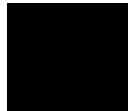


DOC 97-70-10

PROGRAMME FOR
HARMONISED AIR TRAFFIC
MANAGEMENT RESEARCH
IN EUROCONTROL



EUROPEAN ORGANISATION FOR THE SAFETY OF AIR NAVIGATION, EUROCONTROL



PD/3 IOCP - NATS

PD/1+ Trial Report

PHARE/CAA/PD3-5.2.8.4/SSR;1



EUROCONTROL

EUROCONTROL
96 rue de la Fusée
B-1130 BRUXELLES

Prepared by: F. Rainback
D. Hudson
A. Lucas
Date: April 1997

The information contained in this report is the property of the PHARE Participants*. The report or parts thereof may be published and or reproduced on the condition that due acknowledgement of authorship is made by quoting the copyright statement below. The copyright and the foregoing condition on publication and reproduction shall extend to all media in which the information may be embodied.

The information contained in this document is provided on an "as-is" basis and the PHARE Participants shall provide no express or implied warranty of any kind and shall accept no liability whatsoever for or in connection with the use of the information contained in the document.

* The PHARE Participants are:

- the EUROCONTROL Agency;
- the CENA (Centre d'études de la navigation aérienne);
- the STNA (Service technique de la navigation aérienne);
- the NLR (Nationaal Lucht- en Ruimtevaartlaboratorium);
- the RLD (Rijksluchtvaartdienst);
- the LVB (Luchtverkeersbeveiliging);
- the DLR (Deutsche Forschungsanstalt für Luft- und Raumfahrt);
- the DFS (Deutsche Flugsicherung GmbH);
- the UK CAA (Civil Aviation Authority);
- the NATS (National Air Traffic Services);
- the DRA (Defence Research Agency)

Copyright statement:

The copyright in this report vests in the European Organisation for the Safety of Air Navigation (EUROCONTROL); the CENA (Centre d'études de la navigation aérienne); the STNA (Service technique de la navigation aérienne); the NLR (Nationaal Lucht- en Ruimtevaartlaboratorium); the RLD (Rijksluchtvaartdienst); the LVB (Luchtverkeersbeveiliging); the DLR (Deutsche Forschungsanstalt für Luft- und Raumfahrt); the DFS (Deutsche Flugsicherung GmbH); the UK CAA (Civil Aviation Authority); the NATS (National Air Traffic Services) and the DRA (Defence Research Agency).

All rights reserved.

REVISION HISTORY

April 1997 Draft version 1
May 1997 Issue 1 (Including comments from PD3CG members)

	NAME	SIGNATURE	DATE
AUTHORS	F S Rainback NATS PD/3 IOCP Project Leader Douglas Hudson NATS PD/3 IOCP Project Analyst Angela Lucas NATS PD/3 IOCP GHMI Team Leader		
APPROVED	R A Whitaker NATS PD/3 Project Leader		
APPROVED	M Bisiaux PD/3 Project Leader		

EXECUTIVE SUMMARY

The aim of the NATS IOCP was to re-examine further the outcome of the PD/1 trial, in particular addressing tool utilisation and controller workload. Specifically, the project was to use the PUMA workload assessment toolset to identify potential improvements to the PHARE advanced tools, ground human machine interface and procedures which should, if implemented, reduce controller workload. These changes were evaluated through the 'PD/1+ trial', a real-time simulation on the NATS Research Facility.

The PD/1+ trial took place in January and February 1997 at the Air Traffic Management Development Centre, Hurn. The airspace simulated was based on the two UK New En Route Centre sectors 10 and 11/33, each sector being staffed by one planning controller and one tactical controller. Eight London Area and Terminal Control Centre controllers participated in the trial. This report describes the background to the PD/1+ trial, its aims, the design, results, conclusions and recommendations.

The main aim of the PD/1+ trial was to compare workload and tool use between PD/1 and a modified PD/1 system, following the implementation of a number of tool, interface and procedural changes identified through the use of the PUMA toolset. This was achieved through the collection and subsequent analysis of both subjective and objective data in accordance with a pre-defined analysis plan.

In summary, the conclusions are that the modifications reduced overall workload compared to the baseline of PD/1. The changes made to the Highly Interactive Problem Solver no-go zones and the Communications List Window made the biggest contribution to this reduction, however increased familiarity with the system as the trial progressed also played a part.

There are a number of detailed recommendations made with respect to each of the tool and interface changes. However the main recommendation for PHARE Demonstration 3 and any future PHARE trials was that realistic traffic samples, an operational concept and airspace structure must be developed which will maximise the potential of the system under evaluation and gain participant acceptance.

LIST OF CONTENTS

1.	INTRODUCTION	7
1.1	SCOPE	7
1.2	CONTEXT	7
1.3	REPORT STRUCTURE	8
2.	BACKGROUND	8
2.1	WP0: PD/1 VERSUS OTHER EN ROUTE SCENARIOS	9
2.2	WP1, 2 AND 3: PUMA ANALYSIS	9
3.	AIMS AND OBJECTIVES OF THE PD/1 + TRIAL	10
4.	THE TRIALS ENVIRONMENT	11
4.1	THE TRIALS FACILITIES	11
4.2	PHARE ADVANCED TOOLS (PATs)	12
4.3	ADDITIONAL COMPONENTS OF PD/1+	12
4.4	AIRSPACE	15
5.	TRIAL METHOD	15
5.1	TRIAL DESIGN	15
5.2	CONTROLLERS	16
5.3	TRAFFIC SAMPLES	16
5.4	TRAINING	16
5.5	MEASUREMENTS AND DATA COLLECTION	17
5.6	STATISTICAL ANALYSIS METHODOLOGY	17
6.	TRIAL RESULTS	18
6.1	TRIAL ACCOUNT	19
6.2	CONTROLLER WORKLOAD	19
6.3	IMPACT OF CHANGES ON TOOL USAGE: ORG1 VERSUS ORG2	24
6.4	COMPARISON OF CAPACITY BETWEEN ALL THE ORGs	25
6.5	COMPARISON OF SYSTEM USABILITY BETWEEN ALL THE ORGs	26
6.6	FURTHER INVESTIGATION: LEARNING EFFECTS	28
7.	DISCUSSION	28
7.1	OVERALL IMPACT OF THE CHANGES ON CONTROLLER WORKLOAD: ORG1 VERSUS ORG2	28
7.2	COMPARISON OF CONTROLLER WORKLOAD BETWEEN THE REMAINING ORGs	29
7.3	IMPACT OF THE CHANGES TO THE NO-GO ZONES ON CONTROLLER WORKLOAD	29
7.4	IMPACT OF THE CHANGES TO THE CLW ON CONTROLLER WORKLOAD	30
7.5	IMPACT OF CHANGES TO THE HIPS SCROLLING ON CONTROLLER WORKLOAD	30
7.6	IMPACT OF CHANGES TO THE HIPS SPEED PROBLEM SOLVER ON CONTROLLER WORKLOAD	31
7.7	IMPACT OF CHANGES TO THE DEFAULT SCREEN CONFIGURATION ON CONTROLLER WORKLOAD	32
7.8	IMPACT OF CHANGES TO THE COLOURS IN THE NO-GO ZONES ON CONTROLLER WORKLOAD	32
7.9	IMPACT OF IMPROVING THE POINT-OUT FACILITY ON CONTROLLER WORKLOAD	32
7.10	IMPACT OF CHANGES TO THE TRAJECTORY SUPPORT TOOL (TST)	33
7.11	IMPACT OF CHANGES TO THE MIW/MOW	33
7.12	IMPACT OF CHANGES ON TOOL USAGE	33
7.13	COMPARISON OF CAPACITY BETWEEN THE ORGs	34
7.14	COMPARISON OF SYSTEM USABILITY BETWEEN THE ORGs	34
7.15	FURTHER INVESTIGATION: LEARNING EFFECTS	35

8.	CONCLUSIONS	35
9.	RECOMMENDATIONS	36
9.1 RECOMMENDATIONS RELATING TO THE PATS, AND GHMI CHANGES:.....		36
9.2 RECOMMENDATIONS FOR PD/3 AND FUTURE TRIALS:.....		37
10.....	GLOSSARY	38
11.....	REFERENCES	40

Appendices

- A** CHANGES MADE TO THE GHMI, PATS AND PROCEDURES
- B** THE PD/1+ TRIAL DESIGN AND ANALYSIS PLAN
- C** PD/1+ TRIAL RESULTS

1. INTRODUCTION

1.1 Scope

This report details the aims, design, results, conclusions and recommendations of the Programme for Harmonised Air Traffic Management Research in Eurocontrol (PHARE) Demonstration 1+ (PD/1+) trial. The trial was the culmination of the National Air Traffic Services (NATS) PHARE Demonstration 3 (PD/3) Internal Operational Clarification Project (IOCP). This report is self-contained. For the benefit of the PHARE partners, however, references 1, 2, 3, 4 and 5 give detailed internal supporting material concerned with the rest of the work carried out within the IOCP.

1.2 Context

Today's air traffic control (ATC) system in Europe (and elsewhere) is, at times, unable to handle the traffic demands made upon it. Flow restrictions lead to delays during peak periods. The scope for increasing further the capacity of the system through existing ATC methods and technology is limited. Although developments in airspace, routes and sectorisation must, and will be, pursued, changes in the technology and process of ATC must also be envisaged if the necessary capacity gains are to be secured. The limiting factor in much of the present ATC system is the workload of the controller. A means has to be found to help the controller handle more aircraft in a given airspace without a significant increase in workload and without compromising system safety.

One proposed method of increasing controller productivity is by providing 'computer assistance tools' to both the planning and tactical controllers and by the use of datalinks for air-to-ground communication. The provision of such automated assistance to the controllers will support them in the resolution of conflicts and in the planning of efficient use of the airspace. The introduction of datalink to communicate between the airborne systems and ground environment will remove some of the current communication load from the controller and, in addition, will enable the use of data from onboard the aircraft to improve the precision of the ground system's model of aircraft performance which is used for track prediction and conflict prediction. In providing such support, it is necessary to ensure that the tasks removed from the pilot and controller are those which are best executed by computer, and those which remain are those best executed using the flexibility and adaptability of human skills.

It is the need to address some of these issues that has shaped developments within the PHARE Programme.

A series of real-time simulations, or PHARE demonstrations, was planned to evaluate the introduction of computer assistance tools and datalink in en route and terminal movement area (TMA) airspace, each demonstration being led by one of the PHARE partners.

The first PHARE demonstration, PD/1, was hosted by NATS on the NATS Research Facility (NRF), then located at DRA Malvern but now located at the Air Traffic Management Development Centre (ATMDC), Hurn. The trial was conducted over a period of eight weeks towards the end of 1995 (reference 6). The primary aim of PD/1 was to investigate the introduction of computer assistance tools and aircraft equipped with datalink and a 4D flight management system (4D FMS) within en route airspace. The computer assistance tools were designed to assist controllers in planning, and then implementing, conflict-free trajectories through the airspace. It was envisaged that these tools, together with the datalink equipped aircraft, would reduce the controllers' workload and thus help to increase airspace capacity.

The PD/1 results did not, however, indicate the anticipated reduction in workload following the introduction of the tools and datalink. While some preliminary assessment of this result was given in the PD/1 Final Report (reference 6), it was apparent that further investigation of the PD/1 results, followed by further development of the tools, would be beneficial to both the PD/3 and overall PHARE work programmes. The NATS' IOCP within the overall PD/3 work programme was launched to address these outstanding issues from the PD/1 trial. As the culmination of the research and development described in section 2 below, the PD/1+ trial formed a major component of the NATS IOCP.

1.3 Report structure

The report is divided into the following sections and detailed Appendices.

Section 2 places the PD/1+ trial within the context of the IOCP as a whole, stating the aims and objectives of the complete project. Each of the work packages leading up to the trial is briefly discussed.

Section 3 concentrates on the specific aims and objectives of the PD/1+ trial.

Section 4 describes the trial facilities, and contains a brief description of each computer assistance tool and the associated components used during the trial.

Section 5 outlines the trial method, encompassing the design, controllers, traffic samples and training. In addition the measurements and data collected are described and the statistical methodology is stated.

Section 6 gives an account of the trial and summarises the results of the measurements made against each of the objectives of the trial.

Section 7 contains a full discussion of the results. The conclusions and recommendations are given in sections 8 and 9 respectively.

Three Appendices are attached: Appendix A details the tool, interface and procedural changes implemented for the trial; Appendix B contains a detailed account of the trial design and analysis plan; and Appendix C records the detailed results.

2. BACKGROUND

The PD/1 trial investigated the application of computer assistance tools designed to facilitate the planning and implementation of conflict-free trajectories within a scenario containing aircraft equipped with datalink and 4D FMS. The results of the PD/1 trial showed that the PHARE Advanced Tools (PATs), together with their associated ground human machine interface (GHMI), created no significant decrease in controller workload relative to the baseline scenario (see reference 6). The analysis of the PD/1 results identified two themes which appeared to be crucial in understanding exactly why the anticipated decrease in workload was not observed. These themes were:

- tool utilisation;
- workload redistribution between controllers.

The NATS IOCP was therefore designed to examine these two areas in detail, through the application of the Performance and Usability Modelling in ATM (PUMA, see section 2.2) toolset and the NRF. PUMA and the NRF are both critical elements of the assessment methodology (reference 7) of DASR, the Department of ATM Systems Research in NATS. The IOCP was therefore designed to:

- examine the PD/1 results using PUMA to identify areas of high controller workload;

- identify modifications to the PATs, GHMI, ATC procedures and NRF training methods which the PUMA analysis suggests would result in a reduction in controller workload;
- implement the suggested modifications within the PD/1 scenario on the NRF;
- examine the effect on controller workload of these modifications by conducting the PD/1+ trial.

The IOCP was split into eight work packages (WP). The preparation, execution, analysis and reporting of the PD/1+ trial were covered by WP5, WP6 and WP7 and it is this work which is described in detail within this report. A brief description of each of the preceding WPs, and hence the background to the PD/1+ trial, is given in the following paragraphs.

2.1 **WP0: PD/1 versus other en route scenarios**

The PD/1 trial used a baseline system against which the effects of introducing computer assistance tools and datalink-equipped aircraft could be measured. This baseline system was derived from the Operational Display and Input Devices (ODID) work programme - a system significantly in advance of those ATC systems currently operational or planned for near-term implementation (e.g. the UK's New En Route Centre, NERC). ODID does not use paper flight strips, relying instead on interactive track data blocks and lists displayed in the Plan View Display (PVD), and use is also made of a limited number of computer assistance tools to aid conflict detection and resolution.

WP0 of the IOCP was designed to investigate the relationship, in terms of controller workload, between the PD/1 baseline system and current day, or near-term, ATC systems. In conducting this work, use was made of previous real-time simulation studies conducted under NATS' Computer Assistance for En Route (CAER) work programme.

The full results of WP0 are presented in a PHARE internal working paper (see reference 1). The overall conclusion was that incorporation of a baseline system representative of current day, or near-term, ATC systems - as typified by the NERC Reference System used in CAER - would, when compared to the ODID derived baseline system for PD/1, produce a similar workload level for the planning controller but a slightly higher workload level for the tactical controller. Therefore, the baseline system as used in PD/1 may continue to be used in PD/3 and may be considered as being representative, in workload terms, of near-term ATC systems.

2.2 **WP1, 2 and 3: PUMA Analysis**

The PUMA toolset was developed for NATS by Roke Manor Research Limited. Using task analysis, it enables the rapid determination of controller workload given a specific operational concept and a specific ATC scenario. It allows researchers to carry out initial investigations to determine those areas of an operational concept which are particularly workload intensive, and it quantifies the effect on workload of any changes to the operational concept or scenario.

The PUMA analysis was divided into three work packages:

- WP1: to identify areas of high workload experienced by the controllers during the PD/1 trials;
- WP2: to identify modifications to the PATs and GHMI which would result in a reduction in those areas of high controller workload;
- WP3: to assess how division of workload between the planning and tactical controllers might better be organised.

This work was conducted using the standard PUMA analysis methodology as described in the PUMA User Guide (reference 8).

A typical area of high controller workload identified during WP1 was, for the tactical controller, frequent overlap of radio telephony (R/T) communications and discussions with the planning controller. A full description of this work package and the results obtained are given in a PHARE internal working paper (reference 2).

WP2 examined the results obtained from WP1 and identified 14 possible changes to the PATs and GHMI which were predicted to help reduce the controller workload. It was not possible to implement all 14 changes within the time available, therefore 9 modifications were chosen based on the expected reduction in workload predicted by PUMA and on the time taken to implement those changes on the NRF.

Modifications were made to:

- the scrolling of the Highly Interactive Problem Solver (HIPS) Altitude Problem Solver (APS) and Speed Problem Solver (SPS);
- the HIPS Speed Problem Solver for aircraft without datalink;
- the colour of the HIPS no-go zones;
- the HIPS danger and no-go zones;
- the Communication List Window (CLW);
- the default window arrangement;
- the point-out facility;
- the Trajectory Support Tool (TST);
- the Message In and Message Out Windows (MIW/MOW).

The PUMA analysis conducted under WP2 is given in reference 3 and a full description of the modifications to the PATs and GHMI is given in Appendix A. These modifications were incorporated into the NRF under WP4.

WP3 assessed the effectiveness of the planner/tactical team and examined whether the division of workload between the two controllers might be better organised. The primary aim of this analysis was to prevent the overlapping of ATC tasks which the PUMA analysis identified as leading to high workload. Three primary ATC procedures were identified for special attention during PD/1+:

- removing the need for discussions between controllers which adversely affected other, more important, tasks;
- the proper co-ordination of R/T and GHMI tasks;
- the completion of specific tasks before starting another task.

Reference 4 contains the detailed PUMA analysis and results from WP3; Appendix A describes those ATC procedure changes assessed during PD/1+.

3. AIMS AND OBJECTIVES OF THE PD/1 + TRIAL

Following the PUMA analysis and the development of the PATs, GHMI and ATC procedure changes, the following specific aim was defined for the PD/1+ trial:

- to investigate the effect on controller workload of the modification to the computer assistance tools (PATs), controller screen display (GHMI) and ATC procedures.

As stated in section 1, the primary aim of PHARE is to examine methods of increasing airspace capacity in order to meet the expected increase in traffic demand over the next 20 years. Within en route airspace, the limiting factor in increasing sector capacity is known to be controller workload. If the workload could be reduced, then the sector capacity should be capable of being increased; therefore, any reduction in controller workload due to the implementation of the PATs, GHMI and ATC procedure modifications should be capable of being translated into an increase in sector capacity.

In addition to the specific aim described above for the PD/1+ trial as a whole, more detailed aims were required in order to analyse specific reasons for any increase or decrease in controller workload, so that only the productive modifications are recommended for incorporation into PD/3. Eighteen such detailed aims, termed herein as low-level objectives, were identified for the PD/1+ trial and are fully described in Appendix B (see section 5.6 for an example).

4. THE TRIALS ENVIRONMENT

This section describes the hardware and software used in the trial.

4.1 The trials facilities

The PD/1+ trial took place on the NRF, a real-time simulator, based at the NATS ATMDC, Hum. See Figure 1.



Figure 1 - The NATS Research Facility

The software and hardware components of the NRF used during the trial were as follows:

NRF Software

- the PATs (see section 4.2);
- a GHMI as used in the PD/1 trial;
- a modified GHMI incorporating the changes made for the PD/1+ trial;
- GENESIS, a 4-D aircraft track generator with Experimental Flight Management System (EFMS) capabilities;
- a surveillance and tracking system;
- a datalink interface;

- a flight plan processor;
- electronic inter-sector co-ordination support;
- supporting databases.

NRF Hardware

- a network of Sony workstations, providing the controllers with an interface to the system;
- two DEC Alphas running GENESIS;
- a cluster of Sun workstations, supporting the simulation and basic software functions;
- Sun workstations providing an interface into GENESIS for the pseudo-pilots;
- an Ericsson phone system providing simulated R/T and telephone communication;
- video recording facilities.

4.2 PHARE Advanced Tools (PATs)

The tools used during the PD/1+ trial were as follows:

- the **Trajectory Predictor** generated trajectories for sets of strategic constraints, which were either prepared by the controller or pre-planned within the traffic sample file;
- the **Conflict Probe** compared all new active trajectories against the active trajectories of all other aircraft for potential conflicts;
- the **Flight Path Monitor** monitored all active aircraft for deviations from their agreed trajectories;
- the **Highly Interactive Problem Solver (HIPS)** provided a graphical interface, integrated into the GHMI, which allowed a controller to identify potential conflicts and then to plan a set of strategic constraints for an aircraft passing through their sector which would avoid those conflicts.

4.3 Additional components of PD/1+

The following additional components of the system used in the PD/1+ trial drew information from the tools described in section 4.2:

- the **Plan View Display (PVD)** provided the processed radar picture containing maps, airways, reporting points, aircraft tracks and labels. The PVD also contained the Radar Tool Box (RTB) and Sector Inbound Lists (SILs);
- a **Track Data Block (TDB)** was attached to each aircraft target via a line on the display, and could be re-positioned relative to the respective aircraft. It contained information such as the aircraft callsign, current flight level, and cleared flight level. The information displayed and the colour of the TDB was dependent upon the aircraft state;
- the **Sector Inbound Lists (SILs)**, within the PVD, provided entry details and planning information for aircraft scheduled to come into the controlled sector of the GHMI;

- the **Message In/Out Windows (MIW/MOW)** displayed text messages to the controller which were received from the neighbouring sectors or were system messages (MIW), or were sent to the neighbouring sectors (MOW);
- the **Extended Label Window** provided additional information on an aircraft, not available within the TDB, such as requested level and origin and destination airfields;
- the **Radar Tool Box (RTB)** allowed the controller to change the display characteristics of the PVD;
- the **Preferences Tool (PRT)** enabled the controller to save and re-load a preferred display configuration;
- the **Augmented Dynamic Flight Leg (ADFL)** provided a two-dimensional representation of an aircraft's trajectory within the PVD, with which the controller could interact to alter constraint points or register a trajectory;
- the **Conflict Risk Display (CRD)** provided a display window for potential conflicts, giving the time to loss of separation and the separation at closest approach;
- the **Conflict Zoom Window** provided a window giving a snapshot view of a conflict in the CRD at the time of closest approach;
- the **Horizontal and Vertical Assistance Windows (HAW/VAW)** provided a plan view or vertical cross-section respectively of the trajectory of the selected aircraft and illustrated any conflicts involving that aircraft;
- the **Communications List Window (CLW)** provided a list of the R/T commands to be given to the pilots of aircraft without datalink;
- the **Trajectory Support Tool (TST)** enabled the controller to validate, reset, propose, accept or register trajectories;
- the **General Tool Box (GTB)** allowed the controller to turn on or off, as appropriate, the main set of tools, such as the CRD, HAW, VAW and the PRT.

An example of the GHMI before modification is shown in Figure 2.

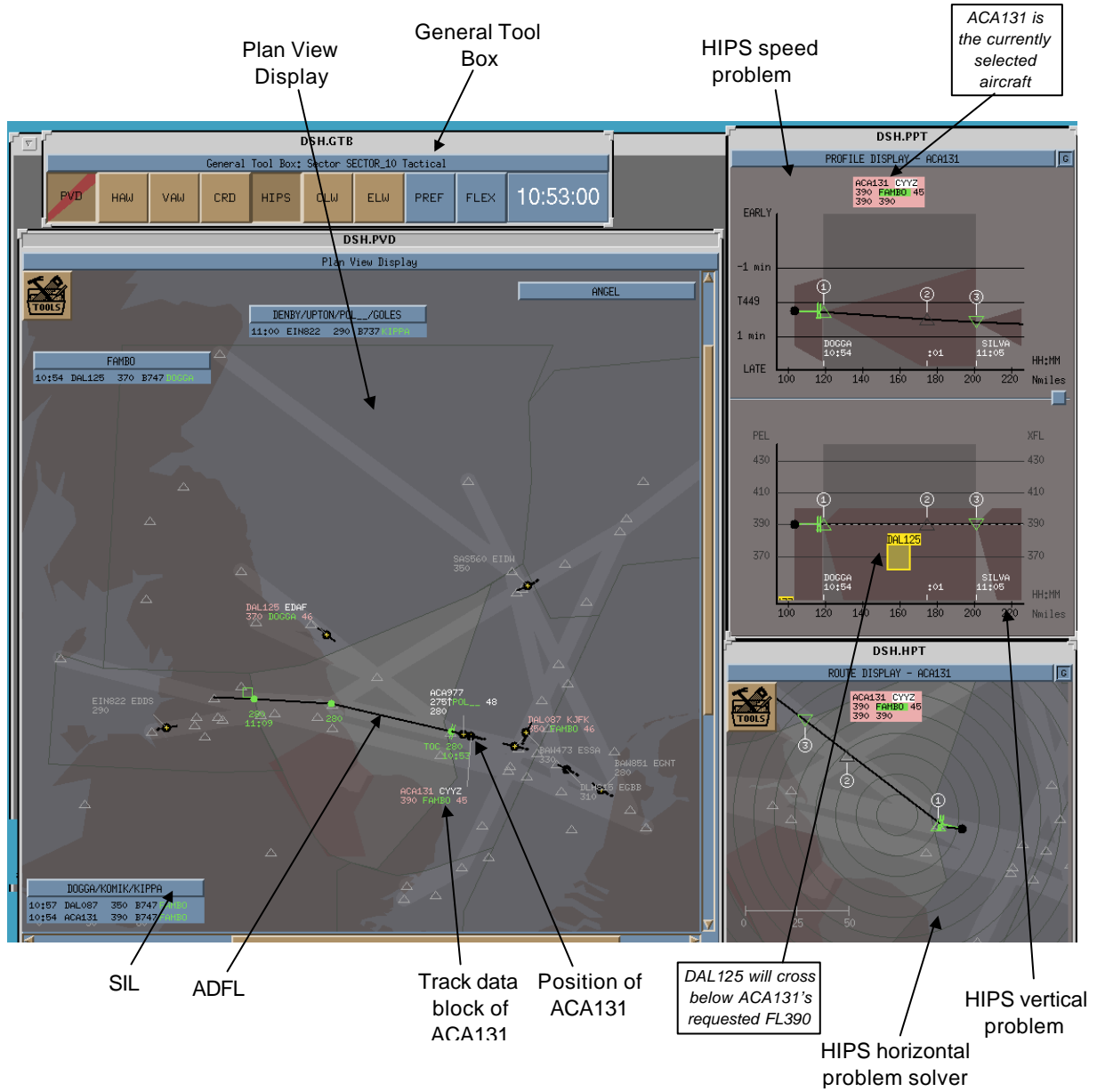


Figure 2 - GHMI (before modification) incorporating the HIPS

4.4 Airspace

The NRF was configured to simulate three NERC North Sea sectors, 10, 11 and 33. Sectors 11 and 33 were banded for the purposes of the trial. Figure 3 shows the airspace and sectors simulated. Each sector had one tactical controller and one planning controller. Sectors 10 and 11/33 are 'fed' by two adjoining feed sectors: one encompassing London (SW), Manchester and Scottish, and the other encompassing London (S), Maastricht, Amsterdam and Copenhagen. It was decided to use the same airspace sectorisation as in the PD/1 trial, although subsequent to this airspace being set up for PD/1, the actual NERC sectorisation has changed.

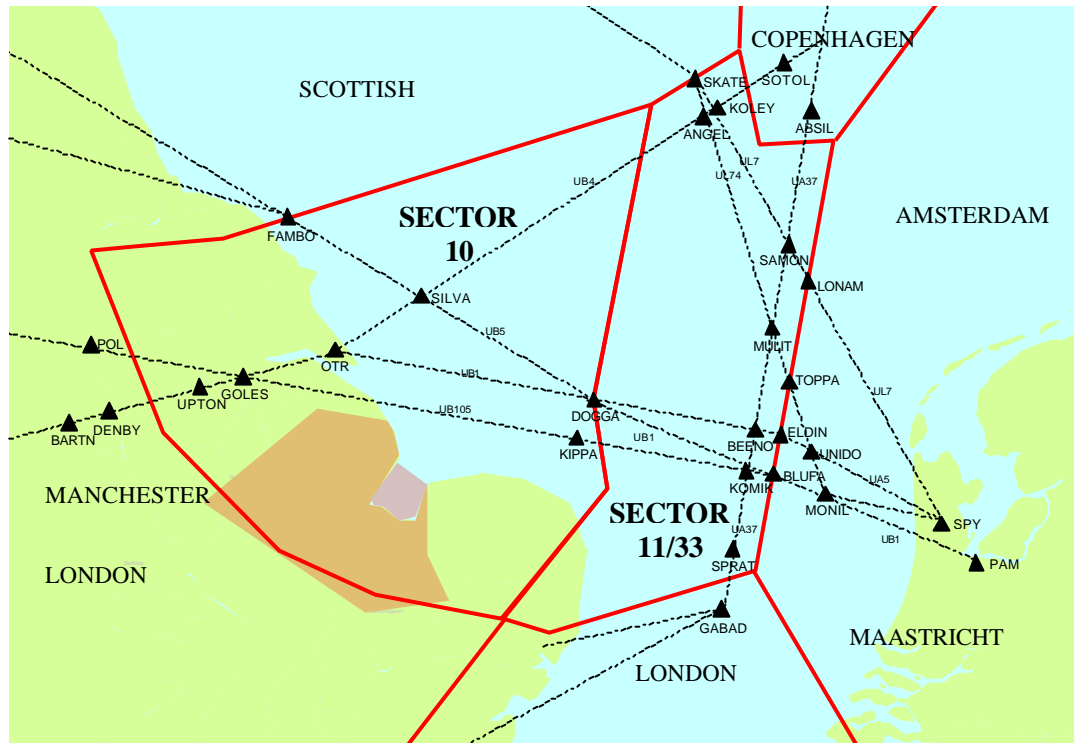


Figure 3 - Airspace for the PD/1+ Trial

5. TRIAL METHOD

5.1 Trial design

The PD/1+ trial was designed to meet the objectives of the trial (section 3), and ran for a total of four weeks from the 13th January 1997 to the 7th February 1997 inclusive. A full account of the trial design can be found in Appendix B of this report.

Four distinct “organisations” (referred to as ORGs) were developed in order to assess fully the impact of the PATs, GHMI and ATC procedure modifications. These ORGs, and the reasons for their development, are described below:

- ORG0 This is the baseline ORG and contains no PATs or datalink-equipped aircraft. The effect of introducing these elements in later ORGs may be compared against this baseline. PD/1+ ORG0 is identical to the PD/1 baseline ORG0.
- ORG1 Includes full PATs and their associated GHMI *without* any modification to the original PD/1 trial specification. 70% of the aircraft are equipped with datalink. PD/1+ ORG1 is therefore identical to the PD/1 trial’s ORG2 (70%). Hence comparison of ORG0 and ORG1 allows a contrast between

the original PD/1 results and those obtained in PD/1+. It is intended that this contrast will give confidence in the PD/1+ trial results.

ORG2 As ORG1 but *with* the modifications to the PATs, GHMI and ATC procedures described previously and detailed in Appendix A. 70% of the aircraft in ORG2 are equipped with datalink. This ORG will allow:

- an evaluation of the introduction of the computer assistance tools when compared to the baseline (ORG0);
- an evaluation of the effect of the modifications when compared to the non-modified system (ORG1).

ORG3 As ORG2 but with 100% of the aircraft equipped with datalink. The original PD/1 trial looked at the impact on workload of having 0%, 30% and 70% of aircraft with datalink. This ORG was introduced to explore further how the proportion of aircraft with datalink affected workload.

Two groups of controllers participated in the trial for a fortnight each. Each group evaluated the organisations in the order presented above, because additional training on the PATs, GHMI and especially the procedure changes had to take place between ORG1 and ORG2.

5.2 Controllers

Eight controllers in total were selected to take part in the PD/1+ trial. These comprised two teams of four controllers, who took part in the trial on separate weeks. The four controllers staffed the tactical and the planner positions of sectors 10 and 11/33. All of the controllers were from the London Area and Terminal Control Centre (LATCC), and hence held current licences. Of the eight, two were familiar with the North Sea sectors 10 and 11/33, the remaining six being more familiar with other current en route LATCC sectors. They were of mixed ages and experience. Some of the controllers had taken part in previous simulations at the ATMDC, although none were familiar with PHARE. The controllers were allocated roles and sectors randomly and remained in these positions for the complete fortnight. Sickness prevented one controller completing the planned trial programme (see section 6.1).

5.3 Traffic samples

Eight traffic samples were developed from two of the original PD/1 samples, namely s_{14_6} and s_{14_7}, each containing 70% datalink equipped aircraft. These two base samples were chosen following discussions with North Sea sector controllers at LATCC, since they best represented traffic rates and patterns experienced at LATCC during the summer of 1996. The two base samples were increased by 25%, 50% and 75%, giving the eight traffic samples for the PD/1+ trial. Variations of these samples were created to cater for the ORGs in which the number of aircraft equipped with datalink differed from 70%. Further details of the traffic samples, their throughput, characteristics and the methods used to increase the levels are included in Appendix B.

5.4 Training

Each of the controllers attended the ATMDC for a day of training prior to the PD/1+ trial, using the standalone computer-based training package developed originally for the PD/1 trial. In addition they received a number of training exercises on the full system prior to each of the ORGs. Controllers were unaware of the proposed GHMI and PATs changes in ORG2 before completing ORG1, since this could have potentially influenced their reaction to ORG1. The aim of the training was to provide the controllers with confidence and

proficiency in the use of the PD/1 system prior to the measured runs for each ORG. Full details of the training given and the samples used during the training exercises can be found in Appendix B.

5.5 Measurements and data collection

For each of the measured exercises a full set of objective and subjective data were gathered for analysis in accordance with the trial analysis plan (reference 9). The measures recorded are outlined below. The trial analysis plan and full details of the measurements taken during the trial can be found in Appendix B.

Subjective data

The subjective data, i.e. those measures representing the opinion of the controllers, were gathered from a variety of sources during the trial. The Instantaneous Self Assessment (ISA) measurement and NASA Task Load Index (TLX) questionnaires (see Appendix B) were used to record the controllers' perceived overall workload. TLX also allowed components of workload such as frustration and time pressure to be analysed. In addition, questionnaires, debriefs and controller comments were used, both to obtain controllers' opinion and, along with a specially-designed psychometric questionnaire, to carry out a usability assessment of each ORG.

Objective data

The objective data, i.e. those measures not recording the controllers' subjective opinion, were recorded automatically either by the NRF system, Ericsson communications system or the track generator. For example, 'log files' were produced containing: details of the number of interactions with a particular tool; a record of the pseudo-pilot inputs, for tracking aircraft throughput; or the duration of R/T messages transmitted. In addition, video recordings of the sector 10 controllers were made, giving a definitive record of the controllers' actions for subsequent use in PUMA analysis.

Observations

During a number of the trial and training exercises for each ORG, the performance of the controllers was observed overtly by specialist observers. The purpose of the observation was to assess, subjectively, controller body language, responses to particular events and the occurrence of certain actions. Throughout the observations any relevant comments were also noted and formally recorded as part of the post-trial analysis (see reference 10).

5.6 Statistical analysis methodology

In the analysis plan a null and alternative hypothesis were developed for each low-level objective of the trial, and the appropriate subjective and objective data required to test the hypotheses were identified. The analysis plan also described the statistical analysis to be conducted in testing the hypotheses.

A typical low-level objective was:

For each controller role, examine the impact on workload of improving the validity of the no-go zones in the HIPS

The null hypothesis for this low-level objective was:

The workload in ORG1 (i.e. without modification) associated with the HIPS not reflecting accurately the no-go zones is the same as that in ORG2 (i.e. with modification)

Alternative hypothesis:

The workload in ORG1 associated with the HIPS not reflecting accurately the no-go zones is greater than that in ORG2

For the measures relevant to each low-level objective (listed in Appendix B), descriptive statistics, usually in the form of a graph or histogram, were produced prior to any statistical analysis being carried out. This gave an indication of each measure's size and range and avoided further statistical analysis if a measure was of little descriptive value. Statistical differences could not, however, be inferred from the graphs alone. Analysis of variance was used to detect differences between the ORGs for the workload and other measures. The Friedman two-way analysis of variance was used, matching controller and traffic sample (see reference 11).

The level of significance used in each of the tests was chosen to be 5%, after consideration of the number of statistical tests planned against each low-level objective and deciding on the acceptable risk of a false positive.

As the questionnaire responses could only be gathered from a maximum of eight controllers, significance testing of the differences between controllers would not have yielded results with any degree of statistical certainty. Hence the data gathered from the questionnaires was interpreted qualitatively, along with comments from the controller debriefing sessions and the specialist observers, to support or reject the null hypotheses.

Further details of the particular statistical tests applied to each measure within each low-level objective can be found in Appendix B.

6. TRIAL RESULTS

The PD/1+ trial was designed to specifically investigate the effect on controller workload of the modifications made to the PATs, GHMI and ATC procedures. This section summarises the results from the trial, presented in full in Appendix C, against this aim. In particular the results relating to the eighteen low-level trial objectives, described fully in Appendix B, are given.

Each low level objective asked a very specific question and hence an appropriate subset of the various types of subjective data and objective data were collected for each low-level objective. As a consequence, results are presented within the following sections for the specific data collected against each objective. There is inevitable overlap between the objectives. For example, any ISA results collected to examine the difference in workload between ORG1 and ORG2 would, in part, have been influenced by each individual change to the GHMI, PATs and ATC procedures. However, it was not possible to quantify the extent to which each change contributed to any observed difference in the ISA results, and therefore ISA was not directly recorded against each individual PATs and GHMI change. Appendix B contains full details of the specific data collected against each objective.

Section 5.6 described the statistical analysis methodology applied. The following sections concentrate on those results which were statistically significant. Where appropriate, debrief comments, observations and questionnaire results are also presented; the complete results obtained from the statistical analysis are given in Appendix C.

Section 6.1 gives an account of the exercises completed during the trial.

Section 6.2 presents the results gathered in relation to the comparison of workload between each of the ORGs and, in addition, the impact each individual GHMI and PATs change had on workload. It was not possible, within the timescales available, to address the impact of each individual ATC procedure change described in Appendix A.

Section 6.3 presents the results of the comparison of tool usage between ORG1 and ORG2, consequently describing how the modifications affected the way the tools were used by the controllers.

Section 6.4 and 6.5 show how capacity and usability varied respectively between the ORGs. Section 6.6 presents the results from the exercise carried out to investigate learning effects during the trial.

The implications of all of the results are discussed in Section 7.

6.1 Trial account

A total of forty measured exercises, out of a planned maximum of forty eight, were completed between Monday, January 13 and Friday, February 7, 1997 inclusive. The forty eight planned runs consisted of two sets of twenty four runs, one per team of four controllers. System problems with seven of the forty completed exercises caused them not to last for the intended duration of one hour; however, the data collected were still used in the analysis, with account being taken of the exercises' shorter duration. All of the forty exercises included in the analysis lasted for at least 50 minutes.

Despite losing a number of the planned measured exercises, sufficient redundancy had been built in to the experimental design to allow the planned analysis to be carried out.

Controller g, the sector 11 planning controller from fortnight two of the trial, was unable to complete the ORG1 exercises due to sickness, and remained unavailable for ORG2 and ORG3. The ATMDC ATCO assisting with the trial was able to take over the position for the remainder of the fortnight, allowing the trial to continue despite this setback. The statistical analysis has excluded the measures from controller g and the ATMDC ATCO. The controller 'sample' thus comprises two controllers for each of the sector 10 tactical, sector 10 planner, and sector 11 tactical positions, and one sector 11 planner.

Comments made during the debrief sessions by one of the planner controllers, indicated that the ISA scores recorded did not accurately reflect his perceived workload. It appeared, from observation, that this controller had been selecting the nearest button ("under-utilised") to stop the flashing. The ISA scores from this controller have not, therefore, been included in the analysis; however, the same controller's TLX scores corresponded well with the debrief comments and were not excluded.

In addition to the planned exercises, it was possible to include an additional exercise at the end of the trial to investigate the effects of learning. Following completion of ORG3, an ORG1 exercise was re-run to examine whether reductions in workload were caused by increasing familiarity with the system and simulated traffic. Results from this additional exercise were not used elsewhere in the analysis.

6.2 Controller Workload

To enable the comparison of workload between the ORGs, a subset of the measures detailed in Appendix B were used in each case. The subjective measures used were ISA, TLX, questionnaire responses and debrief comments. Objective measures used were a range of data recorded automatically by the NRF system and percentage time spent by the tactical controller on R/T. In addition, the workload results were further analysed using the PUMA toolset.

Figure 4 and Figure 5 show how the controllers perceived their workload to vary between the different ORGs. Figure 4 uses overall TLX as the measure and Figure 5, mean ISA. Mean ISA is presented for consistency with the PD/1 results. The vertical bars on the

graph represent the 25th and 75th percentile range of the measure, and the centre point represents the median. It must be emphasised, that statistical differences cannot be inferred from these graphs alone, since the statistical testing used matched pairs of data, rather than characteristics of each data set as a whole.

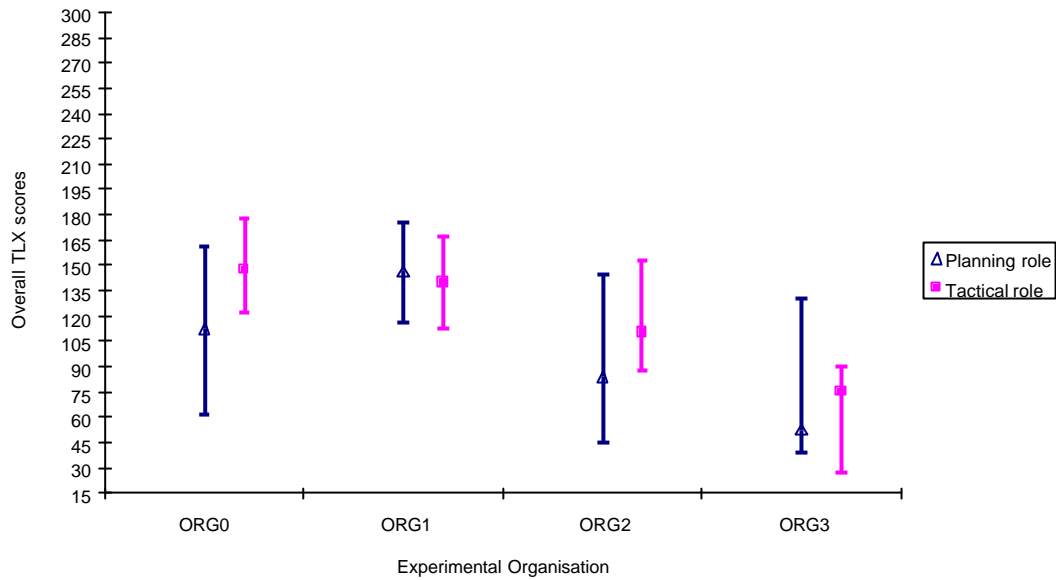


Figure 4 - Comparison of overall TLX scores

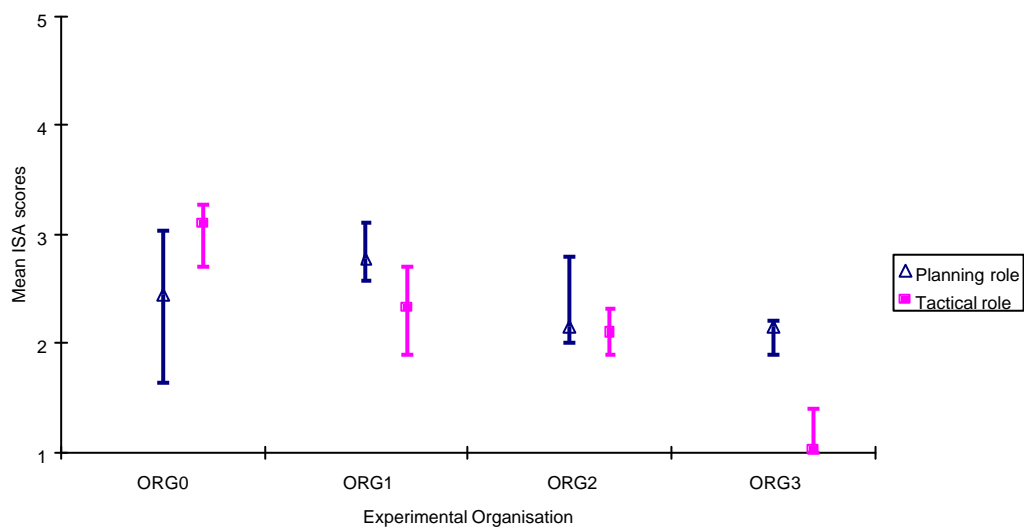


Figure 5 - Comparison of Mean ISA for each ORG

ORG1 versus ORG2

Since the primary aim of the PD/1+ trial was to assess the impact on workload of the modifications to the PATs, GHMI and ATC procedures, the most important comparison of workload was that made between ORG1 and ORG2. The impact of the individual PATs and GHMI changes, which will have influenced these overall results, is discussed in Section 7.

Overall, ISA responses in ORG2 were significantly lower than in ORG1 for both the tactical and planning roles.

The time pressure felt by both the planning controllers and tactical controllers, as measured by TLX, was low in both ORG1 and in ORG2. The differences between ORG1 and ORG2 were not significant. Both the planning and tactical controllers experienced high levels of frustration in ORG1, which was reduced significantly in ORG2 for the tactical controllers, but not the planning controllers.

The questionnaire results suggested that the planning controllers had perceived a reduction in workload from ORG1 to ORG2, as a result of the modifications made to the PATs in ORG2. The tactical controllers were however split in their opinion. However, both the planning and tactical controllers considered there to have been no change in their workload between the two ORGs as a result of the GHMI changes. The biggest impact on the planners' workload was felt to have been the changes made to the HIPS no-go zones, whilst on the tactical controllers' workload it was the removal of unnecessary CLW messages.

Automatic recordings made by the NRF system enabled both the number of release alerts per exercise, and the time taken by the planning controllers to register each aircraft, to be analysed. Concise definitions of both these recordings are contained within Appendix B. The number of release alerts per exercise was significantly lower in ORG2 than in ORG1, on average decreasing from 8 to 5. Planning controllers took significantly less time in ORG2 to the first register of each inbound aircraft than in ORG1, dealing with aircraft almost 27% more quickly on average.

There was no evidence of a reduction in the percentage of time spent using R/T.

Analysis of the PUMA workload assessment tool (WAT) output confirmed that there was a decrease in workload from ORG1 to ORG2 for both controller roles, particularly the planning role. During the video analysis it was observed that modifying aircraft trajectories took the planning controller far less time. Task analysis of the planning role also demonstrated a reduction in the amount of verbal communication. Figure 6 and Figure 7 show the PUMA WAT output for the planning controller. The x-axis represents time and the y-axis represents "W/INDEX units", a comparative measure of workload. The greater plot density and higher peaks in ORG1 is representative of greater workload than in ORG2.

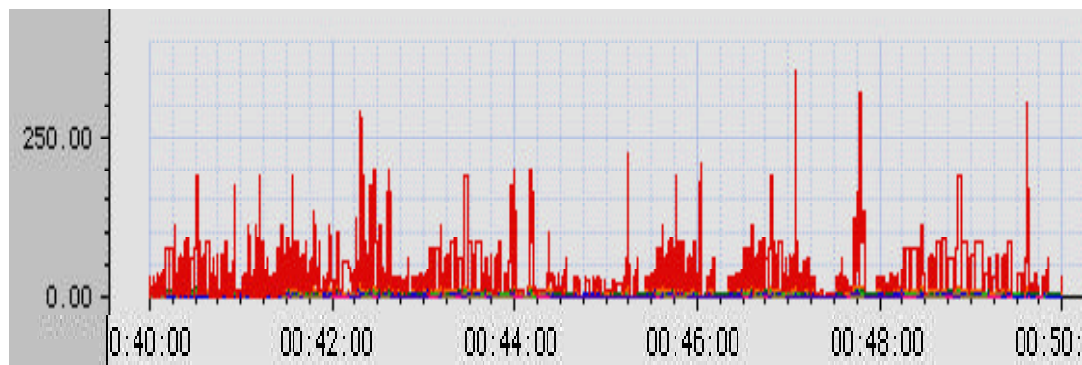


Figure 6 - PUMA WAT output - ORG1, planning controller

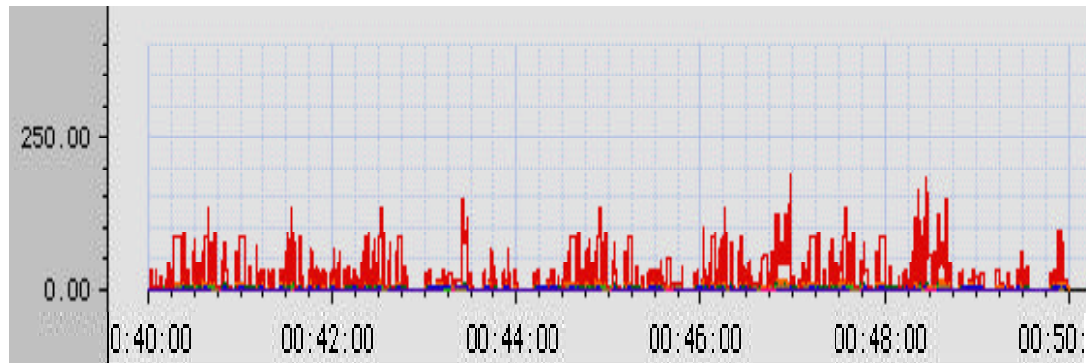


Figure 7 - PUMA WAT output - ORG2, planning controller

ORG0 versus ORG1

Comparing ORG0 and ORG1 allows a contrast to be made between the original PD/1 results and those obtained in PD/1+.

For the planning controllers, there was a significant increase in both ISA and overall TLX in ORG1 compared with ORG0. The TLX frustration scores from the planning controllers were high in ORG1, with 50% of the scores being 17 or greater (on a scale of 1 to 20). For the tactical controllers, there was a slight but significant decrease in ISA, indicating lower workload in ORG1.

Questionnaire responses indicated that, for both the planning and tactical controllers, there was mixed opinion as to whether the introduction of PATs and their associated interface in ORG1 had reduced controller workload. However, the tactical controllers indicated unanimously that they considered it was safer to control aircraft in ORG0 than in ORG1.

Percentage time spent by the tactical controllers in R/T decreased significantly, from 22% to 18% on average, in ORG1 compared with ORG0, but the number of release alerts increased significantly, from 3 to 8 per exercise on average.

The corresponding results from the PD/1 trial agreed with PD/1+ for the planning role: an increase in workload from PD/1 ORG0 to PD/1 ORG2 (70% - equivalent to PD/1+ ORG1). For the tactical controllers, however, the PD/1 trial results exhibited no significant change in workload between these ORGs.

ORG0 versus ORG2

Comparing ORG0 with ORG2 allows the evaluation of the introduction of computer assistance tools when compared to the baseline ORG0.

For the tactical controller, ISA responses in ORG2 were significantly lower than in ORG0. For the planning controller, however, there was no evidence of a difference in workload between these ORGs. The results from TLX indicated no significant differences between ORG0 and ORG2 for either controller. The questionnaire responses from the majority of tactical controllers indicated that they perceived their workload to have been less in ORG2 than in ORG0.

There was no statistically significant difference between the ORGs in the number of release alerts or in the amount that R/T was used.

ORG2 versus ORG3

ORG3 was introduced to explore further how the proportion of aircraft with datalink affected workload; the ORG3 scenario contained 100% of the aircraft with datalink compared to ORG2 with 70% datalinked.

For the planning controllers, there was a slight but statistically significant decrease in ISA, going from ORG2 to ORG3, but no significant difference in the TLX results between the two ORGs.

For the tactical controllers, the ISA responses and the overall TLX scores in ORG3 were significantly lower than those in ORG2.

In ORG3 the tactical controllers spent only 14% of their time using the R/T, on average, compared to 17% in ORG2; a significant difference. There was no difference in the number of release alerts.

Questionnaire and debrief comments by the tactical controllers, suggested that they felt that with all aircraft equipped with datalink they had nothing left to do. One controller said “The use of HIPS by the planning controller took away my need to plan, resulting in a reduction in workload. However this also took away my ability to keep a mental picture, so time saved was used trying to generate this picture.” This suggests that tactical controllers were possibly under-loaded in ORG3.

ORG0 versus ORG3

It was desirable to test explicitly whether the overall benefit from the changes and the move to purely datalink equipped aircraft, ORG3, had reduced workload compared to the baseline, ORG0.

Neither the ISA responses nor the overall TLX scores of the planning controllers differed significantly between ORG0 and ORG3. For the tactical controllers, the ISA responses and the overall TLX scores in ORG3 were significantly lower than those in ORG0.

The majority of the tactical controllers indicated, through their questionnaire responses, that their workload was significantly reduced in ORG3. The opinion of the planning controllers was split, however.

In ORG3 the tactical controllers spent only 14% of their time using the R/T, on average, compared to 22% in ORG0; a significant difference. There was no difference in the number of release alerts.

Impact of individual changes on controller workload: ORG1 versus ORG2

The assessment of the impact on workload of the individual changes is based largely on the views of controllers expressed through questionnaires, comment sheets and debriefs, and is not subject to statistical significance testing.

In their questionnaire responses, all three planning controllers and one of the tactical controllers indicated that the biggest improvement of ORG2 over the ORG1 system was the smaller no-go zones. They all said that this change to the system helped reduce the workload and frustration that they had experienced in ORG1. Although the workload of the tactical controllers was not directly affected by this particular GHMI modification, they all indicated that the increase in accuracy in the size of the no-go zones was an improvement.

In their questionnaire responses, the tactical controllers indicated that the biggest improvement of ORG2 over the ORG1 system was the removal of the unnecessary messages from the CLW, but noted that there were still a few redundant messages and messages missing. They liked the clearer text font used in ORG2. All four tactical controllers said that these two changes helped reduce the workload and frustration that they had felt in ORG1. During ORG1 they were observed to complain about the redundant and unintelligible messages in the CLW.

The controllers found that the modification to the scrolling in the HIPS Altitude Problem Solver (APS) and speed problem solver (SPS) took some getting used to. In particular, they found it scrolled too quickly when executing a large change in flight levels, making it difficult for them to capture the level they required. Opinion was split among the controllers as to whether the scrolling had been improved. Although one controller said that it actually increased workload, three controllers found that its impact was to decrease their workload.

Accidental attempts to interact with the HIPS SPS for aircraft without datalink were very rare, in fact a total of six occasions in the ORG1 measured exercises. Some of the controllers did not understand the SPS, and one controller pair reported that time constraints generated in the SPS for datalink equipped aircraft were not being carried out, and that they therefore gave up trying to use it. The 'greying out' of the display for aircraft without datalink in ORG2 did meet with general approval, as the controllers found it reinforced the distinction between aircraft with and without datalink. Despite this, the general view was that the greying out had no impact on their workload.

All of the controllers found their own arrangement of windows to be superior to the default configuration presented in ORG1. One planning controller said their own arrangement led to more economical mouse movements, but it was not feasible to measure this quantitatively.

The changes made to the colours of the no-go zones in HIPS in ORG2 met with general approval, but only one controller felt that frustration had decreased as a result. Overall, the opinion was that the change was superficial. Similarly, the controllers also found the change to the point-out facility to be superficial.

Opinion was split on the changes to the TST. Four controllers found it was improved in ORG2 but the other three found it had become worse, and no controller felt it had reduced workload. Controllers noted that it took a while to get used to the new order of buttons, and were observed to hover persistently over the wrong button, waiting for it to become activated.

Three controllers found that the MIW and MOW had improved from ORG1 to ORG2, but only one found that the modifications had reduced workload. The mixture of proposal and handover messages meant that both tactical and planning controllers found the MOW confusing to use, in both ORG1 and ORG2.

6.3 Impact of changes on tool usage: ORG1 versus ORG2

In addition to a reduction in workload, it was envisaged that the changes made to the GHMI, between ORG1 and ORG2, would encourage and facilitate tool use.

The number of times both the HIPS and ADFL were used to modify a trajectory was extracted from the NRF system recording for each exercise in ORG1 and ORG2. There was no evidence of a change between ORG1 and ORG2 in the number of times new trajectories were planned using HIPS. The ADFL was only used rarely to modify trajectories, both in ORG1 and in ORG2.

The number of times the ADFL was invoked was also extracted from the NRF system recording for each exercise. The planning controllers and the tactical controllers both invoked the ADFL significantly fewer times in ORG2 than in ORG1. The average number of invocations per exercise was down from 28 to 22 for the planning controllers, and down from 6 to 3 for the tactical controllers.

Figure 8 shows how the controllers' usage of HIPS varied from ORG1 to ORG2. The measure used is the average percentage of aircraft trajectory registers which had edits

made in each HIPS display. The remaining aircraft trajectory registers were accepted as proposed without modification. The graph gives an indication of the size and range of each measure, and is an initial comparison between the ORGs. Statistical differences cannot be inferred from the graph alone: the statistical tests employed used matched pairs rather than averages (reference 11). The vertical bars on the graph represent the 25th and 75th percentile range of the measure, and the centre point represents the median.

Figure 8 indicates that usage of the HIPS APS was greater than that of the HPS, which in turn was greater than that of the SPS, which was rarely used. Usage of the APS did not vary much between ORG1 and ORG2, for either controller role. For the planning role, HPS usage clearly increases from ORG1 to ORG2; on average by a factor of almost 3. The increase in usage was statistically significant. On average, the planning controllers used the HIPS far more than the tactical controllers.

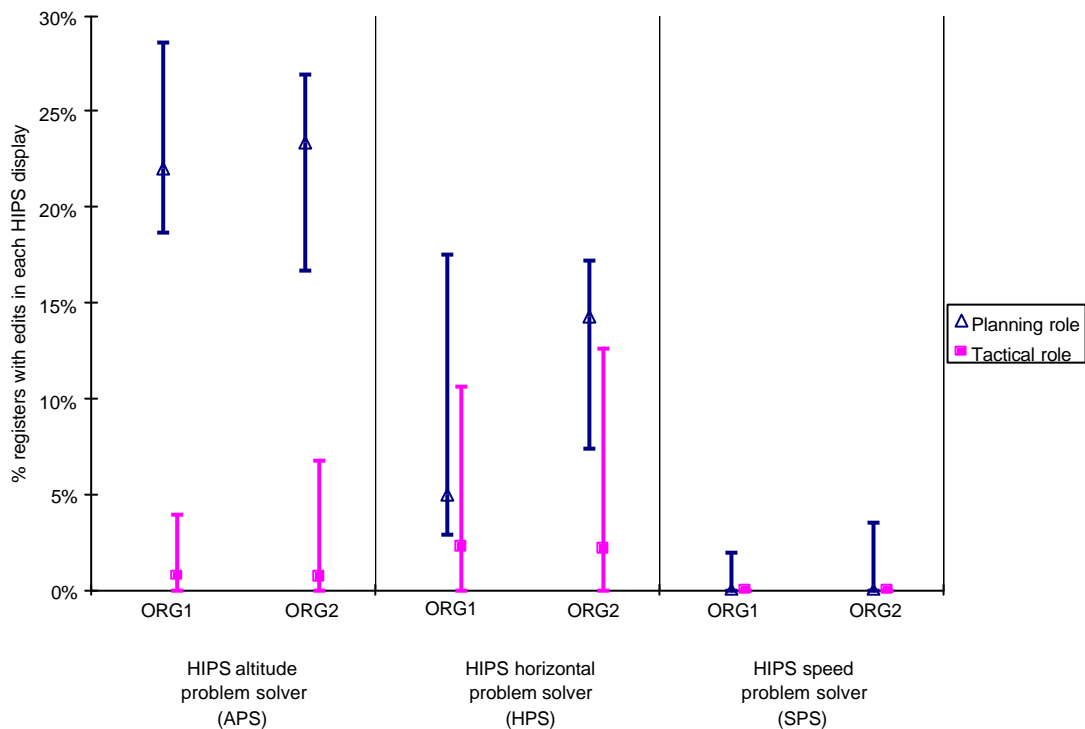


Figure 8 - Comparison of HIPS usage

In their questionnaire responses, five controllers indicated that they could have made more use of the tools during ORG1. The same five found the additional training they received before ORG2 to be useful in teaching them to use the workstation features properly, and only one felt he/she could have used them more during ORG2.

6.4 Comparison of capacity between all the ORGs

One of the recommendations that emerged from Work Package 0 of the IOCP, reference 1, was that work should be undertaken to explore the relationship between workload and throughput. It was suggested that a threshold of throughput may exist, which in some way reflects the capacity of the sector under a certain experimental organisation. For the purposes of the comparison between the ORGs, the ‘sustained aircraft capacity’ is defined to be the average number of aircraft assumed on frequency during these periods when an ISA score of 3 was given. That is the level at which the controllers workload is considered comfortable and sustainable. Application of this definition must be applied with care - ISA

3 is liable to be applied to a range of workload situations, and care must be taken to ensure that the full range has been accounted for.

For the planning controllers, there was no statistically significant difference between the sustainable capacity for each ORG; the questionnaire responses supported this result.

For the tactical controllers, ORG3 had a significantly greater sustainable capacity than the other ORGs. In fact, for three of the tactical controllers, workload remained in the ISA 1 and ISA 2 categories, even in the busiest traffic samples. There was no evidence of any difference in sustainable capacity between ORGs 0, 1 and 2.

6.5 Comparison of system usability between all the ORGs

Questionnaires are one of the battery of tools and techniques recommended by the Usability Assessment Methodology (see reference 12). Software Usability Measurement Index (SUMI) questionnaires were used to measure quantitatively how usable the controllers found the different ORGs in the PD/1+ trial.

SUMI analysis is split into five different aspects: efficiency, affect, helpfulness, control, and learnability, as well as giving the global usability score, thus allowing diagnosis of any problematic aspects of usability in each ORG. See Appendix C for a brief definition of these categories.

The global scale is the most reliable scale in SUMI. It has a mean of 50 with a standard deviation of 10; therefore, 68% of the population will lie between 40 and 60. The global scale is a good measure for comparisons between ORGs and is a general usability benchmark. If an ORG is to be considered state-of-the-art, then the target for the global score should be greater than 60. ORGs scoring between 50 and 60 are by definition above average, ORGs scoring 50 and 40 are below average.

Any of the other sub-scales which are at or below 50 indicate that they are considered poor for that aspect. Sub-scales at or below 40 indicate the need for remedial action. Good software will achieve scores of 60 or more in most sub-scales. As a rule of thumb, reasonable acceptable commercial software rarely falls below 55 in most sub-scales.

Scores below 20 or above 70 for any SUMI scale indicate that the user has responded in an atypical fashion.

In a research environment, the above interpretation of the SUMI scores may be slightly relaxed. A definition of the SUMI score more relevant to the maturity of the software may be employed. For software of the PD/1+ maturity, above 40 is understood to be acceptable.

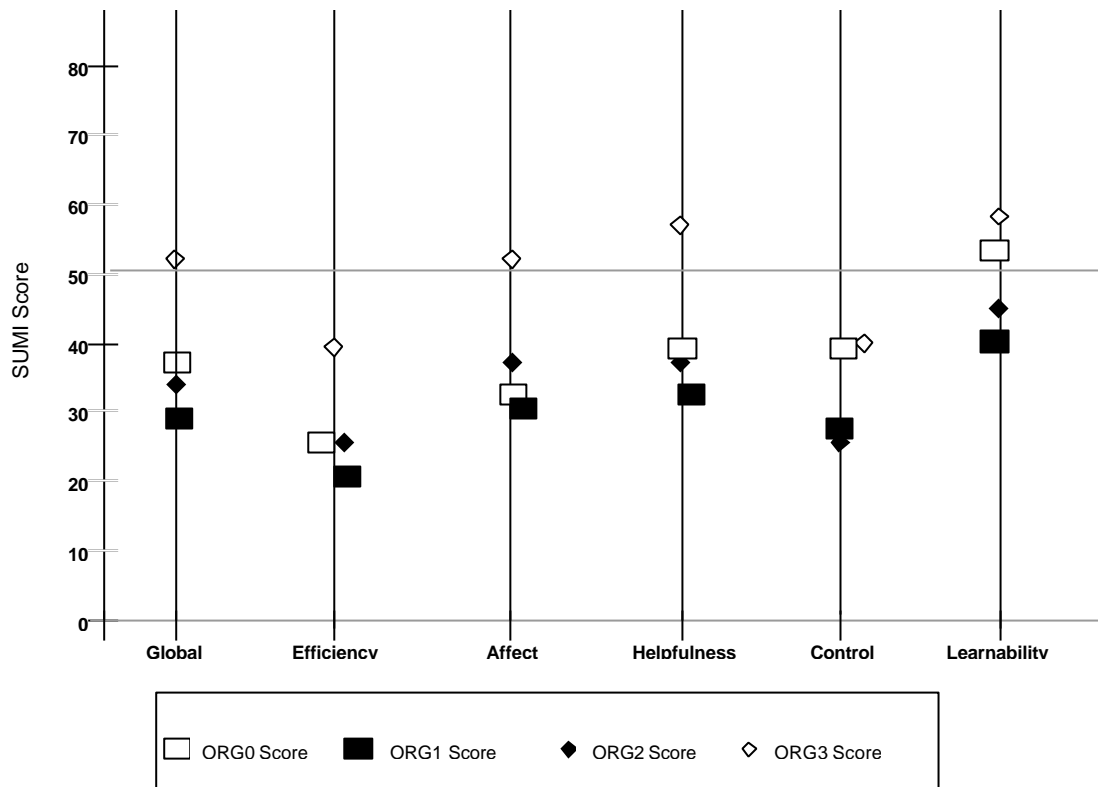


Figure 9 - Summary of SUMI scores

The results of the SUMI analysis is presented in Figure 9, a complete description is given in Appendix C. Using the global scale, the controllers rated the overall usability of all four ORGs as well below standard for state of the art commercial software. In fact, only the ORG3 system was judged to be above average, and ORGs 0, 1 and 2 were judged as unacceptable and to require remedial action. The results of this individual sub-scales are summarised below:

- Efficiency is the extent to which the controllers felt the software enabled tasks to be performed quickly, effectively and economically: only ORG3 proved satisfactory and ORGs 0, 1 and 2 scored very poorly and require remedial action;
- Affect is a measure of “likeability”, namely how much the controllers enjoyed interacting with the system: ORGs 0, 1 and 2 were judged to be below average and ORG3 was rated as above average;
- Helpfulness reflects the impact of the visual cues and feedback offered: ORG3 was judged to be above average, and ORG0 was almost acceptable in this category. ORGs 1 and 2 were rated well below average;
- Control indicates how easy the controllers found it to make the software do what they wanted: ORG0 and ORG3 scored highest, yet the controllers felt their level of control to be only just satisfactory. ORG1 and ORG2 scored very poorly and require remedial action such as reducing the complexity of task sequences, improved labelling of components and faster system response time;
- Learnability indicates how easy the controllers felt the system was to learn: ORG0 and ORG3 were found to be easiest to learn, being above average, indeed ORG3 was considered excellent in this respect. The controllers also found ORG1 and ORG2 acceptably easy to learn.

6.6 Further investigation: learning effects

Many comments were made by controllers stating that they felt ORG2 to be easier to control than ORG1, because by the time they conducted the ORG2 runs they were more familiar with the system and “knew what they were doing”. Due to time pressure on the trial programme, it was only possible to re-run a single extra ORG1 exercise at the end of the fortnight to test these comments. The effects of learning on workload and tool usage efficiency were investigated by comparing between this exercise and the earlier run of the same ORG1 exercise.

One controller was not available in the re-run exercise due to sickness. Moreover, further data were excluded because it was judged that another controller was not taking this re-run exercise in the trial seriously. Hence there was little data for testing the effect of learning. There was no evidence of a difference between the two exercises for the four objective measures tested. Analysis of the ISA scores showed that only one controller perceived the workload to be lower in the re-run exercise.

See Appendix C for a full description of the results of the learning effects investigation.

7. DISCUSSION

This section draws together the results of both the objective and subjective data and in addition the results from the PD/1+ usability study. The discussions relating to the usability of the PD/1+ software rely upon the subjective data gathered during debriefs, observations and comments logged by the controllers themselves. A complete record of this data can be found in reference 10.

A separate subsection has been devoted to each of the individual GHMI and PATs changes, thus enabling a discussion of the relative usability of each change to be included. Although each low level objective was addressed separately during the analysis, they can not always sensibly be separated in this discussion.

It should be noted, when reading this section that ISA values from only two of the planning controllers were available. As detailed in section 6.1, one planning controller was lost to sickness before the completion of ORG1 and remained unavailable for ORGs 2 and 3. A second planning controller appeared, from comments made during the debrief sessions, not to be selecting an ISA score that accurately reflected the perceived workload.

The conclusions of the discussion are summarised in Section 8.

7.1 Overall impact of the changes on controller workload: ORG1 versus ORG2

The majority of objective and subjective measurements decreased significantly from ORG1 to ORG2 for both the tactical and planning roles. This is supported by the PUMA analysis. The contributing factors to this reduction in workload are discussed in the following sections in more detail.

For the planning role, controller comments and observations suggest that a significant factor behind this workload reduction was the change in the size of the no-go zones; planning controllers took less time in ORG2 to both register each inbound aircraft, 27% less on average, and to execute fully a trajectory modification, also 27% less.

For the tactical role, the removal of irrelevant messages from the CLW had the biggest impact, reducing their levels of frustration significantly. There was no evidence of a reduction in percentage time spent in R/T communication, but this is not surprising since in ORG1 these spurious messages were ignored rather than being passed to the pilots

For both roles, controller comments suggest that increased familiarity with the system had some impact on workload, since ORG2 followed ORG1 and training was given between the two ORGs. Learning effects are discussed in more detail in section 7.15. The objective data gathered to explore this effect were insufficient to confirm or refute these comments with confidence.

7.2 Comparison of controller workload between the remaining ORGs

Comparing ORG0 with ORG2, reinforced by ORG0 versus ORG3, suggests that conflict-free planning has shifted workload from the tactical to the planning controllers. This results in the tactical controller having a lower overall workload, a reduction not seen in the results from the PD/1 trial. The planner's workload is no higher in ORG2 than ORG0 (it was higher in PD/1), suggesting that the advanced tools (when modified as per ORG2) allow the planner to absorb the additional work. The effects of increasing familiarity with the system can not be isolated from these results, but the implication is that PD/1+ has achieved the aim of reducing overall workload.

7.3 Impact of the changes to the no-go zones on controller workload

In the comparison between the workload in ORG1 and ORG2, it was found from the questionnaire responses that, for all of the planning controllers, there was a perceived reduction in workload as a result of the smaller no-go zones. This may account, in part, for the reduction in the overall TLX workload scores between ORG1 and ORG2 for the planning controllers, see Figure 4 on page 20. However, since only one of the tactical controllers agreed that the smaller no-go zones reduced workload, it is unlikely that this particular change contributed to any reduction in tactical controller workload between ORG1 and ORG2.

During ORG1, the controllers considered that the excess size of the HIPS no-go zones discredited the use of the tool. This was especially true for the planning controllers, who frequently had to plan aircraft trajectories avoiding no-go zones which, in reality, they could have planned the aircraft through without losing separation. As a result "usable" airspace was unrealistically blocked off and hence the capacity of the sector was reduced. The frustration this caused led to the controllers disregarding the no-go zones on occasions.

Reducing the size of the no-go zones undoubtedly reduced the frustration and workload experienced; however, in reducing the size of the no-go zones to the minimum separation standards, no "margin of error" remained. If it appears from the size of the no-go zone that airspace can be used safely, then with the requirement to constantly improve sector capacity that airspace will undoubtedly be used. However, if the aircraft trajectory was planned just to miss the no-go zone and it subsequently deviated from its expected trajectory, then it would immediately be in conflict. Planning aircraft just to miss the no-go zone was a particular problem for aircraft without datalink, which tend to have a less accurate adherence to the planned trajectory. One possible solution proposed was to use different parameters within the HIPS tool for aircraft equipped with 4D FMS and datalink, and aircraft without. This would give a margin of error for aircraft without datalink, and not rely on the controller to consider the category of aircraft when carrying out conflict-free planning.

As a result, during ORG2 aircraft required closer monitoring than in ORG1 which may have generated additional workload for the tactical controller. In ORG3, where all the aircraft were equipped with 4D FMS and datalink and thus maintained their trajectories precisely, monitoring was not a dominant issue, therefore the smaller no-go zones posed less of a problem. This is likely to be a significant factor in the reduction of workload and frustration experienced by the tactical controllers in ORG3.

7.4 Impact of the changes to the CLW on controller workload

In the comparison between ORG1 and ORG2, it was found from the questionnaire responses that, for three of the four tactical controllers who participated in both ORGs, the changes made to the CLW were the major contributor to a reduction in workload and frustration. This is supported by their overall TLX scores, which were lower in ORG2.

In ORG1 many irrelevant messages were displayed in the CLW, each of which had to be assessed for its relevance and a decision taken on whether to pass the message or ignore it. These messages were a major source of dissatisfaction for the tactical controllers in ORG1 and, to a lesser degree, the planning controllers who, on occasion, referred to the messages to maintain their situation awareness.

The passing of CLW messages in a timely manner was imperative if the trajectory planned for the aircraft was to be maintained. Passing messages late resulted in frustration for both tactical and planning controllers, and increased workload for the tactical controller as the aircraft had to be re-planned. This was often the case in ORG1.

In ORG1 controllers complained that they spent too much time trying to “decipher” (i.e. read, interpret, and distinguish relevant from spurious information) the contents of the CLW. The changes made to the font in ORG2 contributed to the reduction in workload and frustration experienced by the tactical controllers and the increased ‘like-ability’ and ‘helpfulness’ of ORG2; however, there were still comments during ORG2 that the time display in the title bar needed to be more legible. Suggestions were also made during ORG1 and ORG2 that the tactical controller should be prompted that a message was due to be passed to an aircraft if the timing of the message was critical to maintaining planned separation. Alternatively a ‘buffer zone’ could be introduced around the no-go zone within which the message should be passed. From a usability point of view this would increase ‘control’.

However, the improvements to the CLW unintentionally introduced other problems: on occasions, messages were generated in the CLW concerning entry conditions to future sectors. Hence, although the CLW had been improved and workload and frustration reduced for the tactical controller as a result, some messages in the CLW were still irrelevant and therefore unnecessary workload was still involved in distinguishing these.

7.5 Impact of changes to the HIPS scrolling on controller workload

In comparing ORG1 and ORG2 it was found, from the questionnaire results and debrief sessions, that there was a mixed opinion as to whether this change had reduced workload. Any time saved in the change to the scrolling was counteracted by the fine tuning required to select the correct level or time. One planning controller found this change actually increased workload and frustration.

From a usability point of view, during ORG1 it was observed that controllers became familiar with the scrolling of the HIPS profile displays. Although interaction with the scrolling displays had been highlighted as a problem during the PUMA analysis in work package two of the IOCP, none of the controllers reported any difficulties with the scrolling during ORG1, either through debriefs, questionnaires or comments. It is perhaps not surprising, therefore, that this change met with a mixed reaction from the controllers.

On initial exposure to the modified scrolling mechanism, the controllers immediately reported that the displays scrolled too quickly and were harder on the eyes, the ‘like-ability’ of the system being reduced. As they scrolled down, the display would move up and therefore scrolling was much faster than expected, making it difficult to select the required level or time. On occasions the controllers would scroll past the required point,

causing them to subsequently return to the aircraft trajectory and fine tune it. The 'control' of the system had been reduced.

While the modified scrolling mechanism was not immediately intuitive, it gained some measure of acceptability through use, and hence two planning controllers and one tactical controller recorded an eventual reduction in workload and frustration.

7.6 Impact of changes to the HIPS speed problem solver on controller workload

Any comparison of the SPS use between ORG1 and ORG2, with respect to workload, was virtually impossible due to the lack of use of this window by controllers in both ORGs. Only one of the planner controllers actually attempted to use the HIPS SPS, and thus appreciated why the changes had been made. However he stated that they did not reduce his workload or frustration in ORG2. The changes made did, however, meet with controller approval, not because they reduced workload but because it improved the 'helpfulness' of the system. This is discussed in the following paragraphs.

Controllers reported that they were unable to understand the HIPS SPS, could not get it to work, and more often than not reduced it in size so as to be useless. 'Greying out' the SPS for aircraft without datalink did however, highlight, the lack of a distinguishing feature for aircraft without datalink in the HIPS during ORG1. Hence, in principal, the change was approved by the controllers as it reinforced the distinction between aircraft with and without datalink in ORG2.

The controllers also questioned the relevance of the SPS in the HIPS, commenting that speed restrictions are normally given to aircraft to maintain separation and not to establish it. This appears to be a misunderstanding of the *predictive* nature of the HIPS display, in which a no-go zone usually indicates that aircraft which are currently separated will come into conflict in the future. Thus a speed constraint is to maintain, not establish, separation. Their views therefore indicate the difficulty of learning the SPS.

7.7 Impact of changes to the default screen configuration on controller workload

This GHMI modification represented a subtle improvement to the display and all controllers agreed that having the facility to, and being encouraged to, re-organise their screens was beneficial. Having established their own arrangement, it was stated that they were unlikely subsequently to change it. One planning controller suggested that their own arrangement led to more economical mouse movements. This would have led to a reduction in workload, though this was not directly stated by the controller, nor was it feasible to objectively measure workload in relation to the controller's own arrangement. It was not clearly established, therefore, that the facility to reorganise the screens, and the subsequent 'optimisation' of the display by each controller, led to a reduction in workload, though it was considered to be a positive change, improving the usability of the system.

Forcing the controllers to consider how they used the various aspects of their display, and hence establish an optimum layout for their role and sector, can only be beneficial. For example, a discussion during a debrief regarding the best position for the MIW/MOW concluded that it was at eye level above the PVD.

All of the controllers preferred their own window arrangement, which they derived and saved as preferences during ORG2.

7.8 Impact of changes to the colours in the no-go zones on controller workload

Six out of the seven controllers felt that the colours in ORG2 were an improvement over ORG1, but that this had not directly reduced their workload. These opinions may be driven more by the difficulty of measuring the impact of such GHMI changes rather than an actual lack of impact: controllers naturally find it difficult to identify if changes which do not obviously affect the timing or difficulty of tasks actually have an impact on their workload.

Frustration in ORG2 was lower than in ORG1 (see Figure 4) and, from a usability point of view, ORG 2 was more 'likeable'. Given the comments above, the best that can be deduced is that the improved colours in ORG 2 are likely to have contributed to these improved results, but the extent of the contribution is impossible to measure.

7.9 Impact of improving the point-out facility on controller workload

In comparing ORG1 and ORG2 it was found, from the questionnaire results, that the change to the point-out facility had no impact on controller workload. It was used very rarely in either ORG, since it was considered by the controllers not to have the correct functionality, as is discussed in the following paragraphs.

From a usability point of view the controllers felt that the feature pointed out was easy to miss, and pens and fingers were used instead. In general, in its present format, despite the colour change, it was considered superfluous between two adjacent controllers: "if I see something highlighted, I'll ask why it is highlighted and we might as well just talk".

From a 'like-ability' point of view, although there were mixed feeling about the benefits of the new colour, discussion was more centred around the concept itself. It was felt that a point-out facility would be of use between two sectors located at either end of an operations room. An aircraft would be pointed out, followed by a telephone call to discuss the problem. An annoying feature was also identified: the point out had to be cancelled by both controllers on their own screens. If this was omitted by either controller, the point out remained with the aircraft into the next sector.

The 'helpfulness' of the system was improved in the view of at least one controller, the brighter colour of the point-out facility being used as an aide memoir. With no paper flight

strips available, this was a facility which a number of the controllers felt was otherwise missing from the current system.

7.10 Impact of changes to the trajectory support tool (TST)

In comparing ORG1 and ORG2 it was found, from the questionnaire and debrief results, that two tactical controllers and two planner controllers considered the new TST an improvement, the remainder considering it a detrimental step. None however considered it to have reduced workload, believing there to be other more important issues to be resolved before a reduction in workload could be expected: the need to validate and register a trajectory through two separate buttons; the time taken to activate the buttons; and the need to close the TST manually. These problems reduced the 'efficiency' of the system, in usability terms.

The system was not helpful to controllers when a trajectory had been rejected, since no reasons were given by the system. It is possible that this could increase the controllers frustration; however no evidence was found to support this.

Controllers noted that it took a while to get used to the new order of the TST buttons in ORG2. This may well be a reflection of the trial design, since the controllers had got used to the "inefficient" order in ORG1 and been confused by the "improved" order in ORG2.

7.11 Impact of changes to the MIW/MOW

In comparing ORG1 and ORG2 it was found, from the questionnaire and debrief results, that two of the three planning controllers and one tactical controller considered that the changes to the MIW/MOW were an improvement. The remaining three tactical controllers felt that there had been no change.

A tactical controller is unlikely to refer to the MIW unless it is to assume aircraft; however, this can also be achieved through the TDB. It is therefore possible that, for the three tactical controllers who recorded no change, the MIW and MOW were little used, the controllers preferring to assume aircraft through the TDB. The one tactical controller who considered that there had been an improvement also recorded a reduction in workload, since in this case the MIW was being used to assume aircraft. However, the reduction in workload may also have been due the improved positioning of the MIW during ORG2. The tactical controller had positioned the MIW in the centre of the display adjacent to the CLW.

From a usability point of view, both the 'likeability' and 'helpfulness' of ORG2 were improved through the changes made to the font used in the MIW and MOW. This is supported by controller comments that they spent "too much time trying to decipher the contents of the window" within ORG1.

7.12 Impact of changes on tool usage

The results suggest that perhaps the biggest impact on the amount of tool usage between ORG1 and ORG2 was from the change made to the size of the no-go zones in the HIPS displays. With smaller areas to avoid when planning, it was far easier to resolve conflicts by giving a heading change in ORG2 than in ORG1. This was demonstrated by the significant increase in the number of times the Horizontal Problem Solver was used in ORG2, when compared to ORG1. Interestingly, the HPS was not used at the expense of the APS, with controllers instead giving altitude changes coupled with heading changes.

During both ORG1 and ORG2, the ADFL was used on very few occasions to modify a trajectory, the controllers finding the additional information presented within HIPS more useful in planning a conflict free trajectory. The ADFL was used to merely register a trajectory for an aircraft whose plan did not need modification. In ORG2 the improvements

made to the HIPS no-go zones encouraged both the planning and tactical controllers to use the HIPS in the place of the ADFL. Hence the number of times the ADFL was invoked fell significantly in ORG2.

7.13 Comparison of capacity between the ORGs

In attempting to measure any difference in capacity between the ORGs, the number of aircraft on frequency at the time an ISA score of 3 was given were recorded. It was envisaged that this measure would give an indication of the sustainable capacity, the ISA 3 category representing a sustainable level of workload. The results indicate no difference in capacity measured in this way between each ORG for the planning controllers and no difference between ORGs 0, 1 and 2 for the tactical controller. However it was not possible to carry out the planned ORG 3 analysis for the tactical controllers, since at no time did a tactical controller select any category but 1 or 2. This would suggest that in ORG 3 the sustainable capacity had not been reached for the tactical controllers at least.

The use of ISA responses to give an indication of sustainable capacity is relatively new, and needs further investigation. If there are no ISA 4s selected, then the number of ISA 3s may not be representative, since there would be no indication whether the controller could have taken more aircraft without reaching ISA 4. Equally if there were only one or two ISA 3s selected the statistical validity of the data should be questioned. It may also be necessary to consider the peak number on frequency and the impact that this has on capacity. The results from this analysis should, therefore, be treated with some caution.

7.14 Comparison of system usability between the ORGs

Of all the ORGs, only ORG3 was judged to be above average in terms of usability. A number of issues were raised by the controllers which were considered to have had an impact on this result.

Although the controllers' perception of workload associated with both tactical and planner roles had decreased in each of ORGs,1,2 and 3 compared with ORG0 it was at the expense of the tactical controllers' skilled input: the tactical role was reduced to that of a mouthpiece for the communications. It was suggested that the PD/1 concept is better suited to an environment in which all aircraft are datalink equipped, such as ORG3. Datalink equipped aircraft are more likely to keep to their planned trajectories than those without datalink. Hence ORG3 proved to be the most efficient of the ORGs. However, it is also the most wasteful of controller resources as the tactical controller is under-utilised.

The controllers expressed concern at the lack of a visual reference. The information previously contained in strips was dispersed across the system, reducing the 'control' they felt.

It was proposed that the PD/1 concept would be best exercised in a future airspace concept. Controllers acceptance of the concept and system was affected by the apparent lack of consideration that had been given to the design of the airspace being simulated, with respect to the future time frame in which it was expected to operate. Accordingly, both the usability and trial evaluation were hindered by employing the PD/1 concept in contemporary (i.e. NERC) airspace.

The size of the sectors, the restrictions placed on aircraft to fly within current route structures, and the imposition of standing agreements, all served to undermine the concept under evaluation. One controller commented during a debrief that "The system is being dragged back because it is being used in a scenario of traditional ATC. It would be more positive to use it in open skies, if only we didn't have to stick to pre-determined tracks and

adhere to beacons. You have got to let us fly in straight lines; more economical and airline friendly.”

No adverse comments were received in respect of the realism of the traffic samples. If the controllers had perceived problems, then the usability and trial evaluation would have been further hindered.

7.15 Further investigation: learning effects

From the objective results gathered it would appear that learning effects were not an issue between ORG1 and ORG2 of the trial; however, from controller comments and the fact that training exercises existed between the two ORGs, it would seem reasonable to presume that learning effects existed and contributed to any reduction of workload recorded. One controller’s ISA scores were apparently affected by increased familiarity, but one should be wary of generalising from this result from a single controller. The extent to which learning effects contributed overall cannot be quantified, since one exercise provided insufficient data to explore this effect fully.

8. CONCLUSIONS

The conclusions are summarised as follows:

- The workload in ORG2 was less than the workload in ORG1 for both planning and tactical controllers. This reduction was due, for the main part, to the changes made to the CLW in the case of the tactical controllers, and the changes to the size of the no-go zones in the case of the planning controllers; (Section 7.1)
- The smaller no-go zones contributed to a reduction in workload for the planning controller, but not the tactical controller; (Section 7.3)
- The smaller no-go zones encouraged controllers to plan separations which were ‘tighter’ than they felt comfortable with, and which could prove problematic for aircraft without 4D FMS and datalink; (Section 7.3)
- The removal of irrelevant messages from the CLW and the improved font reduced both workload and frustration for the tactical controllers, and increased the usability of the system; (Section 7.4)
- The time of issue of messages over the R/T from the CLW is critical; (Section 7.4)
- From the information gathered, changes made to the HIPS display scrolling could potentially reduce overall controller workload, if further thought were given to the speed of scrolling; (Section 7.5)
- There was a lack of understanding of the need for, or intended use of, the SPS in HIPS; (Section 7.6)
- Changes made to the SPS in the HIPS window did not contribute to a reduction in controller workload; (Section 7.6)
- In terms of usability, the introduction of a distinguishing feature in HIPS for aircraft without 4D FMS and datalink increased the ‘helpfulness’ of the system¹; (Section 7.6)

¹ It should be noted that this was a side-effect of the greying out of the SPS in HIPS.

- Changes to the default screen configuration did not contribute directly to a reduction in workload. They were, however, considered to be a positive change, improving the usability of the system; (Section 7.7)
- Changes to the colours of the no-go zones in HIPS did not reduce controller workload. They did, however, improve the 'likeability' of the system from a usability point of view; (Section 7.8)
- The change made to the colour of the point-out facility did not reduce controller workload; (Section 7.9)
- The changes made to the TST did not reduce controller workload, there being a number of perceived problems with the use of the TST as a whole; (Section 7.10)
- The changes made to the font in the MIW and MOW did not reduce controller workload, but did improve the usability of the system; (Section 7.11)
- Increased familiarity with the system may have been partly responsible for the reduction in controller workload; (Sections 7.1,7.15)
- Changes made to the HIPS no-go zones increased the amount of HIPS tool usage in ORG2, at the expense of the ADFL; (Section 7.11)
- In terms of usability, only ORG3 proved to be above average; (Section 7.14)
- PD/1+ (i.e. ORG2) reduced overall workload compared to the baseline of PD/1; (Section 7.13)
- The traffic samples were adequately realistic; (Section 7.14)
- The use of a contemporary operational concept and airspace structure hindered the outcome of the trial from a usability point of view; (Section 7.14)

9. RECOMMENDATIONS

The recommendations have been broken down into two categories; those which are quite specific and relate in particular to the PATs and GHMI changes made during the PD/1+ trial, and relate to the effect these had on workload and system usability; and a number of broader recommendations relating to the overall PHARE concept and future PHARE Demonstrations, in particular PD/3.

9.1 Recommendations relating to the PATs, and GHMI changes:

- The no-go zones within the HIPS should be no larger than is necessary to safely maintain the current separation standards;
- The separation parameters used within the HIPS tool for aircraft equipped with and without datalink should be reviewed;
- More attention should be given to the contents of the CLW, the timing of the messages as they appear in the CLW and the time constraints within which they should be issued;
- From a usability point of view, the speed of scrolling within the HIPS needs further investigation;
- Consideration should be given to whether the SPS within HIPS is required, and under what conditions it should be used;
- A dedicated method for identifying aircraft without datalink is required in the HIPS;

- A default screen configuration should be used which positively encourages the controller to optimise the display for their role and sector;
- Consider adapting the point-out facility or providing an alternative aide memoire facility to the controller;
- Consider reducing the number of buttons on the TST, by combining validate and register, to reduce the time involved in the validate and register process;
- Provide more feedback when the validate or register is unsuccessful, and remove the need to close the TST after use.

9.2 Recommendations for PD/3 and future trials:

- Future trial designs should be such that the effects of learning are minimised and can be quantified if present;
- Realistic traffic samples, an operational concept and an airspace structure must be developed which will maximise the potential of the trial or demonstration and gain participant acceptance.

10. GLOSSARY

4D FMS	4D flight management system
ADFL	Augmented Dynamic Flight Leg
APS	Altitude Problem Solver
ATC	Air Traffic Control
ATM	Air Traffic Management
ATMDC	Air Traffic Management Development Centre
CLW	Communications List Window
CRD	Conflict Risk Display
DASR	NATS Department of ATM Systems Research
DRA	Defence Research Agency
GHMI	Ground Human Machine Interface
HAW	Horizontal Assistance Window
HIPS	Highly Interactive Problem Solver
HPS	Horizontal Problem Solver
IOCP	Internal Operational Clarification Project
ISA	Instantaneous Self Assessment
LATCC	London Area and Terminal Control Centre
MIW	Message In Window
MOW	Message Out Window
NATS	National Air Traffic Services Ltd
NERC	New En Route Centre
NRF	NATS Research Facility
ORG	Organisation
PATs	PHARE advanced tools
PD/1	PHARE Demonstration 1
PD/3	PHARE Demonstration 3
PHARE	Programme for Harmonised ATM Research in EUROCONTROL
PRT	Preferences Tool
PUMA	Performance and Usability Modelling in ATM toolset
PVD	Plan View Display
R/T	Radio Telephony
RTB	Radar Tool Box
SIL	Sector Inbound List
SPS	Speed Problem Solver
SUMI	Software usability Measurement Index

TDB	Track Data Block
TLX	NASA - Task Load Index
TMA	Terminal Movement (Control) Area
TST	Trajectory Support Tool
VAW	Vertical Assistance Window
WAT	Workload Assessment Tool

11. REFERENCES

1. Replacement of PD/1 baseline with the NERC Reference System, Eurocontrol PD3/NATS/IOCP/WP001/1.1, S. Kay, February 1997
2. PUMA Based Task Analysis (WP1), Eurocontrol PD3/NATS/IOCP/WP002/1.0, M. Hook, April 1997 (Draft)
3. Analysis of PHARE PD/1 Tasks (WP2), Eurocontrol PD3/NATS/IOCP/WP003/1.0, A. Kilner & M. Hook, April 1997 (Draft)
4. PUMA Based Analysis of Tactical and Planner Controller Workload (WP3), Eurocontrol PD3/NATS/IOCP/WP004/1.0, A. Kilner, April 1997 (Draft)
5. PUMA Analysis Summary Report, Eurocontrol PD3/NATS/IOCP/RPT001/1.0, A. Kilner, April 1997 (Draft)
6. PD/1 Final Report, Eurocontrol PHARE/NATS/PD1-10.2/SSR;1.0, NATS, October 1996
7. Assessment Methodology For The NATS Research Facility, NATS DASR\TRIAL-MF\RPT\004, R. Whitaker et al, March 1997 (Draft)
8. PUMA Version 2.2c User Guide, Roke Manor Research Ltd report no. X72/HB/1320/000, M. Hook, 1996
9. NATS PD/3 IOCP: Project Plan - Analysis Plan Supplement, NATS DASR\PD3IOCP\AP\001\1.0, D. Hudson, January 1997
10. PD/3 IOCP Raw Data, NATS DASR\PD3IOCP\TN\024, A. Lucas & L. Clarke, March 1997
11. Nonparametric Statistics for the Behavioural Sciences (2nd edition), McGraw-Hill, S. Siegal & N.J. Castellan Jr, 1988
12. NRF Usability Assessment Methodology, NATS R&DD report in preparation, A. Lucas, 1997
13. PD/1 Usability Audit, NATS DASR\PD1\RPT\002, L. Clarke, 1996
14. PD/3 IOCP GHMI Modifications, NATS DASR\PD3IOCP\TN\025, A. Parkins, January 1997
15. W/INDEX: A predictive model of operator workload (Applications of Human Performance Models to System Design, edited by G.R. McMillan et al, pp. 81-89), Plenum (New York), P.A. North & V.A. Riley, 1989