and timely • providing complete information in each clearance	1, 2, 3, 4, 5, 6, 7, 8, NA
Listening to pilot read backs and requests: • Correcting pilot read back errors • Acknowledging pilot or other controller requests promptly • Processing requests correctly in a timely manner	1, 2, 3, 4, 5, 6, 7, 8, NA
Overall communicating scale rating	1, 2, 3, 4, 5, 6, 7, 8, NA

Table B30 - FAA (modified) Separation and Control Hiring Assessment (SACHA) Competency Assessment (Adapted from Sollenberger et al, 1997)

B15 - FAA AT-SAT OTS

AT-SAT High Fidelity Simulation Over The Shoulder (OTS) Rating Form

Administrative Information - Page 1

Scenario Number: HFG 1 2 3 4 5 6 7		Lab Number: 1 2
Position: 1 2 3 4 5 6 7 8 9 10	Participant ID Number:	Rater ID Number:

AT-SAT High Fidelity Simulation Over The Shoulder (OTS) Rating Scales

Rating Dimensions		Rating Scale: Below Average	Rating Scale: Fully Adequate	Rating Scale: Exceptional
A. Maintaining Separation		① ②	③ ④ ⑤	⑥ ⑦
• Checks separation and evaluates traffic movement to ensure separation standards are maintained	• Considers aircraft performance parameters when issuing clearances			
• Detects and resolves impending conflictions	• Establishes and maintains proper aircraft identification			
• Applies appropriate speed and altitude restrictions	• Properly uses separation procedures to ensure safety			
• Analyzes pilot requests, plans and issues clearances	• Issues safety and traffic alerts			
B. Maintaining Efficient Air Traffic Flow		① ②	③ ④ ⑤	⑥ ⑦
• Accurately predicts sector traffic overload and takes appropriate action	• When necessary, issues a new clearance to expedite traffic flow			
• Ensure clearances require minimum flight path changes	• Reacts to/resolves potential conflictions efficiently			
• Controls traffic so as to ensure efficient and timely traffic flow				
C. Maintaining Attention and Situation Awareness		① ②	③ ④ ⑤	⑥ ⑦
• Maintains awareness of total traffic situation	• Reviews and ensures appropriate route of flight			
• Recognizes and responds to pilot deviations from ATC clearances	• Scans properly for air traffic events, situations, potential problems, etc.			
• Listens to readbacks and ensures they are accurate	• Remembers, keeps track of, locates, and if necessary orients aircraft			
• Assigns requested altitude in timely manner	• Descends arrivals in timely manner			
• Keeps data blocks separated	• Accepts/performs timely handoffs			
D. Communicating Clearly, Accurately, and Efficiently		① ②	③ ④ ⑤	⑥ ⑦
• Issues clearances that are complete, correct, and timely	• Communicates clearly and concisely			
• Makes only necessary transmissions	• Uses correct call signs			
• Uses standard/prescribed phraseology	• Uses appropriate speech rate			
• Properly establishes, maintains, and terminates communications	• Listens carefully to pilots and controllers			
• Avoids lengthy clearances	• Issues appropriate arrival and departure information			

Figure B5 - FAA AT-SAT Over The Shoulder (OTS) Observation Form (Manning et al, 2001)

E. Coordinating	① ②	③ ④ ⑤	⑥ ⑦

- Performs handoff and pointout procedures correctly
- Effectively coordinates clearances, changes in aircraft destinations, altitudes, etc.
- Provides complete/accurate position relief briefings
- Performs required coordinations effectively
- Initiates and receives handoffs and pointouts in an efficient and effective manner
- Processes flight plans/amendments as required

F. Performing Multiple Tasks	① ②	③ ④ ⑤	⑥ ⑦

- Shifts attention between several aircraft when necessary
- Keeps track of a large number of aircraft/events at a time
- Prioritizes activities effectively
- Communicates in a timely fashion while performing other actions
- Returns to what he/she was doing after an interruption

G. Managing Sector Workload	① ②	③ ④ ⑤	⑥ ⑦

- Handles heavy, emergency, and unusual traffic situations effectively
- Stays calm, focused, and functional in busy and stressful conditions
- Responds to imposed airspace restrictions
- Responds to traffic management constraints/initiatives
- Handles unexpected situations effectively (e.g., computer/communication failures)
- Deals effectively with situations for which there may not be clearly prescribed procedures
- Uses contingency or "fall-back" strategies effectively

H. Overall Performance	① ②	③ ④ ⑤	⑥ ⑦

Figure B6 - FAA AT-SAT Over The Shoulder (OTS) Observation Form continued (Manning et al, 2001)

APPENDIX C – OBSERVATION SHEETS

C01 - PRELIMINARY STUDY OBSERVATION SHEET 1

Start of Observation:		End of Observation:		No. observed
Sector:		Did a bandbox or split occur		
Position (T, P, A)		M/F		

HPF	Behavioural Marker/Indicator	No. observed
Situation Awareness, Attention , Focus	Focus & Concentration 4 Fast response to issues requiring action, decisive, clarifies situation, deals well with uncertainty, posture alert and attentive, not easily distracted, attention to detail	
	Strip scan 4 Checks through strips 4 runs finger/hand/pen down strips, cocks strips	
	Workstation scan 4 scan of main workstation components; strips, radar, SIS	
	Returns to a previously interrupted task	
	Where applicable, Electronic Decision Support tools are used (e.g. vector lines)	
	Focus & Concentration 4 Preoccupied, distracted, fixated with specific/minor tasks, daydreaming/automaton , on autopilot/switched off (black holeing), hunched/tunnelled into workstation	
	Awareness 4 appears confused, unable to concentrate, struggling to find aspects of system	
	Very low amount of RT for traffic situation	
User State: Stress Management, Fatigue, and Comfort Frustration, Morale, Motivation	Stays Calm & Relaxed 4 unemotional, not hunched into workstation, no verbal frustration	
	Adapts to pace of task 4 When quiet: sits back/legs crossed, reads paper/book etc, not wearing headset (due to low workload)	
	Positive & responsible attitude: constructive, supportive, approachable, enthusiastic, polite, friendly, relaxed	
	Poor concentration 4 , fidgety, restless, distracted, fiddling with systems/stuff, fidgety	
	Angry / Stressed 4 snappy, grumpy, irritated, excessive swearing, verbal frustration, Looking flustered, rosy cheeks	
	Physically Stressed 4 overly leaning into workstation, tense body posture	
	Uncomfortable / Fatigued 4 rubbing eyes/face, stretching, yawning, signs of discomfort when moving	
	Care and consideration 4 Displays a lack of care and respect for equipment (even vandalism)	
Decision Making, Planning, and Workload Management	Maintains strips accuracy (updated by pen or keyboard) 4 Clearances, level changes etc	
	Picks up the pace 4 as traffic level/complexity increases	
	Keep on top of RT loading 4 doesn't miss calls, or ask a/c to 'standby'	
	Demonstrates multi-tasking and divided attention	
	Seeks assistance when workload increases	
	Task Rate/Response Rate 4 Inferior, delayed, or no response to actions and requests, Excessive stalling tactics/hesitation and task dropping, fixated on one task, over focus on easy tasks, Does not keep on top of RT loading, misses calls, asks a/c to standby	
	Excessive/inappropriate levels of help and requests for assistance	
Teamwork, Teambuilding and Team Support	Team member attitude 4 Shows appreciation, happy to receive help, gives positive/constructive feedback, enthusiastic, easily approachable	
	Team problem solving 4 Helps others, alert each other to points of interest, problems, erroneous info, acknowledges prompts from others	
	As a Member of team 4 Overly competitive, patronising, negative, dominant, insular, unsociable	
	Team attitude 4 isolated team members, poor team mood	
Communications	Clarity of Comms 4 clear, concise, timely, not rushed, and authoritative, standard/correct phraseology used	
	Defers calls 4 Phone calls are deferred, RT parties told to stand by when responding to other information/tasks.	
	Quality & Accuracy of Comms 4 Incomplete, or incorrect read backs are not accepted, unclear/ambiguous messages are questioned, non-standard phraseology is	
	Verbal/Non-verbal communications 4 unclear, too long, poorly timed, too quick, to quiet/weakly conveyed	
	More than two instructions are given in the same transmission.	
When Aircraft Calls	Quickly Finds Flight Data when aircraft calls	
	Consults FPS during transmission or reception of information.	
	Amends or annotates FPS during transmission.	

Figure C1 - Preliminary Study Observation Sheet 1

C02 - PRELIMINARY STUDY OBSERVATION SHEET 2

Start of Observation:		End of Observation:	
Sector:		Did a bandbox or split occur	
Position (T, P, A)		M/F	

HPF	Behavioural Marker/Indicator	No. observed	Detail
	Scanning and checking through Strips		Checks through strips 4 runs finger/hand/pen down strips, cocks strips
	Uses Electronic Decision Support tools		e.g. vector lines
	Strips used as memory aids:		Cocks strips, places above placeholders
	Declutters/reconfigures display		moves around TDBs and other display material to aid view ability
	Very low amount of RT for traffic*		
	Adapts to pace of task		When quiet: sits back/legs crossed, reads paper/book etc, not wearing headset (due to low workload)
	Positive & responsible		Attitude: constructive, supportive, approachable, enthusiastic, polite, friendly, relaxed
	Angry / Stressed*		Snappy, grumpy, irritated, excessive swearing, verbal frustration, Looking flustered, rosy
	Poor concentration / Fatigued*		Restless, distracted, fiddling/fidgety, rubbing eyes/face, stretching, yawning, signs of discomfort when moving
	Care and consideration*		Displays a lack of care and respect for equipment (even vandalism)
	Task Rate/Response Rate*		Inferior/delayed,/no response to actions/requests/RT. Excessive stalling, fixated on one task/ easy tasks
	Excessive/inappropriate levels of help needed*		Continued requests for assistance
Teamwork, building & Support	Team member attitude & Support		Shows appreciation, gives & receives help, positive & enthusiastic
	As a Member of team*		Overly competitive, patronising, negative, dominant, insular, unsociable
	Clear Comms		clear, concise, timely, not rushed, and authoritative, standard/correct phraseology used
	Defers calls		Phone calls are deferred, when engaged in RT
	Strives for Quality Comms		corrects read backs, repeats if required, unclear/ambiguous/non standard messages questioned
	Verbal/Non-verbal communications*		unclear, too long, poorly timed, too quick, to quiet/weakly conveyed
	Sorts/orders strips		
	Maintains strips		(updated by pen or keyboard) 4 Clearances, level changes etc
	Bins Strips		Disposes of strip/holder when finished
	Quickly Finds Flight Data		when aircraft calls
	Consults/Annotates FPS		During transmission or reception of information.
	Amends FPS after transmission*		
Quiet Indications	Relaxed posture		Sits back from workstation, arms on arm rest, twisted/swivelled away from radar
	Asks LAS for Bandbox		
	Comms embellished from minimum		Chatty with pilot, offers pilot option of levels/routes
Busy Indications	Seeks assistance when workload increases		Requests split of sectors, flow control

Figure C2 - Preliminary Study Observation Sheet 2 ('* denotes negative marker)

C03 - PRELIMINARY STUDY OBSERVATION SHEET 1 & 2 DIFFERENCES

Preliminary Study (First Behavioural Marker Set)		Preliminary Study (Second Behavioural Marker Set)	
Situation Awareness, Attention , Focus	Focus & Concentration 4 Fast response to issues requiring action, decisive, clarifies situation, deals well with uncertainty, posture alert and attentive, not easily distracted, attention to detail		Situation Awareness, Attention , Focus
	Strip scan 4 Checks through strips 4 runs finger/hand/pen down strips, cocks strips	Scanning and checking through Strips Checks through strips 4 runs finger/hand/pen down strips, cocks strips	
	Workstation scan 4 scan of main workstation components; strips, radar, SIS		
	Returns to a previously interrupted task	Strips used as memory aids: Cocks strips, places above placeholders	
	Where applicable, Electronic Decision Support tools are used (e.g. vector lines)	Uses Electronic Decision Support tools e.g. vector lines	
	Focus & Concentration 4 Preoccupied, distracted, fixated with specific/minor tasks, daydreaming/automaton , on autopilot/switched off (black holeing), hunched/tunnelled into workstation		
	Awareness 4 appears confused, unable to concentrate, struggling to find aspects of system	Declutters/reconfigures display moves around TDBs and other display material to aid view ability	
	Very low amount of RT for traffic situation	Very low amount of RT for traffic	
User State: Stress Management, Fatigue, and Comfort Frustration, Morale, Motivation	Stays Calm & Relaxed 4 unemotional, not hunched into workstation, no verbal frustration		User State: Stress Management, Fatigue, and Comfort Frustration, Morale, Motivation
	Adapts to pace of task 4 When quiet: sits back/legs crossed, reads paper/book etc, not wearing headset (due to low workload)	Adapts to pace of task When quiet: sits back/legs crossed, reads paper/book etc, not wearing headset (due to low workload)	
	Positive & responsible attitude: constructive, supportive, approachable, enthusiastic, polite, friendly, relaxed	Positive & responsible Attitude: constructive, supportive, approachable, enthusiastic, polite, friendly, relaxed	
	Poor concentration 4 , fidgety, restless, distracted, fiddling with systems/stuff, fidgety	Poor concentration / Fatigued Restless, distracted, fiddling/fidgety, rubbing eyes/face, stretching, yawning, signs of discomfort when moving	
	Angry / Stressed 4 snappy, grumpy, irritated, excessive swearing, verbal frustration, Looking flustered, rosy cheeks	Angry / Stressed Snappy, grumpy, irritated, excessive swearing, verbal frustration, Looking flustered, rosy cheeks	
	Physically Stressed 4 overly leaning into workstation, tense body posture		
	Uncomfortable / Fatigued 4 rubbing eyes/face, stretching, yawning, signs of discomfort when moving		
	Care and consideration 4 Displays a lack of care and respect for equipment (even vandalism)	Care and consideration 4 Displays a lack of care and respect for equipment (even vandalism)	
		Task Rate/Response Rate Inferior/delayed,/no response to actions/requests/RT. Excessive stalling, fixated on one task/ easy tasks	
		Excessive/inappropriate levels of help needed Continued requests for assistance	

Table C1 - Preliminary Study Observation Sheets 1 & 2 Differences - Part 1

Preliminary Study (First Behavioural Marker Set)		Preliminary Study (Second Behavioural Marker Set)	
Decision Making, Planning, and Workload Management	Maintains strips accuracy (updated by pen or keyboard) 4 Clearances, level changes etc		
	Picks up the pace 4 as traffic level/complexity increases		
	Keep on top of RT loading 4 doesn't miss calls, or ask a/c to 'standby'		
	Demonstrates multi-tasking and divided attention		
	Seeks assistance when workload increases		
	Task Rate/Response Rate 4 Inferior, delayed, or no response to actions and requests, Excessive stalling tactics/hesitation and task dropping, fixated on one task, over focus on easy tasks, Does not keep on top of RT loading, misses calls, asks a/c to standby		
	Excessive/inappropriate levels of help and requests for assistance		
Teamwork, Teambuilding and Team Support	Team member attitude 4 Shows appreciation, happy to receive help, gives positive/constructive feedback, enthusiastic, easily approachable	Team member attitude & Support 4 Shows appreciation, gives & receives help, positive & enthusiastic	Teamwork, Teambuilding and Team Support
	Team problem solving 4 Helps others, alert each other to points of interest, problems, erroneous info, acknowledges prompts from others		
	As a Member of team 4 Overly competitive, patronising, negative, dominant, insular, unsociable	As a Member of team 4 Overly competitive, patronising, negative, dominant, insular, unsociable	
	4 Team attitude 4 isolated team members, poor team mood		
Communicatio ns	+ Clarity of Comms 4 clear, concise, timely, not rushed, and authoritative, standard/correct phraseology used	Clear Comms 4 clear, concise, timely, not rushed, and authoritative, standard/correct phraseology used	Communicatio ns
	+ Defers calls 4 Phone calls are deferred, RT parties told to stand by when responding to other information/tasks.	Defers calls Phone calls are deferred, when engaged in RT	
	4 Quality & Accuracy of Comms 4 Incomplete, or incorrect read backs are not accepted, unclear/ambiguous messages are questioned, non-standard phraseology is challenged	Strives for Quality Comms: corrects read backs, repeats if required, unclear/ambiguous/non standard messages questioned	
	4 Verbal/Non-verbal communications 4 unclear, too long, poorly timed, too quick, to quiet/weakly conveyed	Verbal/Non-verbal communications 4 unclear, too long, poorly timed, too quick, to quiet/weakly conveyed	
	4 More than two instructions are given in the same transmission.		

Table C2 - Preliminary Study Observation Sheets 1 & 2 Differences - Part 2

When Aircraft Calls	+ Quickly Finds Flight Data when aircraft calls + Consults FPS during transmission or reception of information. + Amends or annotates FPS during transmission.

Preliminary Study (Second Behavioural Marker Set)	
Quickly Finds Flight Data when aircraft calls	Management of Flight Strips/Flight Data
Consults/Annotates FPS: During transmission or reception of information	
Amends FPS after transmission	
Sorts/orders strips	
Maintains strips (updated by pen or keyboard) 4 Clearances, level changes etc	
Bins Strips: Disposes of strip/holder when finished	
Relaxed posture Sits back from workstation, arms on arm rest, twisted/swivelled away from radar	Quite Indications
Asks LAS for Bandbox	
Comms embellished from minimum Chatty with pilot, offers pilot option of levels/routes	
Seeks assistance when workload increases Requests split of sectors, flow control	Busy Indications

Preliminary Study (First Behavioural Marker Set)

Table C3 - Preliminary Study Observation Sheets 1 & 2 Differences - Part 3

C04 - MAIN OBSERVATION SHEET VERSION 1

Date & Time	
Observe	
Length of Observation	
Other	

Undertaking the task

Best practice ______

Interface Scan _____

Attitude & Mood:

Negative comments ______________

Positive comments ______________

Apologetic ____________________

Social ________________________

Frustrated ____________________

Communications & Verbal Commentary:

Tongue-tied ___________________

Self Affirm ___________________

Delays ________________________

Cool Calm~~~~~~~~~~~~~~~~~~~~~~

Irritated _____________________

Confusion _____________________

Decisive ______________________

Physical Posture & Body Language:

Pace Fast _____________________

Pace Slow _____________________

Phys. Frustration ______________

Fatigue _______________________

Adjust MMI ____________________

Fidgets **Inputs and Interaction with HMI and**

orkstation:

Muscle memory __________ ________

Overconfident ___________ ________

Slow Hesitant~~~~~~~~~~~~

Play Sandpit~~~~~~~~~~~~~

Dual Tasking ____________ ________

Confident control _______

Incorrect actions _______

Surprise _______________ ________

Quickly locates _________ ________

Automatic ______________ ________

Periphery ______________ ________

Plans Ahead ____________ ________

Keeps Info open _________ ________

Input device tick off __

Cyclic Scan ____________ ________

Quirks _________________ ________

Interaction with others:

Affirm Before ____

Affirm After ____

Reactionary _____

Team Aware ______

Team Contribute __

Team Short Snappy

Figure C3 - Main Observation Sheet - Version 1

C05 - MAIN OBSERVATION SHEET VERSION 2

Location	
Date	
Start Time/Stop Time	
Person Observed	
Length of Observation	
Other	

Attitude & Mood:

Negative comments ______________

Positive comments ______________

Apologetic____________________

Social__________________________

Communications & Verbal Commentary:

Verbal Frustration_____________

Confusion______________________

Tongue-tied____________________

Self Affirm____________________

Delays__________________________

Cool Calm~~~~~~~~~~~~~~~~~~~~~~

Decisive_______________________

Standbys_______________________

Physical Posture & Body Language:

Pace Fast______________________

Pace Slow______________________

Phys. Frustration _____________

Fatigue________________________

Adjust MMI_____________________

Fidgets **Inputs and Interaction with HMI and**

orkstation:

Muscle memory___________________

Overconfident___________________

Slow Hesitant~~~~~~~~~~~

Play Sandpit~~~~~~~~~~~~

Dual Tasking____________________

Serial Tasking__________________

Confident control_______

Incorrect actions_______

Surprise_______________ ________

Quickly locates_________________

Automatic______________ ________

Periphery______________ ________

Plans Ahead____________ ________

Keeps Info open_________________

Input device tick off___

Cyclic Scan____________ ________

Quirks_________________ ________

Interaction with others:

Affirm Before_______

Affirm After_______

Reactionary________

Team Aware_________

Team Contribute_____

Team Short Snappy~~~

General Queries_____

Technical Discussion

Figure C4 - Main Observation Sheet - Version 2

C06 - MAIN OBSERVATION SHEET VERSION 3

Location/Unit		Date	
Start Time		Person Observed (name or anonymous)	
Stop Time		Length of Observation	
Other Factors:			

Attitude & Mood:

Negative comments______________

Positive comments______________

Apologetic ____________________

Social ________________________

Communications & Verbal Commentary:

Verbal Frustration ____________

Confusion _____________________

Tongue-tied ___________________

Self Affirm ___________________

Delays ________________________

Cool Calm~~~~~~~~~~~~~~~~~~~~~~

Indecisive Hesitant~~~~~~~~~~~~

S t a n d b y s

Physical Posture & Body Language:

Pace Fast _____________________

Pace Slow _____________________

Phys. Frustration______________

Fatigue _______________________

Adjust MMI ____________________

Fidgets________________________

Inputs and Interaction with HMI and orkstation:

Overconfident__________ ________

Slow Hesitant~~~~~~~~~~~~

Play Sandpit~~~~~~~~~~~~~

Dual Tasking___________ ________

Serial Tasking_________ ________

Incorrect actions______

Surprise ______________ ________

Automatic Memory _______________

Periphery _____________ ________

Plans Ahead ___________ ________

Cyclic Scan ___________ ________

Quirks ________________ ________

Interaction with others:

Affirm Before_______

Affirm After ______

Reactionary _______

Team Aware ________

Team Contribute_____

Team Short Snappy~~~

General Queries_____

Technical Discussion

____

____

____

____

Figure C5 - Main Observation Sheet - Version 3

C07 - MAIN OBSERVATION SHEET VERSION 4

User Behaviour & System Interaction Observation Sheet

Location/Unit		Date	
Start Time		Person Observed (name or anonymous)	
Stop Time		Length of Observation	
Other Factors:			

Attitude & Mood:

Negative comments______________

Positive comments______________

Apologetic ____________________

Social ________________________

Communications & Verbal Commentary:

Verbal Frustration ____________

Confusion _____________________

Tongue-tied ___________________

Self Affirm ___________________

Delays ________________________

Cool Calm~~~~~~~~~~~~~~~~~~~~~~

Indecisive Hesitant~~~~~~~~~~~~

Standbys ______________________

Physical Posture & Body Language:

Pace Fast _____________________

Pace Slow _____________________

Phys. Frustration______________

Fatigue _______________________

Adjust MMI ____________________

Fidgets________________________

Inputs and Interaction with HMI and orkstation:

Overconfident __________ _______

Slow Hesitant~~~~~~~~~~________

Play Sandpit~~~~~~~~~~~~~______

Dual Tasking ___________ _______

Serial Tasking _________ _______

Incorrect actions ______

Surprise _______________________

Automatic Memory ______ ________

Periphery ______________________

Plans Ahead ____________________

Cyclic Scan ____________________

Quirks _________________________

Interaction with others:

Affirm Before ______

Affirm After ______

Reactionary _______

Team Aware ________

Team Contribute ____

Team Short Snappy~~~

General Queries ____

Technical Discussion

Low Dexterity ______

Clumsy _____________ ___________

Frantic approach ___

Leans over

Nervous Voice ______

Nervous Physical ___

Spatial reference __

____

____

Figure C6 - Main Observation Sheet - Version 4

C08 - GUIDANCE SHEET (FOR MARKER SHEET V.4)

Behaviour Reference Sheet (observation sheet version 4)

Attitude & Mood:

Negative comments	Comments regarding the task/interface/HMI	Negative comments
Positive comments		Positive Comments
Apologetic	Critical of own performance	Apologetic of own performance
Social	Jovial laughing, Relaxed and chatty	Talks socially in quieter periods

Communications & Verbal Commentary:

Confusion	Confusion & Lack of understanding	(Verbal queries 4 why, what, where, when etc)
Tongue-tied	Clarity of Comms	Gets tongue tied in RT comms
Self Affirm	Self affirmation of doing things right/right function	Nods head, ok I understand, I'm getting there
Delays/Repeats	Pauses and Delays	Er's um's, oh, um, alright um, pauses, delays, "standby", "say again, I missed that"
Standbys		Uses 'standbys' as a delaying tactic.
Frustrated/ Irritated	Tone of Voice	frustrated, angry, irritated, edgy
Indecisive		confused, not sure if doing things right

Interaction with others:

Affirm Before	awaits affirmation	Before an action
Affirm After		After an action
General Queries/Questions		Asking broad questions, rather than specific technical queries regarding the operation of the system
Reactionary	Reacts to instructor/other's prompts-	Oh right, yes, I see, ah ok
Team Aware	Responds to comments from other team members when engaged in their own primary task	Displays the ability to undertake dichotic listening
Team Contribute	Works Integrally with team	Offers suggestions and discusses options with team members
Team Short/Snappy	Snappy/short with surrounding staff	When busy, and when been put under pressure
Tech. Discussion		

Interaction with HMI and workstation:

Muscle memory	Movement and control of HMI	Demonstrates spatial/muscle memory of interface and layout functions
Overconfident		Overconfident, fast, but makes mistakes and select wrong functions, performs unnecessary tasks
Slow/Hesitant		Slow and Hesitant, indecisive, unsure of actions, moving to control a function then moving back, requires multiple attempts to drive the
Play/Sandpit		Plays and experiments with system to see how it responds and behaves, and to rehearse actions
Dual Tasking		Dual Tasking/Multi Tasking 4 whilst engaged in RT will also drive HMI and input data
Confident control		Confident smooth flowing control. Deft command of HMI
Incorrect actions	Knowledge of HMI	Using HMI incorrectly 4 wrong clicks, taps but no action on interface, can't find right function etc
Surprise		Surprised by behaviour of HMI
Quickly locates		Quickly located required functionality and information when required
Automatic		Displays 'automaton' type actions when progressing through a sequence of steps

Figure C7 - Observer Guidance Sheet - Part 1 (For Marker Sheet Version 4)

Interaction with HMI and workstation (Continued):

Periphery	Scanning & Planning	Picks up activities (e.g. changes in HMI) on the periphery
Plans Ahead		Plans task ahead 4 Opens up windows in advance, leaves cursor in the position needed for the next action or an action that they need to return back to, highlights all strips requiring QNH update prior to departure
Keeps Info open		keeps information windows active and open with applicable info during read back
Input device tick off		Hovers pen/cursor over info on read back to confirm/tick off
Cyclic Scan	Maintains cyclic work pattern	of looking out of the tower, checking arrivals and departures on radar, checking surface radar, checking the EFPS display(s), and other ancillary displays as appropriate (e.g. lighting panels, ATIS)
Quirks	Demonstrates idiosyncrasies & quirky mannerisms	when driving the HMI (e.g. tapping across taxi way windows in a 3, 2, 1, motion, moves the cursor round in circles especially on-screen objects, adjusts the windows so they line up absolutely perfectly

Physical Posture & Body Lanauaae:

Pace Fast	Adapts to pace of task	Relaxes when quiet, sits back, crosses arms
Pace Slow		Sitting up alert and attentive
Phys. Frustration	Physical frustration	gestures of waving hands about, blowing air out
Fatigue	Fatigue	rubbing face, yawns, rubs eyes, "I'm tired", looks at watch
Adjust MMI	Adjusts workstation/HMI	
Fidgets	Fidgets	Taps pen, wriggles about, taps leg

How to undertake the Observations:

Use the observation sheet in order to record each instance a specific listed behaviour is exhibited by the person observed; with a maximum limit of ten observations against each individual behaviour. Use the double sided reference sheet provided if uncertain as to a behaviours definition. The reference sheet contains detail on each individual marker.

You are encouraged to note down any comments, queries or other observations surrounding the markers themselves. Please consider the merits of each marker, and their individual utility. Please also note down any situational elements (environment, task etc) that may have had an influence on events and any addition behaviours they may have spotted which provide insight as to the relationship and level of user development they have achieved with the ATC system to which they are using.

Figure C8 - Observer Guidance Sheet - Part 2 (For Marker Sheet Version 4)

C09 - TIMELINE OF BEHAVIOURAL MARKERS

Category of Markers	Observation sheet version			
	Version 1 Chapter 6 Observations 1-26	**Version 2** Chapter 6 Observations 26-52	**Version 3** Changes made following Chapter 7 Inter rater review	**Version 4** Changes made following Chapter 8
Undertaking the task	Best practice	Removed		
	Interface Scan	Combined with cyclic scan		
Attitude & Mood	Negative comments			
	Positive comments			
	Apologetic			
	Social			
	Frustrated	Verbal Frustration: (Frustrated, Irritated)		
Comms & Verbal Commentary	Irritated			
	Confusion			
	Tongue-tied			
	Self Affirm			
	Delays (Delays/Repeats)			
		Standbys		
	Cool/Calm		Removed	
	Decisive		Changed to: Indecisive/hesitant	
Physical Posture & Body Language	Pace Fast			
	Pace Slow			
	Phys. Frustration			
	Fatigue			
	Adjust MMI			
	Fidgets			
Inputs and Interaction with HMI and Workstation	Surprise			
	Overconfident			
	Slow/Hesitant			
	Play/Sandpit			
	Dual Tasking			
		Serial Tasking		
	Periphery			
	Cyclic Scan/Maintains Global SA			
	Quirks			
	Incorrect actions			
	Confident control		Automatic Memory: (Muscle Memory, Automatic, Quickly Locates, Confident control)	
	Muscle memory			
	Quickly locates			
	Automatic			
	Plans Ahead		Plans Ahead: (Keeps Info Open, Plans Ahead)	
	Keeps Info open			
	Input device tick off		Removed	
				Low Dexterity
				Clumsy
				Frantic approach
				Leans over
				Nervous Voice
				Nervous Physical
				Spatial reference
Interaction with others	Affirm Before			
	Affirm After			
		General Queries		
		Tech Discussion		
	Reactionary			
	Team Aware			
	Team Contribute			
	Team Short/Snappy			

Table C4 - Behavioural Marker & Observation Sheet Version History

APPENDIX D – DATA ANALYSIS & GENERAL MATERIAL

D01 - PRELIMINARY OBSERVATIONAL STUDY DATA

Category	Behavioural Markers	total	Mean/Hr.
Situation Awareness, Attention , Focus	+ Focus & Concentration	18	2.87
	+ Strip scan	100	16.64
	+ Workstation scan	6	1.03
	+ Returns to a previously interrupted task	6	0.99
	+ Where applicable, Electronic Decision Support tools are used	105	17.94
	4 Focus & Concentration	1	0.12
	4 Awareness	0	0
	4 Very low amount of RT for traffic situation	0	0
User State: Stress Management, Fatigue, and Comfort, Frustration, Morale, Motivation	+ Stays Calm & Relaxed	21	4.09
	+ Adapts to pace of task	34	5.96
	+ Positive & responsible attitude	13	2.20
	4 Poor concentration	0	0
	4 Angry / Stressed	0	0
	4 Physically Stressed	0	0
	4 Uncomfortable / Fatigued	1	0.26
	4 Care and consideration	0	0
Decision Making, Planning, and Workload Management	+ Maintains strips accuracy	443	76.09
	+ Picks up the pace 4 as traffic level/complexity increases	10	1.91
	+ Keep on top of RT loading 4 doesn't miss calls, or ask a/c to 'standby'	1	0.18
	+ Demonstrates multi-tasking and divided attention	8	1.75
	+ Seeks assistance when workload increases	0	0
	4 Task Rate/Response Rate	0	0
	4 Excessive/inappropriate levels of help	0	0
Teamwork, Teambuilding and Team Support	+ Team member attitude	51	9.49
	+ Team problem solving	7	1.18
	4 Overly competitive, patronising, negative, dominant, insular, unsociable	0	0
	4 Team attitude	0	0
Communications	+ Clarity of Comms	495	84.79
	+ Defers calls	4	0.43
	4 Quality & Accuracy of Comms	0	0
	4 Verbal/Non-verbal communication	5	0.55
	4 More than two instructions are given in the same transmission.	0	0
When Aircraft Calls	+ Quickly Finds Flight Data when aircraft calls	121	19.67
	+ Consults FPS during transmission or reception of information.	444	76.41
	+ Amends or annotates FPS during transmission.	437	74.98

Table D1 - Preliminary Study Data using Observation Sheet Version 1

Category	Behavioural Markers	total	Mean/Hr.
Situation Awareness, Attention , Focus	Scanning and checking through Strips	78	8.30
	Uses Electronic Decision Support tools	105	17.60
	Strips used as memory aids	16	3.24
	Declutters/reconfigures display	100	19.53
User State: Stress Management, Fatigue, and Comfort Frustration, Morale, Motivation	Adapts to pace of task	11	2.18
	Positive & responsible	11	3.05
	Angry / Stressed	1	0.11
	Poor concentration / Fatigued	0	0.00
	Care and consideration	0	0.00
Decision Making Task/Workload Management	Keep on top of RT loading	1	0.20
	Task Rate/Response Rate	0	0.00
	Excessive/inappropriate levels of help needed	0	0.00
Teamwork, Support	Team member attitude & Support	44	8.13
	As a Member of team	0	0.00
Communications	Clear Comms 4	447	74.15
	Defers calls	0	0.00
	Strives for Quality Comms:	23	4.93
	Verbal/Non-verbal communications	0	0.00
Management of Flight Strips/Flight Data	Upon receipt Sorts/orders strips	115	18.03
	Maintains strips	83	12.51
	Bins Strips:	147	24.45
	Quickly Finds Flight Data	130	21.96
	Consults/Annotates FPS:	353	59.75
Quite Indications	Relaxed posture	7	2.04
	Asks LAS for Bandbox	0	0.00
	Comms embellished from minimum	1	0.23
Busy Indications	Seeks assistance when workload increases	0	0.00

Table D2 - Preliminary Study Data using Observation Sheet Version 2

Category	Behaviour	Potential indicator of?
Physical posture and body language	Finger tapping	sign of thinking/decision making, or simply boredom/habit
	Adjusts chair/keyboard/mouse/HMI to suit	workload indicator?
	Sitting twisted/seat twisted away from radar	indication that complexity/workload is not high.
	Arms on arm rest	this could be added to the leaning back/looking relaxed marker
	Physically sits back/further away from workstation	sign of low workload / user comfort/satisfaction with task/environment
	Very radar focused	same thing as 'black holing'
Strip handling and management	Inching up strips with both hands (working through strips)	Scanning and Situation Awareness
	Playing with strips	Comfortable and highly familiar with paper strips
	Orders/sorts strips when received by ATSA	Planning
	General strip sorting & strip maintenance	not currently listed?
	Cocks strips	This is done as an aide-memoir
	Annotates strips after RT	Does this go against best practice
	Binning strip holders and stacking paper strips	very tactile and auditory feedback to the controller
	Takes RT first then amends strips	Don't think this is the best practice way of doing it.
	Moves around TDBs and other display material to aid view-ability	
Communications	Repeats communications when required, when a/c do not respond	Alert and aware a/c have not replied
	Offers a/c choice to choose level/route	Offering a tailored service to a/c; and indication of low to medium traffic
	Corrects read back errors	Best practice, alert and attentive
	Chatty with pilot	Low workload
Team	Chats socially during quite periods	There are comfortable levels of workload
	General good team atmosphere (laughter, jokes, chatting)	Positive team atmosphere
	Needing info from team members as distracted/unfocused/pre-occupied	Strong team support
Environmental context	Asks supervisor for bandbox	indication of low traffic levels
	Bandbox occurs during observation	indication of low traffic levels
	Removal of bandbox during observation	indication of high traffic levels

Table D3 - Preliminary Study Data Potential new behavioural markers identified

D02 - BEHAVIOURAL MARKERS OF LEARNING & DEVELOPMENT IDENTIFIED

Category	Behaviour observed	ATCO Group 1	ATCO Group 2	ATCO Group 3	What is the value of this marker?
	Demonstrates best practice (e.g. corrects incorrect Radio Telephony performs handover checks etc).			x	Indicates that the ATCO is in command of the situation and system, and able to use best proactive methods in their work
	Performs regular scans of the interface(s), searching and checking information	x	x	x	Regular scanning of displays is a vital component in the formation and maintenance of effective situation awareness
	Negative comments	x	x		Suggests uncertainty from the ATCO of the new system, and a lack of system experience. There is also an element of resistant to change.
	Positive Comments	x	x		Suggests uncertainty from the ATCO of the new system, and a lack of system experience. There is also an element of resistant to change.
	Apologetic of own performance	x			The ATCOs know their performance is not as good as with their current system, and want to perform as well with the new system.
	Talks socially in quieter periods		x	x	Indicates a certain level of ATCO spare capacity
	Laughs with frustration	x	x		Indicates that an ATCO is frustrated with their current level of proficiency and understanding with the system (and that further progress is needed before proficiency is obtained)
	(Verbal queries - why, what, where, when etc)	x	x		Demonstrates confusion and uncertainty as to the results of system behaviour, suggesting knowledge and understanding is still developing.
	Gets tongue tied in RT comms	x			Distracted by the task, and task interface, showing signs of confusion (like a secondary task measure of workload).
	Nods head, ok I understand, I'm getting there, talks to themselves, talks through next steps in task	x			Undertakes verbal narrative, and verbal think out loud
	Er's um's, oh, um, alright um, pauses, delays, "standby", "say again, I missed that"	x	x	x	Slowed responses and pauses suggest system navigation skills and interaction skills still developing.
	relaxed, calm	x	x	x	Comfortable with the task, and the system, and the roles and responsibilities required
	frustrated, angry, irritated, edgy	x	x		Shows frustration and irritation at themselves for their level of performance, wanting to achieve better.
	Confident, Decisive, self assured	x	x	x	Actions are smooth, efficient, and decisive indicating an ATCO in command of the task, and using the system interface
	Swearing, huffing, Apologetic for mistakes & wrong actions	x	x		Verbal frustration from ATCO regarding levels of competence undertaking the task with the new system
	Relaxes when quiet, sits back, crosses arms	x	x		Adapts to the pace of the task as required (relaxes when quiet, sits up when busy)
	Sitting up alert and attentive	x	x	x	
	gestures of waving hands about, blowing air out	x	x		Physical signs of frustration and confusion, (often combined with non-verbal communications)
	rubbing face, yawns, rubs eyes, "I'm tired", looks at watch	x	x		Demonstrates signs of tiredness, and fatigue, and that they have had enough of the task session
	Adjusts MMI to the needs and requirements of the task	x	x	x	Has sufficient capacity, and desire, to adjust the MMI to attain a better workstation setup
	Taps pen, wriggles about, taps leg	x	x	x	Potentially an indicator of restlessness, and distraction

Table D4 - Behavioural Markers identified through observation - According to where observed - Part 1

Notes:

- 'x' indicates behaviour was observed
- '1' denotes Tower ATCO group (<3 hours of simulator experience with EFPS)
- '2' denotes en-route ATCO group (<25 hours of simulator experience with iFACTS)
- '3' denotes a different Tower ATCO group (several years of experience using EFPS)

Category	Behaviour observed	ATCO Group 1	ATCO Group 2	ATCO Group 3	What is the value of this marker?
Inputs and Interaction with HMI and Workstation	Demonstrates spatial/muscle memory of interface and layout functions			x	Strong motor skills and intuitive understanding of interface layout
	Overconfident, fast, but makes mistakes and select wrong functions, performs unnecessary tasks	x	x		Has confidence and knowledge but still developing the underpinning skills needed to perform efficiently and effectively
	Slow and Hesitant, indecisive, unsure of actions, moving to control a function then moving back, requires multiple attempts to drive the HMI	x	x		A lack of knowledge, experience, and skill using the task interface.
	Plays and experiments with system to see how it responds and behaves, and to rehearse actions		x		'sand-pit' behaviour indicative of growing understanding of global actions
	Dual Tasking/Multi Tasking 4 whilst engaged in RT will also drive HMI and input data	x	x	x	Multi-tasking is a core skill encourage as best practice within ATC
	Confident smooth flowing control. Deft command of HMI	x	x	x	Has confidence and knowledge and skill in using the interface
	Using HMI incorrectly 4 wrong clicks, taps but no action on interface, can't find right function etc	x	x		Has a lack of knowledge and skill of the HMI
	Surprised by behaviour of HMI	x	x		
	Quickly located required functionality and information when required	x	x	x	Intuitive understanding of interface layout
	Displays 'automaton' type actions when progressing through a sequence of steps	x		x	
	Picks up activities (e.g. changes in HMI) on the periphery	x		x	Global and peripheral awareness (not tunnelled into the HMI).
	Plans task ahead 4 Opens up windows in advance, leaves cursor in the position needed for the next action or an action that they need to return back to, highlights all strips requiring QNH update prior to departure			x	Proactive in the task, has developed personal strategies to enhance their performance and task management
	keeps information windows active and open with applicable info during read back	x	x	x	has developed personal strategies to enhance their performance and task management
	Hovers pen/cursor over info on read back to confirm/tick off	x		x	
	Looks out of the tower, checking arrivals and departures on radar, checking surface radar, checking the EFPS display(s), and other ancillary displays as appropriate (e.g. lighting panels, ATIS)	x	x	x	Global and peripheral awareness (not tunnelled into the HMI).
	when driving the HMI (e.g. tapping across taxi way windows in a 3, 2, 1, motion, moves the cursor round in circles especially on-screen objects, adjusts the windows so they line up absolutely perfectly			x	Has developed individual 'Quirks', mannerisms and supportive strategies
Interaction with others	Looks for affirmation from colleagues or instructor, before an action	x			Still growing in confidence and looking for support before making an action
	Looks for affirmation from colleagues or instructor, after an action	x			Although performs an action, is under confident and looks to acquire affirmation that it is was appropriate/correct
	Prompted by others 4 Oh right, yes, I see, ah ok	x			Lacking skill and ability 4 still requiring coaching to perform core actions and activities
	Displays the ability to undertake dichotic listening Offers suggestions and discusses options with team members		x	x	Peripheral awareness, awareness of others; spare capacity and global awareness
	Snappy and irritable when busy, and when been put under pressure		x		Being pushed hard and taking out frustrations to surrounding individuals
	Engages in technical discussion	x	x		Keen to reinforce knowledge and explore boundaries of understanding

D03 - FURTHER POTENTIAL BEHAVIOURAL MARKERS

Potential Marker	Indicator of...	Assessment
Clumsiness moving and interacting with paper strips	Demonstrates physical slowness at moving strips and interacting with workstation, movements over emphasised	These markers were all observed within the first day of observation and were considered to offer useful insight into proficiency and as a result they were tracked across the remaining 6 days of observations
Nervous Physically hand shaking	Shaking hands, shoulders, and holding and writing with a pen	
Both hands to move strips (poor MMI motor skills)	Experienced controllers move strips with one hand, quickly and deftly, using two hands to move them is extremely unusual	
frantic writing/frantic task/rushed	Unless in exceptional situations ATCOS approach the task calmly, not frantically	
Nervous or Extremely Quiet Voice	A strong, positive, and commanding voice is required, nervous quiet voices are undesirable	
Points things out to themselves or instructor	Re-enforcing spatial information through physical referencing	
leans right over strips	Tunnelled into the workstation, rather than sitting back and absorbing 'the big picture'	
scans through and sorts strips	Developing systematic processes to the task	This was considered weak, as it provides no strong insight into the task, and is a fundamental ATC activity which occurs all the time
looks relaxed	Comfortable user state	This was considered weak, similar to 'cool and calm' a marker removed from earlier versions of the observation sheet
hand presses different bays	Potentially a quirk using a methodical hand gesture to reinforce a scan	This was only seen on one occasion with only one trainee, and therefore considered of low value
heads down	Tunnelled into the workstation and task and reduced awareness on the periphery	This was considered weak, highly subjective, and therefore considered of low value
makes corrections (RT corrections)	Indication of best safety practise	This has been seen over the course of this research, but extremely infrequently
moves seat to re-centre on strips	Considers their working ergonomics important	This was considered weak, offering low value, and therefore considered of low value
touches strips on receipt	Self styled re-enforcement check	Covered under existing marker - 'automatic/quick'
Self affirm with gestures (hand movements almost illustrating thought process)	Indication of learned knowledge being applied	Rejected - hard to interpret the meaning behind the gestures (only displayed by a single trainee)
holds strip in strip holder	Strong physical bond to strip to maintain mental focus on the immediate task	This was considered weak, highly subjective, and therefore considered of low value
keeps hand on strips during comms	Strong physical bond to strip to maintain mental focus on the immediate task	This was considered weak, large individual difference, and therefore considered of low value
very positive actions pushing things into place	Poor motor skill	Considered to have insufficient mutual exclusivity with clumsy and 2 hands to move strips
uses a pen as a pointer	Growing spatial awareness of the task, and task environment	Considered to have insufficient mutual exclusivity with points things out/spatial reference
taps strip as a memory tap	Self styled re-enforcement check	This was considered weak, highly subjective, and therefore considered of low value
memory aids/cocks strips	Standard best practice in use	Best practice skills are not included within this marker system
fingers through strips	Maintains the strip board	Common, frequent, repetitive, standard tasks, considered to offer low insight into development and therefore rejected
sorts/tidies strips		

Table D6 - Further potential Behavioural Markers identified through observing Trainee ATCOs

D04 - EXPERT ATCO GROUP - INDIVIDUAL BEHAVIOURS

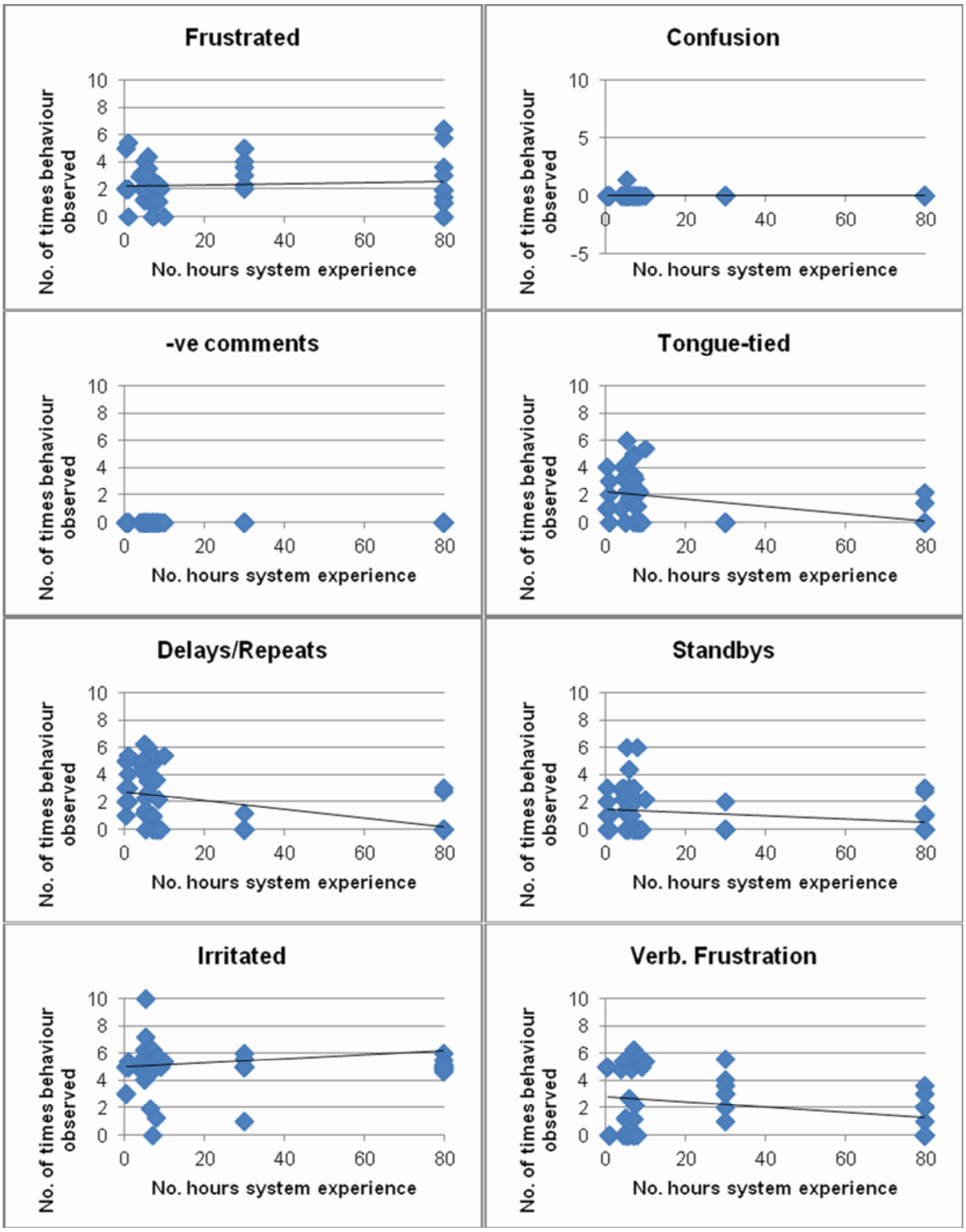


Figure D1 - Expert ATCO Group, Individual Beginner Behavioural Markers

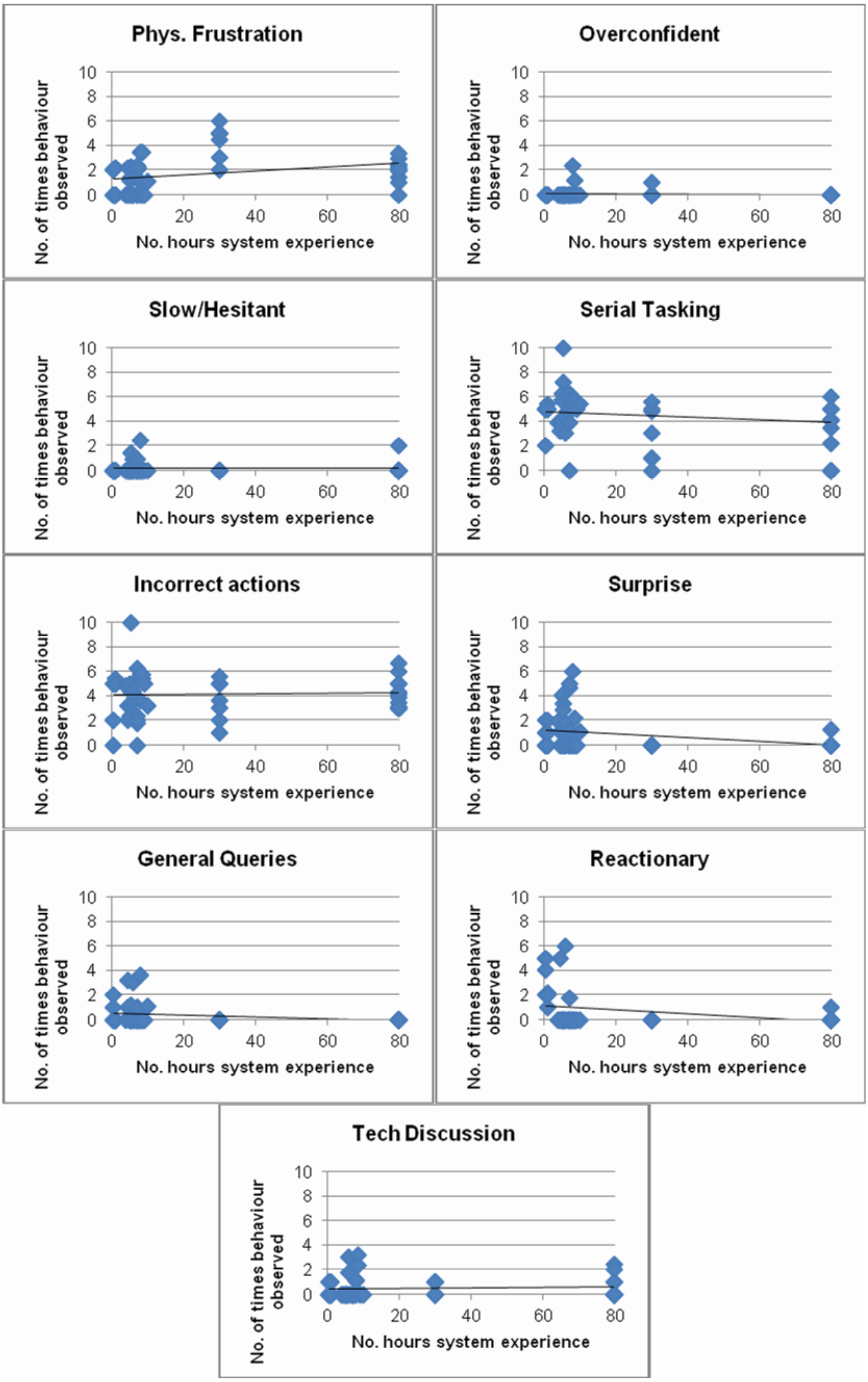


Figure D2 - Expert ATCO Group, Individual Beginner Behavioural Markers continued

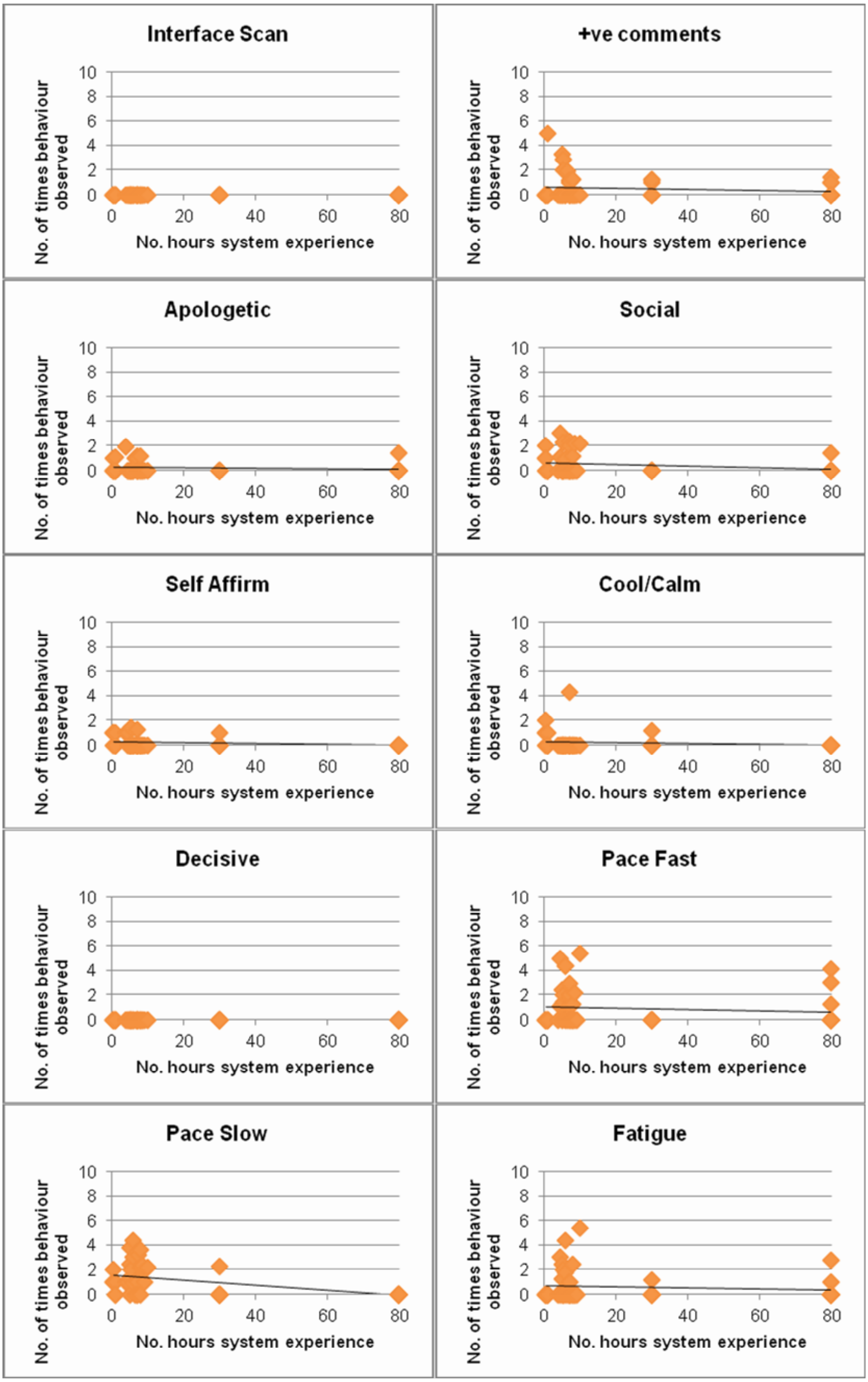


Figure D3 - Expert ATCO Group, Individual Intermediate Behavioural Markers

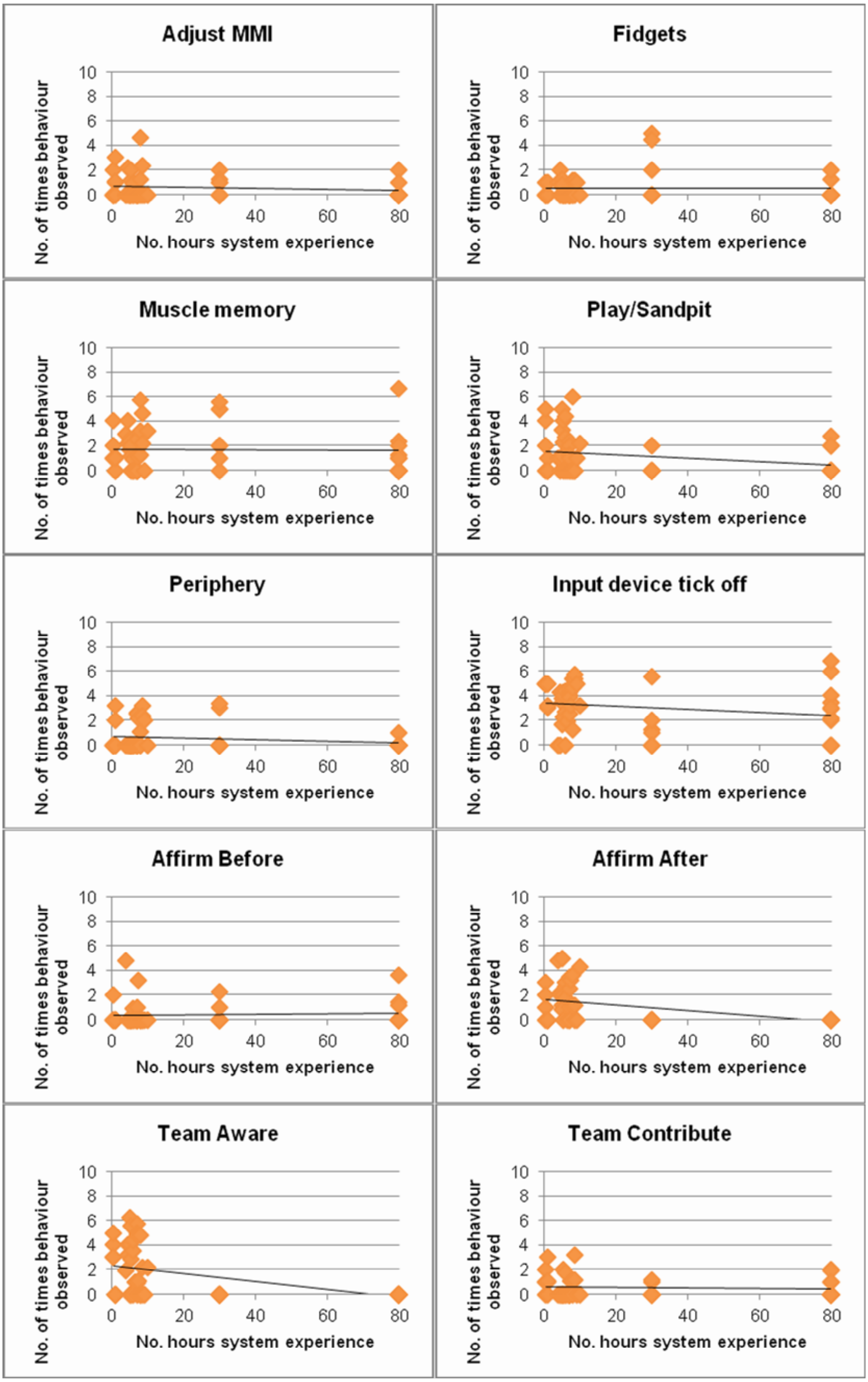


Figure D4 - Expert ATCO Group, Individual Intermediate Behavioural Markers continued

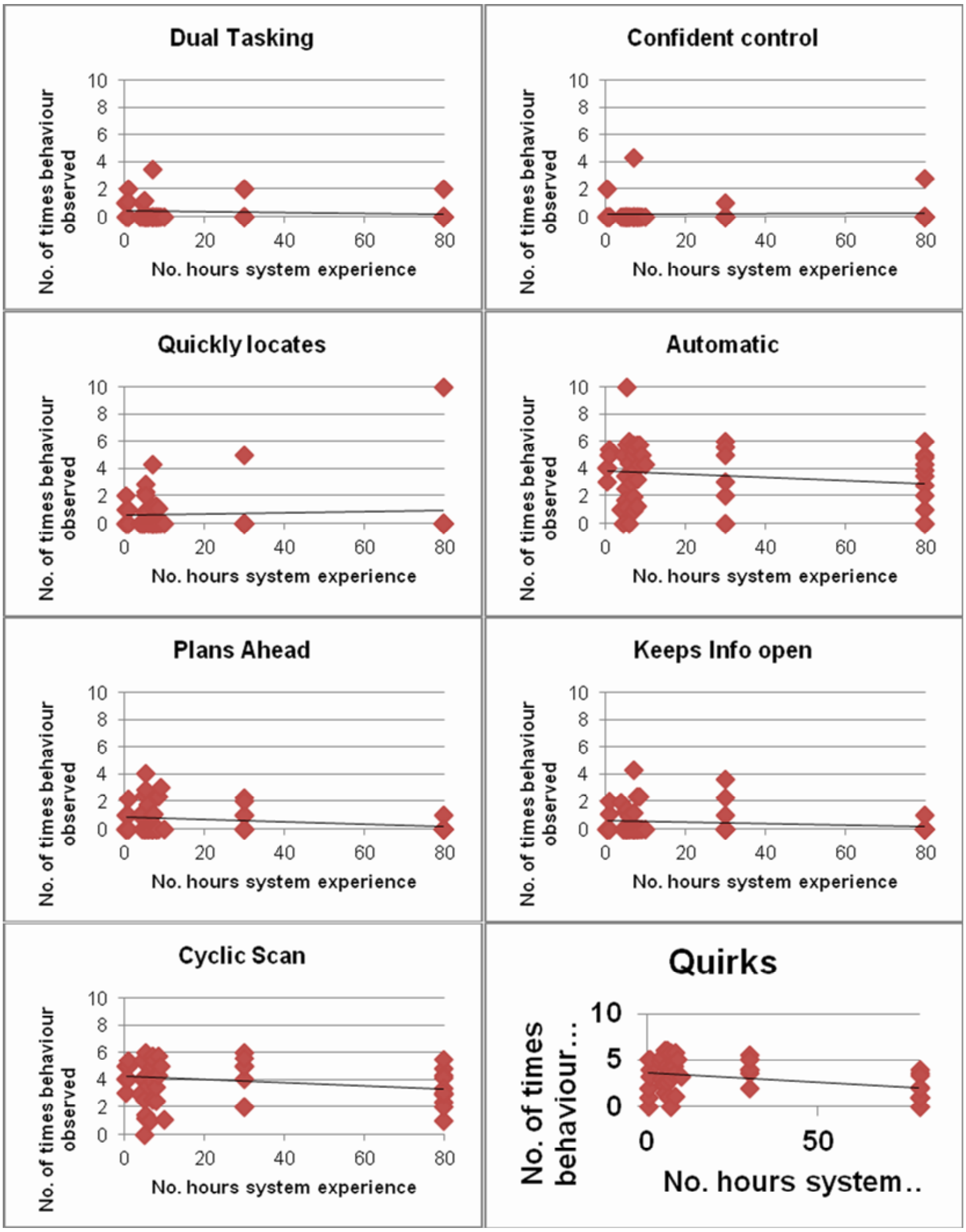


Figure D5 - Expert ATCO Group, Individual Expert Behavioural Markers

D05 - QUALITATIVE INTERVIEW TRANSCRIPT

Q. How did you find the layout of the markers on the sheet, did you find it straight forward to locate the appropriate marker? Q. Did you find that you were hunting for right behaviour on the sheet, how was the design layout?
Observer A: Fine, [the] groupings are appropriate, I had many more interactions with HMI & workstation wonder if because some are duplicates, our natural tendency to look at the measurable, to observe what we understand.
Observer B Answer: I liked the layout, nice and simple, clear. I did have to look up several markers, as they weren't clear, wasn't sure what the difference was. Not sure about the balance of markers across the [six] categories; HMI interaction category contains many more markers, potential duplicates? The imbalance is potentially due to the limitations of what can actually be observed (and hence what markers are contained).
Q. How did you find the observation process, did you find it useful?
Observer A: hmm, not sure if [the] observation is measuring the impact of the system change, or people's individual's behaviours and individual difference. There is a need for base lining. So if it is a relaxed task, and you put a change in it may still be relaxed. If it is a stressful task and you put a change it t may still be a stressful task. Baseline behavioural survey before and after the system goes in. When you use this you will get snap shots. Perhaps the baseline doesn't matter as long as you see a movement (up or down).
Observer B: [I'm] not fully confident using the behavioural observation system, it was hard to identifying correct behaviours from the categories. I think further exposure and experience using the [observational system] would likely improve [my] understanding of the technique. It is hard work and resource heavy, but gaining insight from this which is useful for training and assurance, is backed up questionnaire feedback from validation. Some behaviours were harder to observer than others (the behaviours were subtle, and discrete).
Q. How did you find 30 min as the length of the observation period?
Observer A: Right length of time for the observer, but whether for the observee. If you are looking for influences like fatigue, or degrading performance you'd have to change it. E.g. a very demanding over an hour you would look at the hour. But it is tricky to do for more than 30 min 4 you wouldn't want to do many a day. Fine if you are in a nice sunny tower 4 but in TC would get worn out very quickly.
Observer B: Guidance material provided to observers needs to makes reference to the limitations of a single observation period. Not all behaviours contained may manifest themselves. Also, when during the day or night the observations themselves take place. There may be differences between watches with [some] more formal than others. The timing of observations, where possible, are scheduled to accommodate potential differences across watches.
Q. I noticed you made notes whilst observing (both observers did this) did you find this important? Was there enough space on the sheet to allow you to do this?
Observer A. Absolutely, but It's a note sheet not fixed data entry. The sheet/tool is flexible; I like that. No redesign needed.
Observer B: Notes allow queries to be discussed following the observations, like the definition of certain markers. I also recorded [the] context of a particular behaviour any new markers, or the re-design of specific individual markers. I just scribbled notes on the back.
Q. From your experience making using this sheet, and observing users interacting with EFPS, do you feel you would be able to make a judgement as to the overall level of development of the controller (perhaps using a simple scale)?
Observer A: Yeah, felt that I was able to do so; it would provide a useful summary of the observed person's behaviour.
Observer B: I think that any ratings could be seen as a direct measure of task performance; and I don't feel qualified to provide that.
Q. Were there specific condition, which you may have noted, that you may consider to have had an impact upon the ATCO observed (environmental and traffic situation, state of the user etc)?
Observer A: Difference/dynamic between watches (some formal, some less formal). Finding this out would help determine if the observations made are independent of the unit/unit factors.
Observer B: experience, conditions of the day, state of the user at the time; all effect what behaviour is exhibited. Not sure if we impacted upon controller behaviour, I suppose they [are] regularly observed whilst working, and this doesn't usually affect them too much.

Table D7 - Interview Transcript of HF Observers

Q. How did you find the categories used to categorise the markers? Do you feel they are appropriate?
Observer A: Categories worked quite well. Didn't like muscle memory etc. couldn't try out a lot of them as many of the interaction with others ones are found only in training.
Observer B: seemed fine, nothing to say really.
Q Did you find the markers mutually exclusive, and were they clear to delineate from one another?
Observer A: Not found it hard, as with a freq. you are getting the fact that they are doing it 4 rather than the detail of the freq. count.
Observer B: Not difficult to delineate. Used the reference sheet.
Q What are your views on the five count frequency limit? Q. Was there any pattern to your observation process?
Observer A: sensible to set a limit, allows you to move on. Observe other things. Prevent unintentionally tunnelling upon a single behaviour and focusing upon it to the detriment. Yes, I did find my focus would move on from certain markers once they had been observed.
Observer B: fine. To be honest, the frequency is not the most important element in the [observational] process, but rather the behaviour is identified and recorded. There was a pattern to the observation [but I] can't explicitly state
Q. Did you notice any additional behaviours? Any new markers?
Observer A: Behaviours/sub behaviours yes. Body language, body positions sitting back sitting forwards, hovering pen for a few min [planning ahead, being in control]. First guy sat there, in the chair, leaning back. Showing great awareness in the task. Not interacting with EFPS, but able to provide RT to aircraft without integrating the screen. He had strong picture of the traffic. Contrasted against the other controller who was leaning in to EFPS and looking under confident. Perhaps indicative of someone who's been controlling for many years, with high spatial and cognitive skills.
Observer B. Hunched over the HMIs (behaviour of unfamiliarity).

Table D8 - Interview Transcript of HF Observers continued

D06 - QUANTITATIVE INTER RATER AGREEMENT

Marker	Number of instances					
Negative comments	5	2	0	0	0	0
Positive comments	4	2	1	0	0	0
Apologetic	6	0	0	1	0	0
Social	1	6	0	0	0	0
Confusion	4	1	1	1	0	0
Tongue-tied	4	3	0	0	0	0
Self Affirm	4	2	1	0	0	0
Delays (Delays/Repeats)	4	1	1	1	0	0
Standbys	5	2	0	0	0	0
Cool/Calm	2	1	2	1	0	1
Irritated	7	0	0	0	0	0
Decisive	1	0	2	1	2	5
Verb. Frustration	5	1	0	0	0	1
Pace Fast	5	2	0	0	0	0
Pace Slow	3	1	1	0	1	1
Phys. Frustration	5	1	0	1	0	0
Fatigue	4	2	1	0	0	0
Adjust MMI	2	2	2	1	0	0
Fidgets	2	3	1	1	0	0
Muscle memory	6	0	0	0	0	1
*Overconfident	7	0	0	0	0	0
Slow/Hesitant	6	0	0	0	1	0
Play/Sandpit	4	3	0	0	0	0
Dual Tasking	2	2	2	0	1	0
Serial Tasking	6	1	0	0	0	0
Confident control	0	3	3	0	1	0
Incorrect actions	4	2	0	0	0	1
Surprise	3	3	0	0	0	1
Quickly locates	2	2	2	1	0	0
Automatic	3	0	2	1	0	1
Periphery	4	3	0	0	0	0
Plans Ahead	3	1	0	2	0	1
Keeps Info open	4	2	0	1	0	0
Input device tick off	0	0	0	1	2	4
Cyclic Scan	3	2	0	1	1	0
Quirks	5	1	1	0	0	0
Affirm Before	7	0	0	0	0	0
Affirm After	5	2	0	0	0	0
General Queries	5	1	0	1	0	0
Reactionary	5	2	0	0	0	0
Team Aware	6	0	0	1	0	0
Team Contribute	4	2	1	0	0	0
*Team Short/Snappy	7	0	0	0	0	0
Tech Discussion	3	0	2	2	0	0

Table D9 - Distribution of Behavioural Markers, based upon levels of Inter-Rater Agreement ('*' denotes not observed)

D07 - QUALITATIVE & QUANTITATIVE INTER-RATER COMPARISON

Behavioural Marker	Observer A Comments	Observer B Comments	Level of Inter-Rater Reliability (Table 8.1)	Proposed changes discussed	Actual changes made
Overconfident	No comment	No comment	N/A	Keep	No Change
Team Short/Snappy	No comment	No comment	N/A	Keep	No Change
Negative comments	I observed this several times	No comment	High	Keep	No Change
Positive comments	No comment	No comment	High	Keep	No Change
Social	This was very easy to spot	No comment	High	Keep	No Change
Tongue-tied	No comment	I found this useful information, it was rich listening into comms (especially if you have an insight into them)	High	Keep	No Change
Self Affirm	No comment	No comment	High	Keep	No Change
Standbys	No comment	Is this bad? This can be an indication of workload (but also confusion), so need take into account the traffic levels	High	Clarify 4 Observer Guidance Sheet	Clarification added (standbys due to confusion, hesitation etc)
Irritated	No comment	This is a good one	High	Combine with Verbal Frustration	Combined with Verbal Frustration
Pace Fast	No comment	Not sure how useful this marker is?	High	Revise 4 displays behaviours associated with High/low pace (Adapts To The Pace Of The Task)	No Change
Fatigue	Yes I saw this, I saw someone yawning	No comment	High	Keep	No Change
Play/Sandpit	I found it difficult to differentiate this from 'planning ahead'. In a live environment unlikely to see this (situation depended)	Didn't see much of this behaviour	High	Clarify 4 Observer Guidance Sheet	Clarification added (may be situation dependant)
Serial Tasking	No comment	No comment	High	Keep	No Change
Periphery	Didn't get this but I think it is a good one	No comment	High	Keep	No Change
Quirks	Didn't notice any quirks, although I have seen quirks outside of these observations	No comment	High	Keep	No Change

Table D10 - Qualitative HF Observer comments combined with Quantitative levels of Inter-Rate Agreement

Behavioural Marker	Observer A Comments	Observer B Comments	Level of Inter-Rater Reliability (Table 8.1)	Proposed changes discussed	Actual changes made
Affirm Before	These are only applicable to the training environment, rather than live ops.	No comment	High	Clarify 4 Observer Guidance Sheet	Clarification added (make explicit that these are training environment only)
Affirm After	These are only applicable to the training environment, rather than live ops.	No comment	High	Clarify 4 Observer Guidance Sheet	Clarification added (make explicit that these are training environment only)
Reactionary	These are only applicable to the training environment, rather than live ops.	No comment	High	Clarify 4 Observer Guidance Sheet	Clarification added (make explicit that these are training environment only)
Team Contribute	Fine 4 Got this 4 You need to confirm What is the team (e.g. the person of the runway looking at birds on the radio to the tower 4 is he part of the team (controller with binoculars looking out during communication)	No comment	High	Clarify 4 Observer Guidance Sheet	Clarification added (what constitutes the 'team')
Apologetic	No comment	Didn't get many of these	Medium	Keep	No Change
Confusion	I'm happy with it 4 A good example I saw was "I don't know"	Wasn't sure if this was meant to be physical confusion.	Medium	Clarify 4 Observer Guidance Sheet	Clarification added (verbal confusion')
Delays (Delays/Repeats)	No comment	No comment	Medium	Keep	No Change
Phys. Frustration	No comment	Is this voice frustration?	Medium	Clarify 4 Observer Guidance Sheet	Clarification added (physical frustration only)
Adjust MMI	I don't think I recorded these all of the time	No comment	Medium	Keep	No Change
Fidgets	Got some of these 4 how long is a period of fidgeting?	No comment	Medium	Clarify 4 Guidance Sheet	Clarification added (per fidget period ~ 30 seconds)
Confident control	Confused with verbal cool/calm	No comment	Medium	Revised	Combined with muscle memory, automatic, and quickly locates
Incorrect actions	No comment	No comment	Medium	Keep	No Change
Surprise	No comment	No comment	Medium	Keep	No Change
Quickly locates	Again this is kind of the same as quickly locates /muscle memory / confident control	This is similar to automatic, and muscle memory	Medium	Revised	Combined with muscle memory, automatic, and confident control
Keeps Info open	No comment	No comment	Medium	Keep	No Change

Table D11 - Qualitative HF Observer comments combined with Quantitative levels of Inter-Rate Agreement continued

Behavioural Marker	Observer A Comments	Observer B Comments	Level of Inter-Rater Reliability (Table 8.1)	Proposed changes discussed	Actual changes made
General Queries	These are only applicable to the training environment, rather than live ops.	No comment	Medium	Clarify 4 Observer Guidance Sheet	Clarification added (make explicit that these are training environment only)
Team Aware	Fine 4 Got this one	Had to look this one up to see the difference with this and team contribute	Medium	Keep	No change
Tech Discussion	No comment	No comment	Medium	Keep	No change
Cool/Calm	As we are looking for people out of training to have this, we are looking for professional controller voice	Got this all the time. I Noted in the comments anything in particular	Low	Delete	Marker Deleted
Decisive	How is it different from cool/calm	Basic ATCO comms. This balances with overconfident. Similar to with confident control. A trainee ATCO would not validate without this skill	Low	Revised	Marker changed to Indecisive
Verb. Frustration	Happy with this	Similar to irritated? Could combine with irritated.	Low	Revised	Combine with Irritated
Pace Slow	No comment	Not sure I saw this one.	Low	Revise 4 displays behaviours associated with High/low pace (Adapts To The Pace Of The Task)	No Change
Muscle memory	Didn't like this one, found it hard to differentiate from 'automatic'. Not sure if I did this right	No comment	Low	Revised	Combined with confident control, automatic, and quickly locates
Slow/Hesitant	No comment	Is this the same as 'delays / repeats'	Low	Clarify 4 Observer Guidance Sheet	Clarification added (physical frustration only)

Table D12 - Qualitative HF Observer comments combined with Quantitative levels of Inter-Rate Agreement continued

Behavioural Marker	Observer A Comments	Observer B Comments	Level of Inter-Rater Reliability (Table 8.1)	Proposed changes discussed	Actual changes made
Dual Tasking	This is essentially WAYS RAYL 4 so potentially every time they speak you could put this down	No comment	Low	Clarify 4 Observer Guidance Sheet	Clarification added (this is similar to WAYSRAYL, but it is different as it covers not just writing whilst speaking but also interacting whilst speaking. A less prescriptive to cover other dual task interaction)
Automatic	Again this is kind of a 'quickly locates', 'muscle memory', 'confident control' marker	No comment	Low	Revised	Combined with muscle memory, confident control, and quickly locates
Plans Ahead	What if they are idle clicking a window! Saw plenty of these 4 opening boxes (awaiting for a call to come in) Really need to understand the task on this one. Some of the HMI mouse click stuff on the reference sheet is probably realistically too hard to spot. A good example 4 holding a dibber in anticipation of action.	Difficult one to assess 4 and possibly requires expert task knowledge.	Low	Delete	Marker Deleted
Input device tick off	I Didn't get any of these 4 I saw interacting with strips I know the action but not sure if I saw it. Feel this is maybe linked closely with observers task/system knowledge	Struggled to see this/observe this as it can be very subtle, and depends on where you are positioned as an observer	Low	Delete	Marker Deleted
Cyclic Scan	What does it contribute. You can't say it was a cyclic scan. I see a cyclic scan as like driving "mirror signal". Maybe Maintains awareness/Maintains external awareness/maintains situation awareness. Yeah you could do it that way. Cyclic scan is [more of] a control centre thing.	Discussed changing this to maintains awareness	Low	Revised	Change to " maintains global awareness"

Table D13 - Qualitative HF Observer comments combined with Quantitative levels of Inter-Rate Agreement continued

D08 - TRAINEE ATCO GROUP - INDIVIDUAL BEHAVIOURS

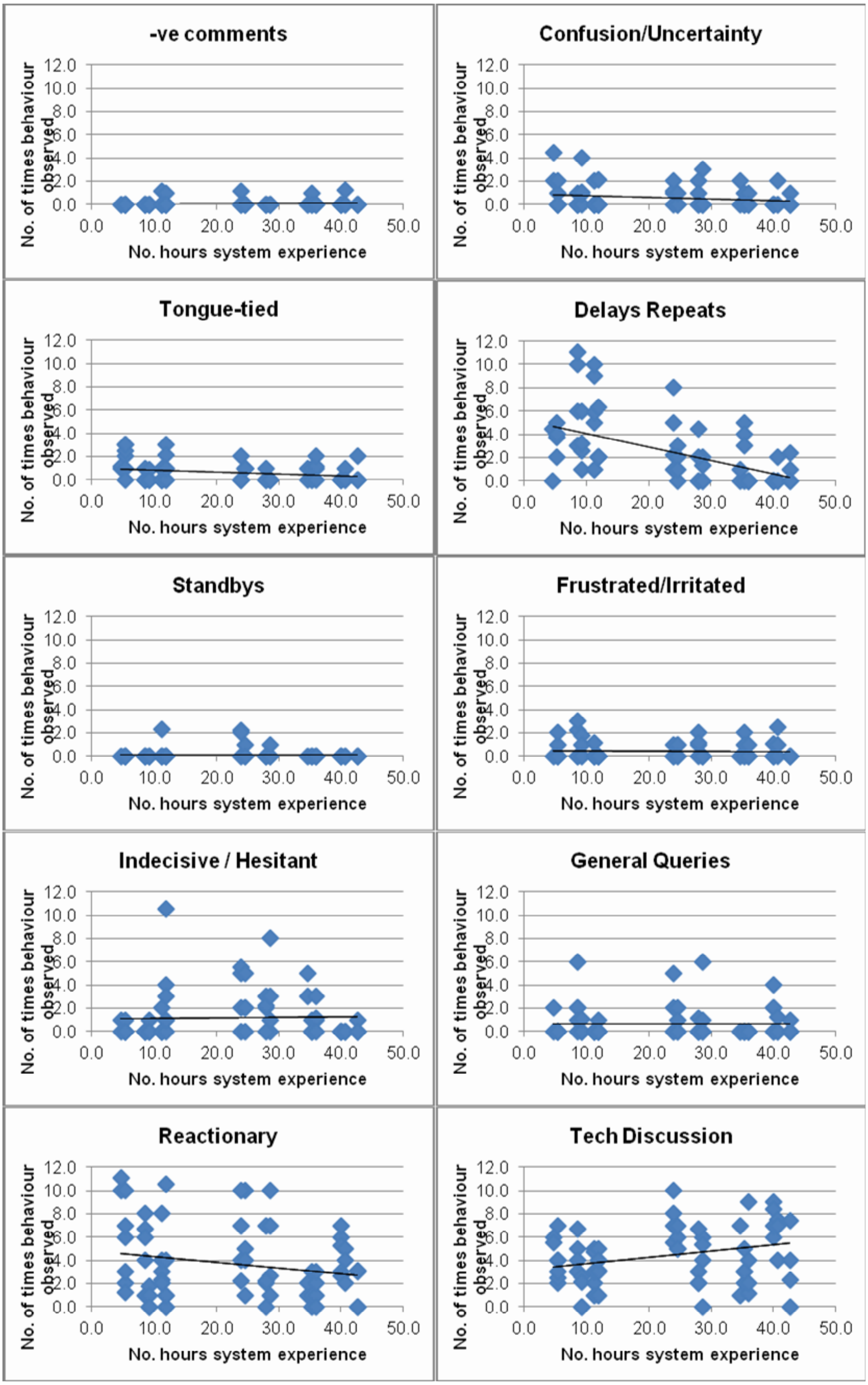


Figure D6 - Trainee ATCO Group, Individual Beginner Behavioural Markers

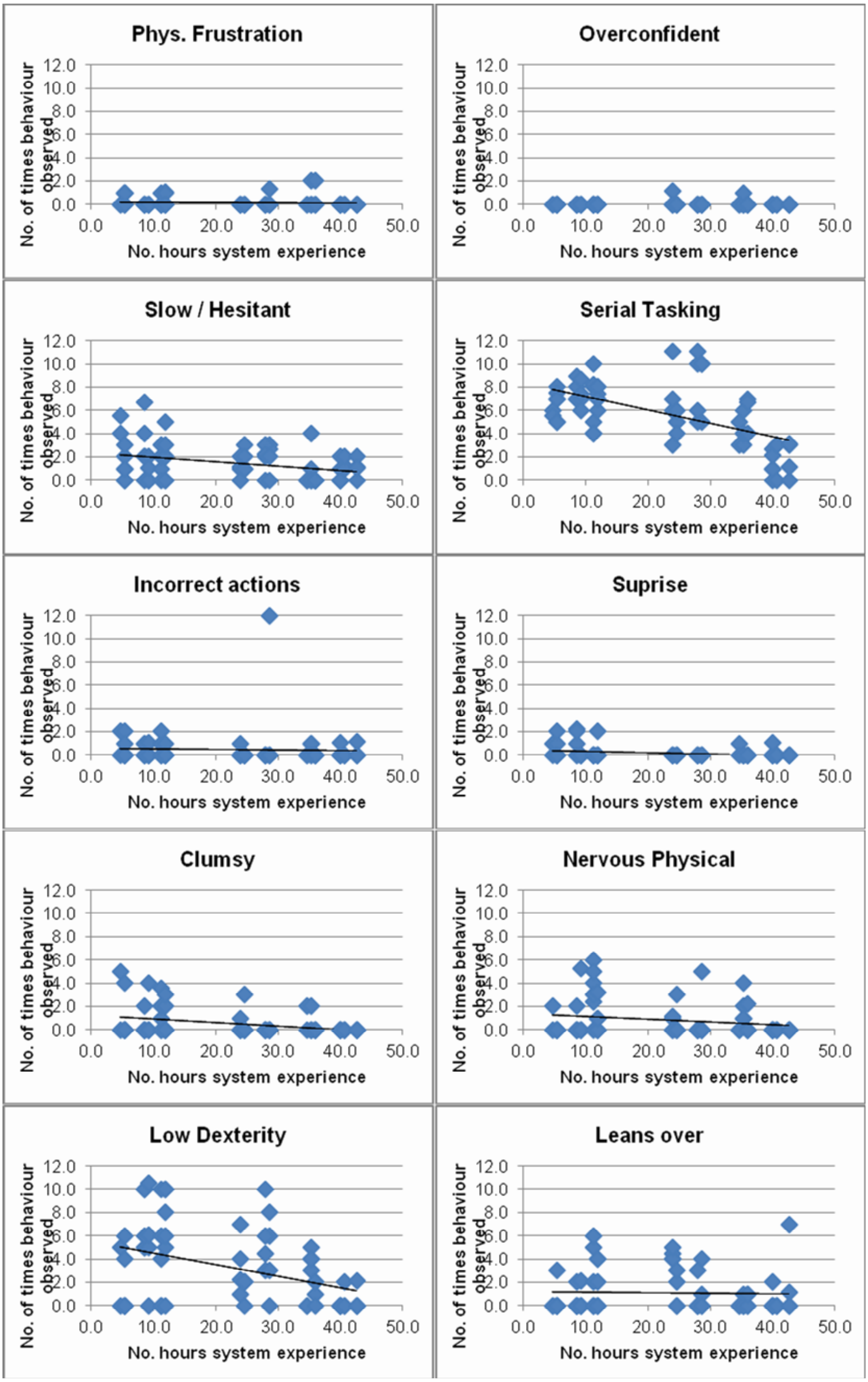


Figure D7 - Trainee ATCO Group, Individual Beginner Behavioural Markers continued

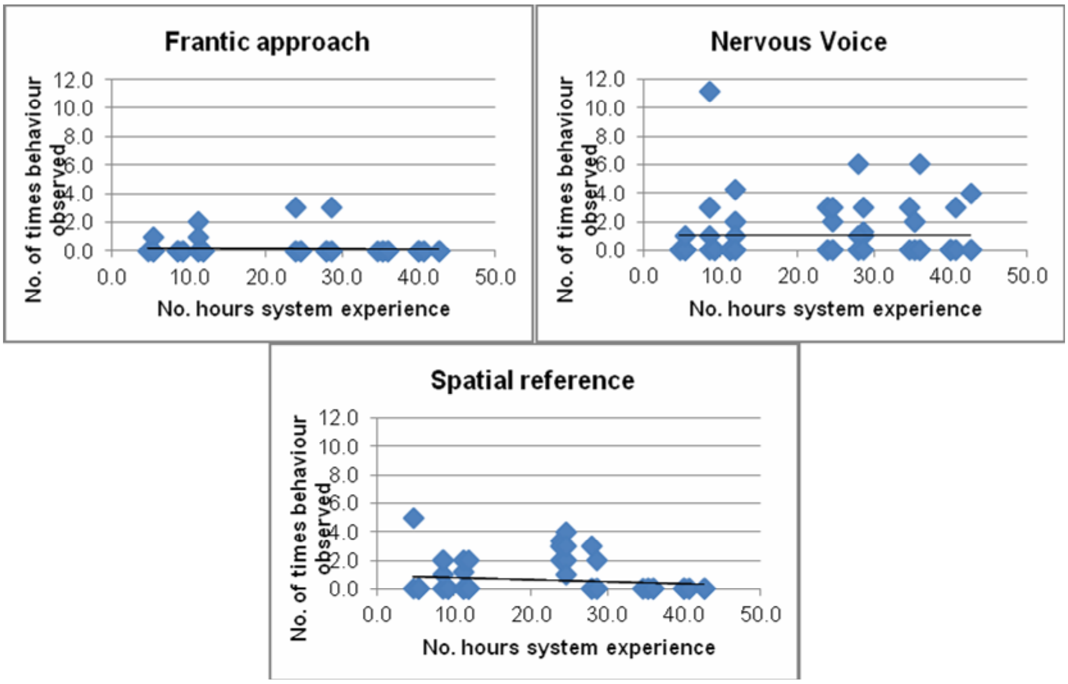


Figure D8 - Trainee ATCO Group, Individual Beginner Behavioural Markers continued

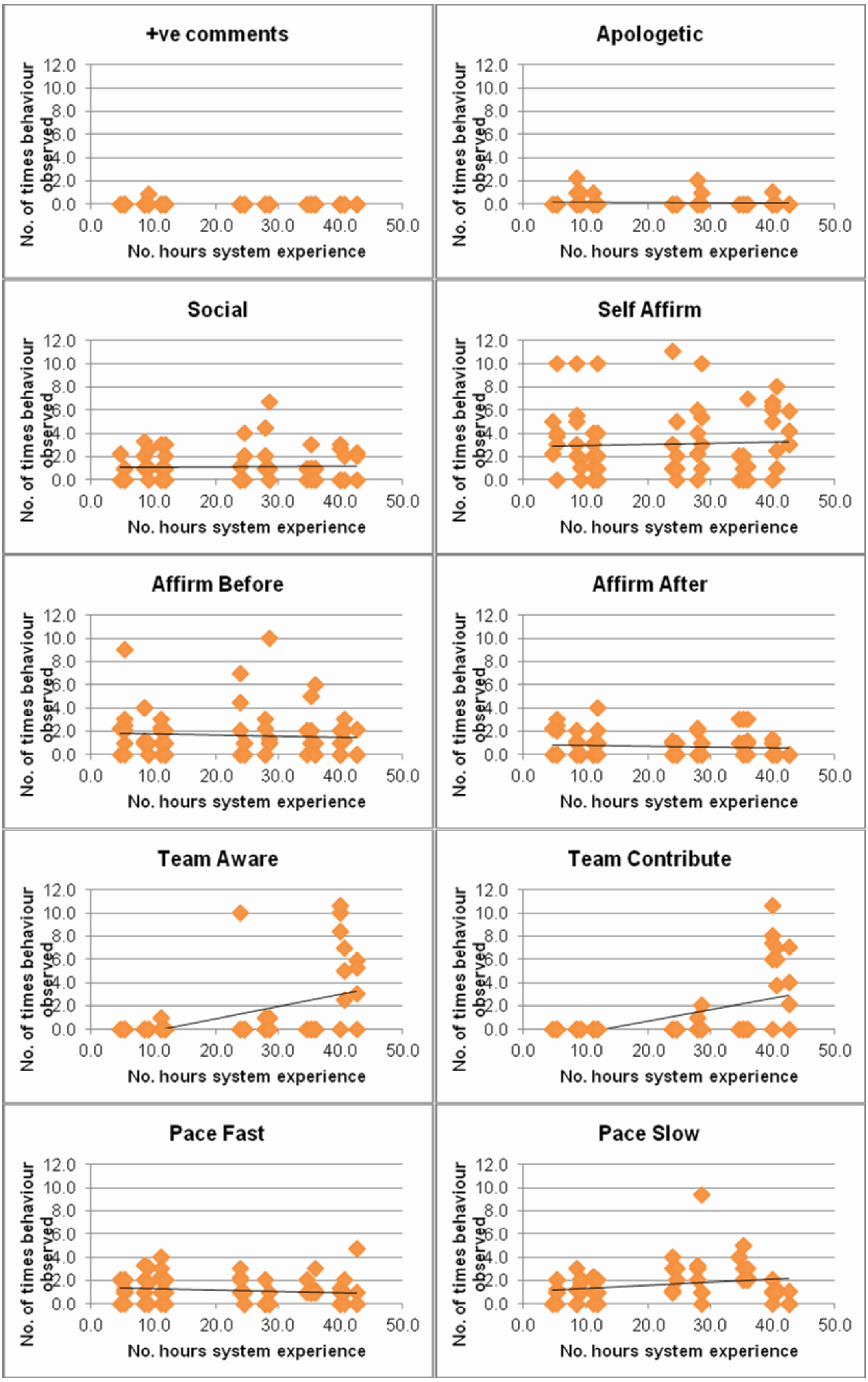


Figure D9 - Trainee ATCO Group, Individual Intermediate Behavioural Markers

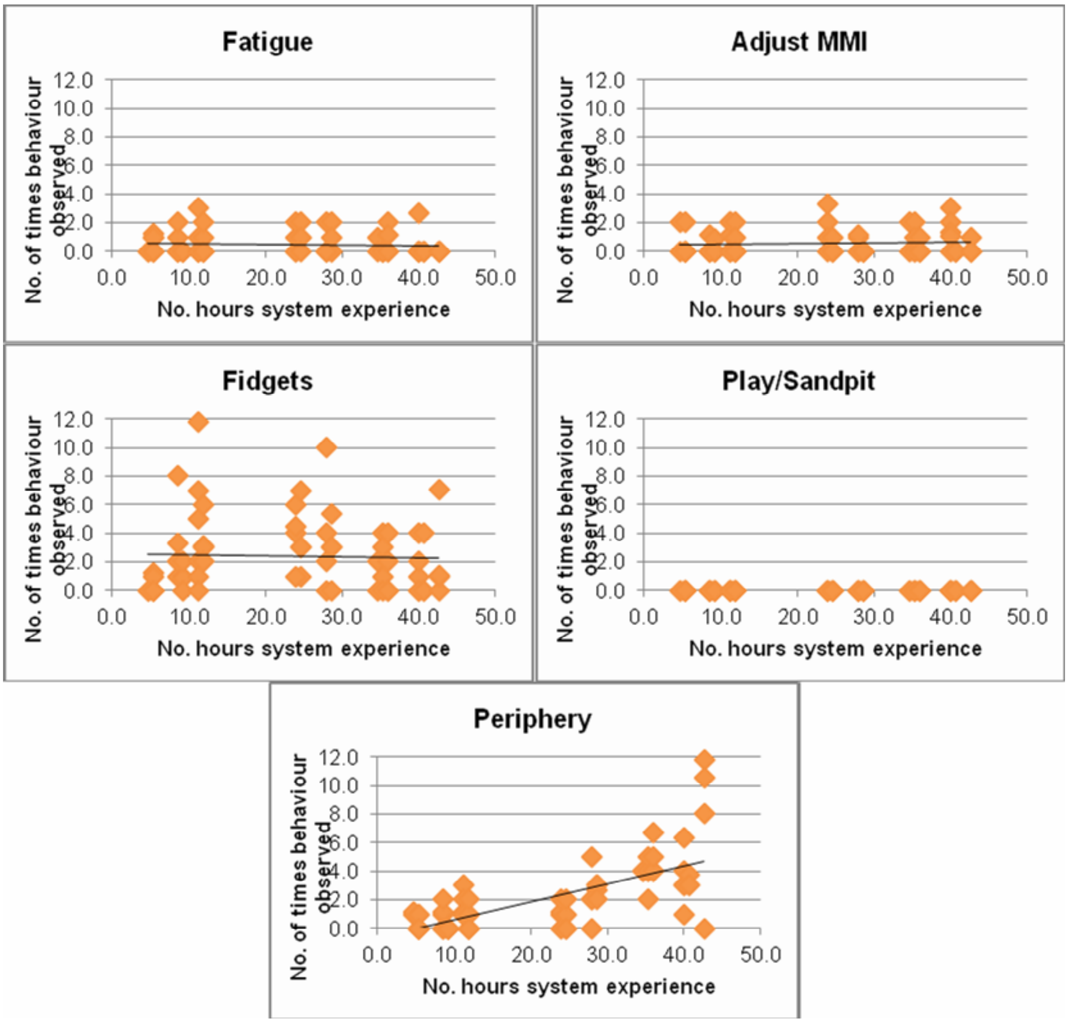


Figure D10 - Trainee ATCO Group, Individual Intermediate Behavioural Markers continued

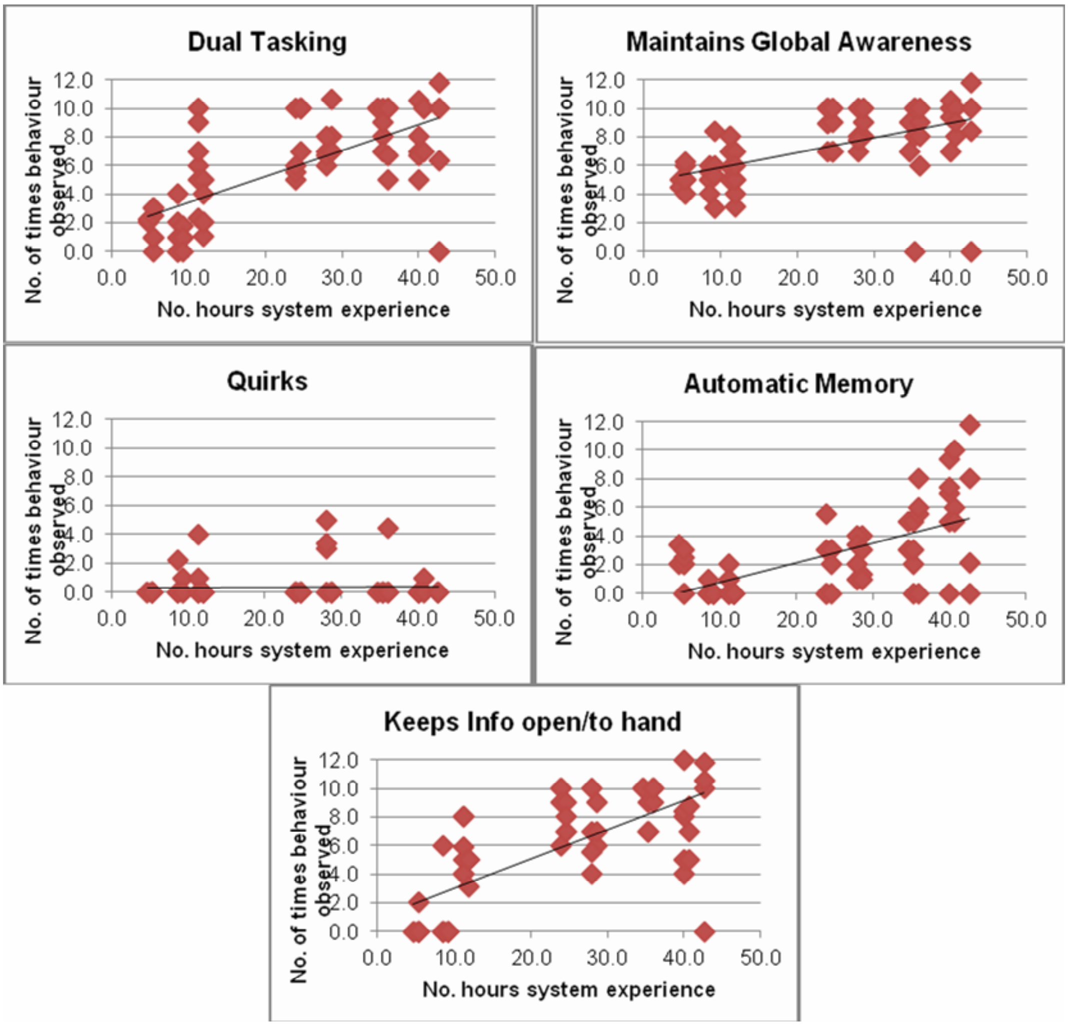


Figure D11 - Trainee ATCO Group, Individual Expert Behavioural Markers

D09 - INDIVIDUAL TRAINEE ATCO RESULTS

Trainee ATCO	alpha		bravo		charlie		delta		echo		foxtrot		golf		hotel		indigo		Juliet	
N	7		6		6		8		6		6		6		7		7		8	
Behavioural Marker																				
Beginner Behavioural Markers																				
Team Short/Snappy																				
Negative comments			.655	.158	-.372	.468	-.372	.468					-.393	.441	.204	.661				
Confusion/Uncertainty	.408	.182	-.655	.158	.034	.949	.034	.949	.541	.268	-.207	.694	.131	.805	-.906	.005*	-.281	.542	-.626	.097
Tongue-tied	-.299	.258	-.676	.140	-.395	.439	-.395	.439	.393	.441	-.091	.864	.802	.055	-.632	.127	-.356	.434	-.192	.650
(Delays/Repeats	-.564	.094	-.880	.021*	-.319	.538	-.319	.538	-.657	.156	-.759	.080	-.559	.249	-.197	.672	-.595	.159	-.783	.022*
Standbys					-.131	.805	-.131	.805					-.439	.383	.000	1.00				
Frustrated Irritated	.204	.330	.507	.305	-.338	.512	-.338	.512	-.677	.140	.207	.694	-.213	.686			-.791	.034*	.171	.685
Indecisive / Hesitant	.335	.231	-.131	.805	-.213	.686	-.213	.686	.432	.392	-.131	.805	.372	.468	-.433	.331	.162	.728	.518	.188
General Queries	.374	.204	.507	.305	-.304	.558	-.304	.558	-.131	.805	.304	.558	-.655	.158	-.359	.430	.612	.144	-.187	.657
Reactionary	.334	.232	-.232	.658	-.203	.700	-.203	.700	-.406	.425	-.257	.623	-.841	.036*	.000	1.00	-.288	.531	.204	.629
Tech Discussion	.847	.008*	-.058	.913	.143	.787	.143	.787	-.880	.021*	.406	.425	-.058	.913	.270	.558	.250	.589	.711	.048
Phys. Frustration					.393	.441	.393	.441	-.655	.158	-.655	.158					.079	.867		
Overconfident					-.131	.805	-.131	.805	.393	.441										
Slow/Hesitant	-.334	.232	-.383	.454	.145	.784	.145	.784	-.270	.604	.304	.558	-.147	.781	-.477	.279	.037	.937	-.593	.121
Serial Tasking	-.487	.134	-.516	.295	-.829	.042*	-.829	.042*	-.829	.042*	-.837	.038*	-.486	.329	-.593	.161	-.234	.613	-.554	.154
Incorrect actions	-.408	.182			.338	.512	.338	.512			-.393	.441	-.655	.158	-.709	.074	-.299	.515	.274	.512
Surprise	-.408	.182			-.828	.042*	-.828	.042*	-.393	.441	-.655	.158	.		-.612	.144	.408	.363	.577	.134
clumsy					-.393	.441	-.393	.441	-.655	.158	-.135	.798	-.778	.069	-.668	.101	-.204	.661	.252	.547
Nervous Physical					-.372	.468	-.372	.468	.393	.441	-.270	.604	-.101	.848	-.523	.229	-.178	.702	-.204	.627
Low dexterity	-.222	.316	-.177	.738	-.812	.050*	-.812	.050*	-.177	.738	.000	1.00	-.147	.781	-.667	.102			-.711	.048*
Frantic pace	.	.		..	131	.805	.131	.805	-.655	.158	-.372	.468								
Nervous voice	-.019	.484	.131	.805	.131	.805	.131	.805	.257	.623	.088	.868	-.655	.158	-.158	.735	.134	.775	.078	.854
Spatial reference	-.316	.245	-.131	.805	-.741	.092	-.741	.092	.131	.805	-.372	.468	-.833	.039*	-.670	.100			.247	.555
leans over	.267	.281	-.393	.441	.116	.827	.116	.827	.324	.531	-.034	.949	-.655	.158	-.445	.317	.204	.661	.156	.712
Affirm Before	-.412	.179	-.383	.454	.029	.957	.029	.957	-.609	.200	-.657	.156	-.213	.686	.617	.140	.255	.582	-.733	.039*
Affirm After	-.148	.376	-.338	.512	-.348	.499	-.348	.499	-.638	.173	-.135	.798	.393	.441	-.204	.661	.236	.610	.051	.904
Intermediate Behavioural Markers																				
Positive comments																			-.247	.555
Apologetic				.	-.414	.414	-.414	.41	-.270	.600							-.474	.280	.577	.134
Social	.473	.14	-.530	.28	.213	.69	.213	.69	-.395	.44	.541	.27	.000	1	-.355	.44	.128	.784	-.350	.395
Self Affirm	-.252	.29	-.290	.577	-.029	.957	-.029	.956	-.580	.227	-.412	.417	.371	.468	-.414	.355	.468	.289	.184	.662
Team Aware	.612	.072	.463	.355	.304	.558	.304	.557			.778	.068	.655	.158	.757	.049*	.612	.144	.577	.134
Team Contribute	.612	.072	.676	.140	.655	.158	.655	.158			.778	.068	.655	.158	.757	.048*	.612	.144	.577	.134
Pace Fast	.185	.345	.393	.441	.522	.288	.522	.288	-.794	.059	.463	.355	-.091	.863	-.670	.099	-.299	.511	.000	1
Pace Slow	.126	.394	.383	.454	.334	.518	.334	.517	.116	.826	-.093	.861	-.698	.123	.074	.874	.691	.086	.275	.509
Fatigue	.327	.236	-.131	.805				.	-.541	.267	-.131	.804	.185	.725	.000	1	.089	.849	-.412	.310
Adjust M M I	.867	.006*	-.270	.604	-.372	.468	-.372	.467	-.393	.441	.093	.861	-.439	.383	-.316	.489	.612	.144	.250	.550
Fidgets	-.306	.25	.213	.686	.486	.329	.486	.328	-.638	.173	-.059	.911	-.257	.622	.335	.462	-.054	.908	.192	.650
Play/Sandpit																				
Periphery	.756	.025*	.714	.111	.829	.042*	.829	.042	.030	.954	.943	.005*	.829	.042*	.691	.085	.727	.064	.712	.048
Expert Behavioural Markers																				
Dual Tasking	.685	.04*	.820	.040*	.754	.084	.754	.084	.203	.699	.698	.123	.551	.257	.360	.427	.821	.023	.843	.009
Automatic/Quick	.482	.137	.829	.042	.928	.008	.928	.008	-.154	.770	.820	.046	.886	.019	.299	.514	.593	.161	.709	.049
Plans Ahead	.955	.000*	.667	.148	.899	.015	.899	.015	.395	.439	.771	.072	.736	.096	.667	.102	.667	.102	.783	.022
Maintains Global SA	.857	.000*	.812	.049	.600	.208	.600	.208	.086	.872	.679	.138	.812	.049	-.036	.938	.714	.071	.921	.001
Quirks	.204	.330	-.393	.441					-.135	.798			.759	.080			-.204	.661		
Dual Tasking	.685	.04*	.820	.040*	.754	.084	.754	.084	.203	.699	.698	.123	.551	.257	.360	.427	.821	.023	.843	.009

Table D14 - Individual Trainee ATCO Spearman's Rho Correlations for Individual Behavioural Markers against simulation exposure (* significance at p≤.05)

D10 - NON-BENIGN VS BENIGN RESULTS

Behavioural Markers		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Neg. Comments	Equal variances assumed	4.981	.038	2.201	19	.040	.84585	.38423	.04165	1.65006
	Equal variances not assumed			2.110	10.377	.060	.84585	.40089	-.04301	1.73471
Pos. Comments	Equal variances assumed	3.241	.088	.823	19	.421	.13373	.16243	-.20624	.47370
	Equal variances not assumed			.804	14.470	.434	.13373	.16630	-.22187	.48933
Apologetic	Equal variances assumed	4.468	.048	-.951	19	.353	-.08264	.08689	-.26450	.09921
	Equal variances not assumed			-1.000	10.000	.341	-.08264	.08264	-.26679	.10150
Social	Equal variances assumed	.747	.398	2.031	19	.056	1.46274	.72011	-.04447	2.96994
	Equal variances not assumed			2.003	16.683	.062	1.46274	.73015	-.07998	3.00545
Confusion	Equal variances assumed	11.570	.003	-1.377	19	.185	-.21251	.15433	-.53553	.11050
	Equal variances not assumed			-1.448	10.000	.178	-.21251	.14680	-.53960	.11457
Self Affirm	Equal variances assumed	1.665	.212	-.384	19	.706	-.14610	.38084	-.94321	.65102
	Equal variances not assumed			-.396	14.571	.698	-.14610	.36897	-.93456	.64237
Delays Repeats	Equal variances assumed	6.180	.022	-1.364	19	.188	-.35930	.26336	-.91052	.19192
	Equal variances not assumed			-1.409	14.362	.180	-.35930	.25495	-.90481	.18622
Irritation Frustration	Equal variances assumed	1.573	.225	1.937	19	.068	1.02463	.52906	-.08270	2.13197
	Equal variances not assumed			1.916	17.345	.072	1.02463	.53468	-.10174	2.15101
General Queries	Equal variances assumed	13.603	.002	1.571	19	.133	.37799	.24060	-.12559	.88157
	Equal variances not assumed			1.504	10.110	.163	.37799	.25139	-.18132	.93730
Reactionary	Equal variances assumed	51.400	.000	2.063	19	.053	.31637	.15339	-.00468	.63743
	Equal variances not assumed			1.961	9.000	.081	.31637	.16130	-.04852	.68127
Team Short%Snappy	Equal variances assumed	5.598	.029	1.052	19	.306	.10526	.10010	-.10425	.31477
	Equal variances not assumed			1.000	9.000	.343	.10526	.10526	-.13286	.34338
Pace Slow	Equal variances assumed	5.344	.032	.888	19	.385	.48914	.55057	-.66323	1.64150
	Equal variances not assumed			.856	11.520	.409	.48914	.57116	-.76109	1.73937
Phys. Frustration	Equal variances assumed	1.568	.226	.687	19	.500	.19574	.28489	-.40055	.79203
	Equal variances not assumed			.678	16.850	.507	.19574	.28864	-.41364	.80513
Over Confident	Equal variances assumed	4.468	.048	-.951	19	.353	-.12121	.12743	-.38793	.14551
	Equal variances not assumed			-1.000	10.000	.341	-.12121	.12121	-.39129	.14887
Slow Hesitant	Equal variances assumed	10.160	.005	-1.291	19	.212	-.37741	.29238	-.98936	.23454
	Equal variances not assumed			-1.338	13.503	.203	-.37741	.28211	-.98458	.22976
Play Sandpit	Equal variances assumed	.023	.881	-.061	19	.952	-.01010	.16554	-.35658	.33638
	Equal variances not assumed			-.061	18.978	.952	-.01010	.16443	-.35429	.33409
Incorrect Actions	Equal variances assumed	4.468	.048	-.951	19	.353	-.07273	.07646	-.23276	.08730
	Equal variances not assumed			-1.000	10.000	.341	-.07273	.07273	-.23477	.08932
Surprise	Equal variances assumed	3.062	.096	-.786	19	.442	-.50606	.64422	-1.85442	.84230
	Equal variances not assumed			-.824	10.543	.428	-.50606	.61426	-1.86521	.85309
Quirks	Equal variances assumed	28.681	.000	1.812	19	.086	1.27837	.70562	-.19852	2.75525
	Equal variances not assumed			1.741	10.892	.110	1.27837	.73427	-.33969	2.89643
Team Aware	Equal variances assumed	1.729	.204	10.993	19	.000*	4.25149	.38676	3.44199	5.06099
	Equal variances not assumed			10.675	13.133	.000	4.25149	.39827	3.39196	5.11102
Team Contribute	Equal variances assumed	5.017	.037	9.221	19	.000	3.83236	.41563	2.96244	4.70229
	Equal variances not assumed			8.959	13.270	.000*	3.83236	.42775	2.91017	4.75456
Quickly Locates	Equal variances assumed	5.638	.028	4.453	19	.000	1.36797	.30718	.72504	2.01091
	Equal variances not assumed			4.629	12.730	.000*	1.36797	.29552	.72817	2.00777
Periphery	Equal variances assumed	20.467	.000	4.766	19	.000	1.36183	.28571	.76383	1.95983
	Equal variances not assumed			4.553	9.771	.001*	1.36183	.29909	.69330	2.03036
Plans Ahead	Equal variances assumed	.752	.397	3.340	19	.003*	2.25404	.67495	.84136	3.66673
	Equal variances not assumed			3.384	18.492	.003	2.25404	.66611	.85726	3.65083
Global SA	Equal variances assumed	4.301	.052	9.821	19	.000*	3.72820	.37961	2.93366	4.52273
	Equal variances not assumed			9.977	18.134	.000	3.72820	.37368	2.94354	4.51285
Dual Tasking	Equal variances assumed	6.163	.023	2.196	19	.041	2.20241	1.00280	.10353	4.30129
	Equal variances not assumed			2.306	10.239	.043*	2.20241	.95488	.08152	4.32330
Serial Tasking	Equal variances assumed	3.803	.066	2.227	19	.038*	.65454	.29394	.03931	1.26977
	Equal variances not assumed			2.188	15.888	.044	.65454	.29912	.02006	1.28902
Fatigue	Equal variances assumed	24.154	.000	5.007	19	.000	3.31722	.66251	1.93058	4.70387
	Equal variances not assumed			4.786	9.883	.001*	3.31722	.69309	1.77043	4.86402
Adjusts MMI	Equal variances assumed	43.051	.000	4.533	19	.000	2.78814	.61511	1.50070	4.07559
	Equal variances not assumed			4.331	9.828	.002*	2.78814	.64370	1.35048	4.22581
Fidgets	Equal variances assumed	7.213	.015	3.639	19	.002	2.76546	.75995	1.17486	4.35605
	Equal variances not assumed			3.559	14.812	.003*	2.76546	.77695	1.10759	4.42332
Tech Discussion	Equal variances assumed	10.676	.004	4.150	19	.001	3.14841	.75874	1.56034	4.73648
	Equal variances not assumed			4.017	12.425	.002*	3.14841	.78380	1.44710	4.84971
Pace Fast	Equal variances assumed	44.128	.000	4.040	19	.001	2.71520	.67207	1.30855	4.12186
	Equal variances not assumed			3.842	9.000	.004*	2.71520	.70674	1.11646	4.31395

Table D15 - T-test Results, Non-Benign vs. Benign Data (* significance at $p \leq .05$)

D11 - NOTICE TO PARTICIPANTS

My name is David Thompson I work in the HF team over at CTC. As part of my time at NATS I am undertaking an Engineering Doctorate at UCL. My research concerns the observation of controllers when working. Specifically I am exploring whether certain overt behaviours may be used to determine how users are engaging with the systems they are using.

I am interested in how behaviours change over time, as users become more familiar with a system. The intended output of my research is a set of behavioural markers which may be used to assess how well a controller is engaging with a system (particularly new systems such as electronic strips as they are introduced into service).

The behaviours I am interested in fall into five categories. As a HF researcher I am focusing on non-technical skills, overt mood, and interaction with others. The five categories are:

Style of Input and interaction with the HMI/MMI
Interaction with others
Physical Posture and Body Language
Attitude and Mood
Communications and Verbal Commentary

With agreement from the college, I am planning to undertake a series of discrete observations during simulator training for the 226 ADI course. These observations will be non intrusive, will not be assessing performance, any data collected are made anonymous, and remain confidential used solely for the purposes of my research (I will be sharing my research findings with the college).

I would hope to spend 20-30 minutes per observation session, watching trainee controller's actions and recording each instance with a tally count when certain behaviour occurs. If you have any questions or would like further information please do not hesitate in contacting me:

David 3. Thompson

Human Factors, Directorate of Safety

Figure D12 - Trainee ATCO Study - Notice to Participants