An Investigation of Three Dimensional Displays for Real-Time, Safety-Critical Command/Control Applications: with application to Air Traffic Control

Brown, Mark Anthony

ISSN 1369-1961

Department of
Computer Science

Technical Report No. 765

An Investigation of Three-Dimensional Displays for Real-Time, Safety-Critical Command/Control Applications:

with application to Air Traffic Control

Mark Anthony Brown

QUEEN MARY
AND WESTFIELD COLLEGE
UNIVERSITY OF LONDON

February 1996

An Investigation of Three-Dimensional Displays for Real-Time, Safety-Critical Command/Control Applications:

with application to Air Traffic Control

by

Mark Anthony Brown

Submitted to the University of London for the degree of Doctor of Philosophy in Computer Science

Department of Computer Science
Queen Mary & Westfield College
University of London
Mile End Road
London E1 4NS

February 1996

Acknowledgements

No man is an island....
John Donne, 1571?–1631.

The author would like to thank the numerous people who have offered support, advice and encouragement over these past years during the Ph.D.

Firstly to my internal supervisor, Mel Slater, formerly of the Department of Computer Science at QMW and my external supervisor, Professor John Vince, formerly Chief Scientist at Thomson Training and Simulation Ltd.

Other QMWites the author would like to thank are David Sweeting of the Department of Aeronautical Engineering for his advice, Shancy Kou, another long-suffering Ph.D. student for her help and support, and Rob Kooper for helping with the finer points of LaTeX, EMACS and DVS. The author would also like to thank the members of the department who 'volunteered' as guinea pigs and experimental subjects without excessive coercion!

At CAA NATS, thanks go to Peter Havelock, Geoff Strutt and Dr. Sarah Sharkey for supporting the project and battling to move it ahead, and to the ATCOs at Heathrow tower, especially Dick Gallop, who submitted cheerfully to my experiments and who helped enormously with many hours of interesting and useful discussion, and who made me feel most welcome on my visits to Heathrow.

The author would like to thank Division Ltd. for the loan of a pair of CrystalEyes LCD glasses, and their technical support.

Thanks also go to the commercial flight crews who have put up with my visits to the flight deck, my questions, and have allowed me to observe landings. These have afforded valuable insights into how the air traffic control system operates from the 'receiving end'.

Final thanks go to my parents, family and friends who have put up with me (no easy task) and have offered support, both financial and moral, throughout.

Abstract

This research investigates the suitability of three-dimensional (3D) display technologies for real-time, safety-critical command/control of objects in 3D space. The study deals with perceptual issues relating to use of 3D displays for such tasks, and examines the benefits of the various displays in terms of task performance. Air traffic control (ATC) is used as a major focus and an example for this research.

Tasks which require the visualisation of objects in 3D space suggest themselves naturally as possible applications of 3D displays. However, to be adopted by a user community, a new display technology must have significant benefits over the *status quo.* New technologies often bring new problems, such as new types of error in their use. Nowhere must the benefits and drawbacks be more thoroughly investigated than in applications where there is little room for error, where the quality of the information presented and the way in which it is interpreted by the human operator impact on the performance of some safety-critical task.

A comparative study of three displays of aircraft position, such as might be used for air traffic control, was carried out for a number of tasks. The display types were: A two-dimensional plan-view format based on the current radar display (serving as a base for comparison), a pseudo-3D "through-the-window" (non-immersive) display, and stereoscopic 3D TTW display. The study used two groups of subjects: air traffic control officers (ATCOs) and non-ATCOs. Basic elements concerned with the reading of information from the display (reading of angles and distances, extraction of altitude and horizontal separation) were identified and investigated. The effect of display format on the memory of a scene was also investigated. Two tasks which required a combination of the basic elements, a conflict detection task and an interceptor control task, were included to see whether the performance in the basic task elements was a predictor for performance in a compound task.

A pilot study carried out with 9 ex-ATCOs confirmed significant differences between performance in tasks carried out in the different display types. The pilot study indicated that a 3D parallel projection yields better accuracy for judgment of angles than a perspective projection, but unexpectedly there was no significant difference in the accuracy of reading distances between the two projections used. The main study confirmed the results of existing research into 3D displays, but suggested that while some ATCOs may have a 3D spatial mental model of the air traffic which they are controlling, there is a tendency to treat the horizontal and vertical dimensions separately for control purposes. If this is the case, it may undermine some of the arguments for adopting 3D displays for ATC.

Additionally, an immersive (virtual reality) 3D display was presented to ATCOs to gain opinions for possible future applications to ATC.

Contents

List of Figures

List of Tables

Chapter 1
Introduction

2D or not 2D: That is the question.
With apologies to William Shakespeare.

1.1 Introduction

1.1.1 Three-Dimensional Displays: The Benefits...

Vision can be an extremely efficient mode of communication; photographers, artists, advertisers and teachers have long used the visual sense to explain concepts and to evoke emotional responses. Perhaps because our visual sense is so important, there has long been a drive to utilise computers to generate and display ever more sophisticated imagery.

Since the real world has three spatial dimensions, three-dimensional (3D) displays, which directly represent distances along the line of sight (LOS) into the display (*depth*), may allow for a more 'natural'[1] display of objects. Since we have evolved to make sense of the 3D visual world, representation of information using 3D images may afford more effective communication of that information to a viewer "by exploiting the analytic and integrative characteristics of visual perception" [WT90]. This may lead to lower workload and increased efficiency in performing a task with that information.

Carrying these assertions further is the idea of *virtual reality* (VR)[2]. In the 1960s, Sutherland wrote a landmark paper describing his 'ultimate display' [Sut66]. This would 'immerse' the viewer in a computer-generated artificial environment complete

[1] The words 'natural' and 'intuitive' seem to be used a lot in the literature when referring to 3D displays, but precisely what is meant by these terms is very difficult to define from a rigorous scientific point of view. However, they will be operationalised later in this thesis.

[2] The definition of 'virtual reality' is the subject of much current debate, with claimants to the title ranging from displays incorporating visually-coupled, interactive graphics presented through a head-mounted display (so-called *immersive* display) to purely textual environments. In this thesis, the term will be used to describe only the former, unless indicated otherwise.

with highly realistic visual, kinæsthetic and tactile feedback, providing the ability to see an object in three dimensions, experience resistance when pushing it and feel its surface texture. Although this ideal is currently not realised, a measure of 'immersion' can be realised today by using head-slaved 3D computer-generated imagery (CGI) presented on a binocular stereoscopic head-mounted display (HMD), giving an image with a compelling impression of 'depth' filling the viewer's field of view and changing as his or her head moves. This requires 3D imagery to be generated at interactive speeds and so requires high performance computing and rendering hardware. Within the past few years, the increasing power of computers, increasing circuit density and the dramatic decrease in cost per unit of performance have brought computers of sufficient power to generate 3D graphics interactively, which are now found in high-end desktop workstations and personal computers and are thus now accessible for a range of applications hitherto prohibited by the cost, size and performance of the hardware required.

The payoffs of appropriate application of 3D graphics are potentially very large. One of the most successful applications is flight simulation. Some flight simulators utilise sophisticated real-time CGI to simulate the view from the cockpit windows. This combines with motion cues and realistic behaviour of the simulated aircraft and its systems to give a high degree of realism, to the extent that the most advanced of these devices allow a pilot to be certificated to fly an aircraft without ever having flown the real thing, with considerable cost savings. Simulators allow potentially hazardous training scenarios to be practiced without leaving the ground, considerably enhancing safety, and also remove weather and other constraints from training [RS86].

Pictorial 3D displays are being investigated in aerospace for more effective presentation of geographical position, flight path information and the disposition of other aircraft to pilots, potentially giving better *situational awareness* and reduced workload and thus improved safety and operational efficiency [HRG84, EMH84, WTS89, SD90, MR90, PG93]. For more down-to-earth applications (literally as well as figuratively!) the ability of 3D graphics to show multidimensional data in a more 'natural, intuitive' way is being used increasingly in the field of *data visualisation* [Ell89, WTS89, WT90], helping scientists and engineers better to understand phenomena such as weather, fluid flows and stresses in structures. 3D displays need not be restricted to concrete spatial entities; they might be used to visualise abstract quantities—for example, the locus of a point in 3D space could represent a process parameter moving in phase space—allowing them to be applied to non-spatial multi-dimensional problems. The æsthetic appeal of 3D graphics alone can be important, such as in the computer games market, where it may contribute considerably to the marketing potential of a game.

1.1.2 ...and the Costs

The æsthetic appeal and 'naturalism' of 3D CGI do not alone guarantee that its application will be successful, however. As Ellis has observed:

1.1 *"Simply casting multidimensional data set into a 2 or 3D spatial metaphor does not guarantee that the presentation will provide insight or a parsimonious description of phenomena implicit in the data."* [Ell89]

There are a number of drawbacks associated with 3D displays:

1. They are often more expensive, in terms of hardware and software resources required, than 2D displays.
2. There are additional design factors, such as optimisation of viewing parameters, in addition to those which apply to 2D displays. Certain aspects of viewing parameter choices can lead to distortions or biases in viewing.
3. The 3D representation creates some ambiguity regarding the precise location of objects along the LOS.
4. The very integration that affords 'holistic' perception may result in reduced precision in reading values along any one particular axis.

[WTS89, EH90, SWK90]. The choice of whether or not to adopt a 3D display is therefore not one which should be taken lightly, nor is it straightforward. The display must be designed to support the task and the human performing the task, and it must be realised that any device employed for human tasks introduces its own characteristic human errors [Hop94], which must be identified and addressed. Failure to understand properly the associated perceptual and technological factors may yield a design so poor that any potential gains are negated. Yorchak and Allison summarise this nicely:

1.2 *"Efficient display designs should consider the capabilities and limitations of the human cognitive system."* [YA85]

Unfortunately, the development of guidelines for the design and use of interactive computer graphic displays has lagged behind the growth of technology for implementing them [YA85], and there is a danger of misapplying the technology, either by using it where inappropriate or by bad implementation in applications where it is appropriate.

1.2 The Research Problem

1.2.1 Motivation

The above suggests that 3D displays may have the greatest benefits for tasks where holistic spatial awareness is critical. One such class of applications may be the monitoring and control of trajectories of objects moving in space. These could be real objects, such as aeroplanes, submarines and ships, or abstract entities. Often, such applications are real-time and may be safety critical or otherwise sensitive to error. Examples are fighter control and air traffic control (ATC), which involve controlling, respectively, fighters intercepting possible hostile intruders, and aircraft flying from place to place. These tasks are real-time—the positions of aircraft are constantly changing; and they are sensitive to certain types of error—a poor interception may result in the hostile getting away or shooting down the interceptor, and poor ATC may result in an 'aluminium shower' (mid-air collision) or 'controlled flight into terrain'.

The real-time and error-sensitive natures of such tasks are what makes this type of application interesting, since they place more stringent constraints on the presentation of information *vis-à-vis* speed of assimilation of a situation and the quality of the perceived data. In a real-time application where critical decisions may have to be made quickly, the required information must be quick and easy to read and difficult to misinterpret.

In ATC, 3D displays might allow more rapid identification of potential hazards through better situational awareness, improved efficiency by allowing separation minima to be reduced safely, visualisation of phenomena such as weather fronts and wake turbulence, and lower workload. However, in order for a new technology to be adopted, it must demonstrate significant benefits over the *status quo* and any benefits have to be weighed carefully against the costs. For example, even if a 3D display gives better performance than a 2D display for a given task, adopting it may require considerable investment in new equipment and training.

Previous studies constructing 3D displays for ATC (e.g. [Str91b, Tam93]) have aimed at demonstrating the feasibility of a 3D display as a novel way of displaying air traffic. However, very few (with the notable exception of [BB91]) have addressed the issue identified by Yorchak and Allison (Quote 1.2), that is whether or not the cognitive limitations of three-dimensional displays make them fundamentally suitable or unsuitable for ATC. Wickens and Todd [WT90] cite two relevant research domains in making the choice between a 2D and a 3D representation: 3D display research and the *proximity compatibility principle*. From 3D display research, two key factors are the costs of position ambiguity along the display LOS and the inherent distortion of distance judgments along axes which are not parallel with the viewing plane. If tasks such as ATC require such judgments to be made with precision, then a 3D rendering may not be suitable. The proximity compatibility principle, asserts that "*tasks of a more integrative nature involving (for example) the comparison between data points will benefit from more 'object-like' displays, whereas tasks requiring the focus of attention on a single dimension or single object will be better served by more separated bargraph or digital displays*".

In order to consider adoption of 3D displays for a task such as ATC, an investigation is required to address some of these issues.

1.2.2 The Research Problem

This thesis describes the investigation of a utilisation of 3D displays for real-time, command/control type applications, concentrating on possible application to ATC. The objectives were to make an exploratory empirical investigation into the use of 3D displays from a human factors viewpoint, comparing them with a 2D plan-view display similar to those currently in use in order to identify possible strengths and pitfalls in using 3D displays and to highlight areas for further investigation.

Three types of 3D display were chosen for the investigation: Pseudo-3D, stereo-3D and immersive 3D.

The conventional method for showing 3D CGI is to render a 2D projection of the 3D scene on a VDU. Such a display is referred to here as a *pseudo-3D* display since it employs only monocular cues as an artist might in a painting or drawing and does not give any sense of 'real world' depth. This is relatively cheap and straightforward to implement.

Research into 3D displays suggests that the fidelity of 3D renderings and the accuracy of distance judgments along the LOS is determined by a weighted additive function of the depth cues employed in the display. For more static displays, such as for ATC, binocular disparity through stereopsis has been found to be a particularly salient cue [WT90]. The presence of disparity might therefore be expected to further highlight the differences between 2D and 3D display formats. However, stereopsis can be expensive to implement. A separate image must be presented to each eye and this imposes a demand on the image generator (IG) which does not exist in pseudo-3D displays, and there are various considerations as to which technology to adopt to display the stereo imagery to the viewer. The effective ones (neglecting severely limited technologies such as the familiar 'red/green glasses') are not cheap. If there are any advantages in the use of a stereoscopic display over a pseudo-3D display, they must outweigh its costs.

One of the most expensive 3D displays, but from some points of view the most effective, is the immersive display, which employs head-slaved visuals. However, this imposes still further demands on the IG.

It was felt early on that the experimental subjects for such an investigation could present a problem. Obviously, using air traffic control officers (ATCOs) would be advantageous, but such individuals are trained in the use of the current 2D displays, and present operational practices are obviously optimised for operation with current equipment. This could lead to a bias in favour of the 2D display. To address this, the investigation was carried out using two subject populations: ATCOs and university students. As the university students had received no training in air traffic control, they therefore should not be strongly biased towards either 2D or 3D displays.

A number of elements of 3D command/control tasks were identified and a series of part-task tests were devised to assess subject performance using the different display types and to gain subjective data. The tasks chosen were those covering important elements of the use of displays for spatial command/control tasks, and for which the format of the display might be expected to have significant influence:

1. **Precision of Observation of Azimuth Angle and Relative Distance.** Two key factors determining whether 2D or 3D displays are suitable for ATC are the precision with which distances and angles may be determined. ATCOs are sometimes required to judge angles for issuing radar vectors and assessing aircraft tracks, and are required to determine distances so that minimum separations may not be violated.

2. **Information Extraction Speed.** Display format (2D or 3D) was expected influence time required for the following tasks which require information to be extracted from the display.

   (a) Height is represented in the 2D plan-view ATC display by a digital readout associated with each aircraft. It was expected that this would impose a greater workload on the ATCO than a 3D display as he or she has to read the digital information and then assimilate this into a 3D mental picture. However, if height information is reproduced pictorially in an integrated 3D display, this may lead to lower workload. It was therefore expected that the speed of selecting, say, the highest and lowest aircraft from a number would be faster for a 3D display than for a 2D display.

   (b) ATCOs sometimes need to determine distances purely in the horizontal plane, for example to ensure separation between aircraft and to determine distances from aircraft to navigation points, runway centrelines or airspace boundaries. The 2D plan-view display presents horizontal information without ambiguity, whereas perceptual ambiguity exists in the 3D display for all distances not in the plane perpendicular to the LOS. It was therefore expected that speed of picking out aircraft with the closest horizontal proximity would be faster for a 2D display than for a 3D display.

3. **Memory of a Scene.** It was postulated that 3D spatial relationships are assimilated faster in a 3D presentation than in a 2D presentation, and retained longer. This might have advantages for ATCOs in picking up the 'mental picture' when starting a shift, and lead to better retention in memory in case of loss of radar service. In order to test this, a task was devised where subjects

were presented with a static scene, in either 2D or 3D, for a fixed period of time, and were then asked to reproduce the scene on a piece of paper. It was expected that a 3D presentation would be recalled with greater accuracy and speed than a 2D presentation of the same scene.

4. **Conflict Detection.** One of the prime tasks of ATC is to guide aircraft safely to their destinations, and this involves preventing collisions between them. It was postulated that a 3D display may lead to faster detection of so-called *conflicts* between aircraft.

5. **Interception Task.** The above part-task tests are passive, in that they merely require the subject to read information from the display, and then compare the variation in speed or accuracy of the reading over the display types. However, the whole is often more than the sum of its parts, so an active synthetic task was introduced which required active participation of the subject. The aim was to see if the results of the part-task tests considered in isolation could apply to a task which was the combination of some of these elements. This task involved controlling an 'interceptor' to 'catch' a manœuvering target.

1.3 Scope

This thesis details an exploratory investigation into aspects of using 3D display technology for the real-time spatial control task of ATC, concentrating on possible human factors associated with the use of 3D displays. This was carried out by empirically comparing 3D displays against a 2D plan-view format similar to those in current operational use in ATC, with the aims of identifying possible strengths and pitfalls in using 3D displays and highlighting areas for further investigation.

This thesis concentrates on whether or not the characteristics of 3D displays regarding human cognitive performance and limitations makes them more or less suitable for ATC-type tasks. The aim is not to develop an optimum ATC display which might be used operationally.

Producing an optimum display is a rather complex engineering task which requires knowledge of the current operational environment and its constraints. However, in this research the view taken is that there is little point in trying to engineer such a display before addressing first the fundamental question of whether or not such displays are suitable for the task. The display formats developed may therefore have severe limitations if they were to be adopted 'as is' for operational purposes; however, they were sufficient for the purposes of this investigation.

From the outset, this research concentrated on display issues rather than on the full operational task of air traffic control. Real ATC is very complex, and simulation of a full working environment requires much expensive dedicated equipment and operators. There are many different tasks within ATC, approach radar and en-route sector control, for example, and choosing evaluation over only one task would not have been representative of other tasks. This research has therefore focused on certain task elements involving the display rather than on any specific ATC task as a whole. This approach has the advantage that studies into elements common to several ATC tasks will yield results that can be applied across a range of tasks rather than to one specific task. Also, other researchers have studied similar elements of the usage of 3D displays and so this enables comparisons to be made. By making the research more generic, it is hoped to be able to draw conclusions that will be applicable to a range of applications within and without ATC than to just a single ATC task.

Similarly, some of the issues involved in adopting 3D displays for ATC, such as modifications to procedures, training and equipment, will not be considered in depth.

Concentrating only on display aspects which can be measured has meant that some crucial parts of the ATC task involving anticipation of future traffic movements have been omitted. This is related to the formation and maintenance of the 'picture', the ATCO's internal model of the current and future state of the air traffic. However, the speed of assimilation of a scene (and thus the picture) was measured in some of the part tasks, and this might be taken as evidence that controllers may build up overall awareness more rapidly with certain types of display.

The task elements chosen have been identified by sifting through previous research and talking informally to air traffic controllers about what problems and benefits they felt a 3D display might bring. It is not claimed that the research covers all possible aspects of the use of 3D displays for ATC; that would take a much longer study. However, this thesis can be seen as a useful exploration into this area, and is expected to form a basis for such later studies.

The task elements are not covered in the same depth that a cognitive psychologist might. For example, in reading azimuth angle, a full study might involve systematically controlling several variables and possibly involve hundreds of observations per subject. Such studies are outside the scope of this thesis. Cognitive studies tend to seek to isolate completely the parameters under study. The results of these studies are often then applied to the real world in which a lot of additional factors may exist. This thesis took the approach of trying to see whether results obtained from cognitive research would be supported if they were cast into a slightly less 'sterile' domain.

The use of a 'non-expert' student population as well as a group of ATCOs may yield interesting comparisons between the two. If students give similar levels of task performance to highly trained individuals in certain display types, this could raise some interesting questions. The subjects responded to requests for volunteers, both by direct contact and by notices, and were all drawn from the college, being largely students of a 'technical' bent (either engineering or science background), probably due to the nature of the experimental work appealing more to individuals from a technically-orientated background than an 'arts' background. The range of subjects selected was therefore not a diverse section of the population as a whole, particularly in terms of age and intelligence. However, the main criterion for the 'non-expert' group was that they should be just that: Non-experts in air traffic control, and these limitations were not felt to be a severe drawback.

1.4 Contributions

This thesis makes the following contributions:

- Aspects of the mental 'picture' formed by ATCOs have been investigated and questions concerning how spatial aspects of the picture are represented, as an integrated whole or as separate horizontal and vertical dimensions, have been considered.

- A simulation of air traffic scenarios for investigating ATC displays has been developed.

- Consideration has been given to possible uses of immersive 'virtual reality' display technology in ATC, in both in radar and non-radar control applications. A novel application of elements of VR technology to ATC has been proposed.
- A comparison of novice and expert subject groups in the performance of tasks with 2D and 3D displays has been carried out. Differences in the way in which the groups perform tasks have been identified, both using the 2D display, where experts would be expected to perform better than novices due to their training, and using 3D displays, which are unfamiliar to both groups.
- An investigation has been carried out of the performance of stereoscopic 3D displays relative to both pseudo-3D and 2D displays. The results indicate that stereoscopic 3D display technology gives some performance benefits over non-stereoscopic displays and 2D displays for tasks where the 3D presentation of a scene would be expected to facilitate the task, but that 3D displays may yield worse performance than 2D displays for tasks where precise determination of horizontal information is required.
- Two possible sources of position ambiguity in 3D air traffic displays have been identified and discussed.

1.5 Overview of this Thesis

The following two chapters present the background for this research. Chapter 2 first examines the problem of Air Traffic Control, viewing it as a problem of 'flow control'. The air traffic controller's mental 'picture' is then discussed, and previous applications of 3D displays to ATC and similar fields are presented. Chapter 3 presents the theory of 3D displays, and shows implications of this theory for the design of 3D display formats.

The next two chapters describe the development of this research up to the main study. Chapter 4 describes the evolution of 2D and 3D display formats for air traffic control, to be used in the experimental comparison of 3D and 2D displays. The chapter concludes with the design of displays used in a pilot study. Chapter 5 then presents the pilot study itself, its results, and the implications for the main study.

The two chapters which follow these describe the design and implementation of the main study itself (Chapter 6) and its results (Chapter 7).

Chapter 8 presents the conclusions of the thesis, describes how the research's contributions and how it met (or failed to meet) its objectives, any implications of the research, and possible future work.

Chapter 2
Air Traffic Control

I wish there was a radar,
Sitting on the ground;
So I could see where aircraft are,
And make them go around.

Roger Bacon

2.1 Introduction

2.1.1 Overview

This section discusses some concepts of visual information displays. The task of Air Traffic Control is then introduced. Limitations of the present 2D display for ATC and some possible benefits of a 3D air traffic display are then discussed.

2.1.2 Visual Representations of Information

Digital computers are very good at manipulating large quantities of numerical data at high speed, and have become indispensable in applications such as modelling fluid flows around objects, nuclear warhead explosions, the weather, stresses in structures and other phenomena which previously could not have been examined in such great detail without complex and costly experimentation (if at all). However, no matter how great the accuracy and speed of computation or the volume of information that can be handled, much of it is useless without some means for it to be presented to and interpreted by *people*. Pressure, temperature and density values enumerated over a grid may be an efficient way for a computer to handle weather information, but humans require isobaric contours, colour-coded isotherms and wind vectors printed on a map to make sense of the data. Presentation of information requires a *display*.

In the vernacular of engineering psychology:

2.1 *displays present information about the state of a man-machine system* [SWK90]

Displays provide information for a person to perform some task. The information may be conveyed through any of the five senses, but the most commonplace display is the visual display, and the reason for this is not difficult to identify: Human vision is a highly developed and sophisticated sense, allowing us to perceive the rich and complex visual environment in which we live. The visual channel can be a highly efficient method of communication: Some concepts are more quickly and effectively understood when presented as visual images rather than, for example, by verbal description and while much of it is serial in nature, requiring fixation of the foveae on items of interest in turn, it also has parallel processing characteristics so that several things may be perceived simultaneously.

Information may be represented by a display using different codes. Different codes tend to use different aspects of human cognition when they are interpreted, and this has implications for the way in which the information is perceived and used by humans. This, in turn, has implications when designing a display to support a particular task. As an example, consider three codes of visual presentation of data: Digital (where numerals show the value of a quantity), analogue (for example, pointer-and-dial displays), and pictorial.

Digital and analogue displays are familiar to most of us. Differences between these include the precision and nature of information conveyed and the potential errors in reading them. Since the angle of a needle may be perceived 'at a glance', a value shown on a pointer-and-dial display may be more rapidly assimilated than a digital display, but a digital display may afford greater precision. Since the movement of a needle may be seen directly, a pointer-and-dial shows useful trend information where the quantity is changing, whereas it is more difficult to perceive the rate of change of a value on a digital readout, and may be impossible where the value is changing rapidly. It is easy to make a gross error reading a digital display by misreading a digit; it is less easy to misread the approximate angle of a pointer. It may be easier to spot discrepancies in a bank of dials which are supposed to be reading the same approximate value than in a bank of digital displays.

Pictures have the potential to condense information in several codes from multiple sources into a single representation. Since they are integrated, they potentially require less cognitive effort to interpret as the viewer no longer has to collect several items of information in different codes from different displays and then carry out the integration mentally. Further, pictures tend to display information in meaningful ways that are compatible with mental models of the world and present the relationships in the data clearly. However, pictures are not necessarily self-explanatory [SWK90].

These are examples of the sort of characteristics which need to be considered when designing a display.

This thesis is concerned with the visual display of 3D spatial information. There are a number of ways such information may be displayed, including:

1. Using multiple two-dimensional pictures (for example, three-view engineering and architectural drawings).

2. Representing the third dimension by a non-spatial code on a two-dimensional pictorial display (for example, representing height by a digital readout or by colour coding on a plan-view radar picture).

3. Representation by a single 3D picture.

Whilst all these methods represent 3D data, only displays utilising (3) are referred to here as *three-dimensional displays.* This thesis investigated representations of 3D spatial information for tasks involving real-time control of objects in space; in particular, the task of Air Traffic Control.

2.2 Introduction to Air Traffic Control

2.2.1 Overview

This section introduces the general task of radar control of aircraft, viewing ATC as a *flow control* problem based on studies into computer-supported cooperative working in ATC by Hughes *et al.* [HHS91, HRS92, HH93, SHRH94].

ATC is a complex system of technological, procedural, legal and other elements. There are many different facets of ATC, and different tasks in different areas (for example, approach control is a different task from en-route or ground control). A full discussion of the air traffic control system is beyond the scope of this thesis: The reader is referred to Duke [Duk92] and Graves [Gra89] for an overview of ATC in the United Kingdom, and to Nolan [Nol90] for a rather more detailed view of how ATC is carried out in the United States.

The prime task in ATC, according to the Manual of Air Traffic Services, which lays down rules and procedures for carrying out air traffic control in the UK, is:

2.2 *to maintain the safe, orderly and expeditious flow of traffic*

or, as the Harper and Hughes quote one ATCO:

2.3 *"Send 'em all to the same place and expect us to stop them hitting."* [HH93]

ATC may be viewed as basically an exercise in flow control and scheduling. Each ATCO is responsible for a particular *sector*, which is a three-dimensional chunk of airspace. Aircraft enter the sector at various positions and times and must be guided to their exit points (e.g. an aerodrome or a point at the boundary of an adjacent sector) whilst observing 2.2 above: *Safe* means that aircraft must be prevented from colliding with each other or the terrain (except in the controlled case of landing!), *orderly* implies that the flow should be organised rather than *ad hoc*, and *expeditious* means that aircraft should be guided to their destinations as quickly and as efficiently as possible, within the constraints of safety and orderliness.

ATC operates within a framework of rules and procedures to ensure quote 2.2. Some rules relate to 'separation'. Aircraft are not permitted to approach each other within prescribed minimum distances. These vary from situation to situation, but an example might be that aircraft must be separated by at least 1 000 ft vertically or 3 nm laterally (in other words, if two aircraft are within 3 nm of each other horizontally, they must be 1 000 ft or more apart vertically, and if they are closer than 1 000 ft vertically, they must be at least 3 nm apart horizontally). If aircraft conform to these minima, they are said to be *separated.* If aircraft have lost or will lose separation at some point in the near future (for example, if two aircraft at the same altitude are converging on the same point to arrive there at the same time, but are currently separated) then they are said to be in *conflict.* Each aircraft can therefore be thought of as carrying a 'bubble' of 'personal space' around with it which no other aircraft may penetrate, and conflicts are situations in which another aircraft either has penetrated the bubble or may do so. Conflicts may be divided into categories, *viz:*

1. Aircraft approaching head-on, either co-altitude, or assigned the same altitude but climbing/descending, or passing through the same altitude (Figure 2.1(a)),

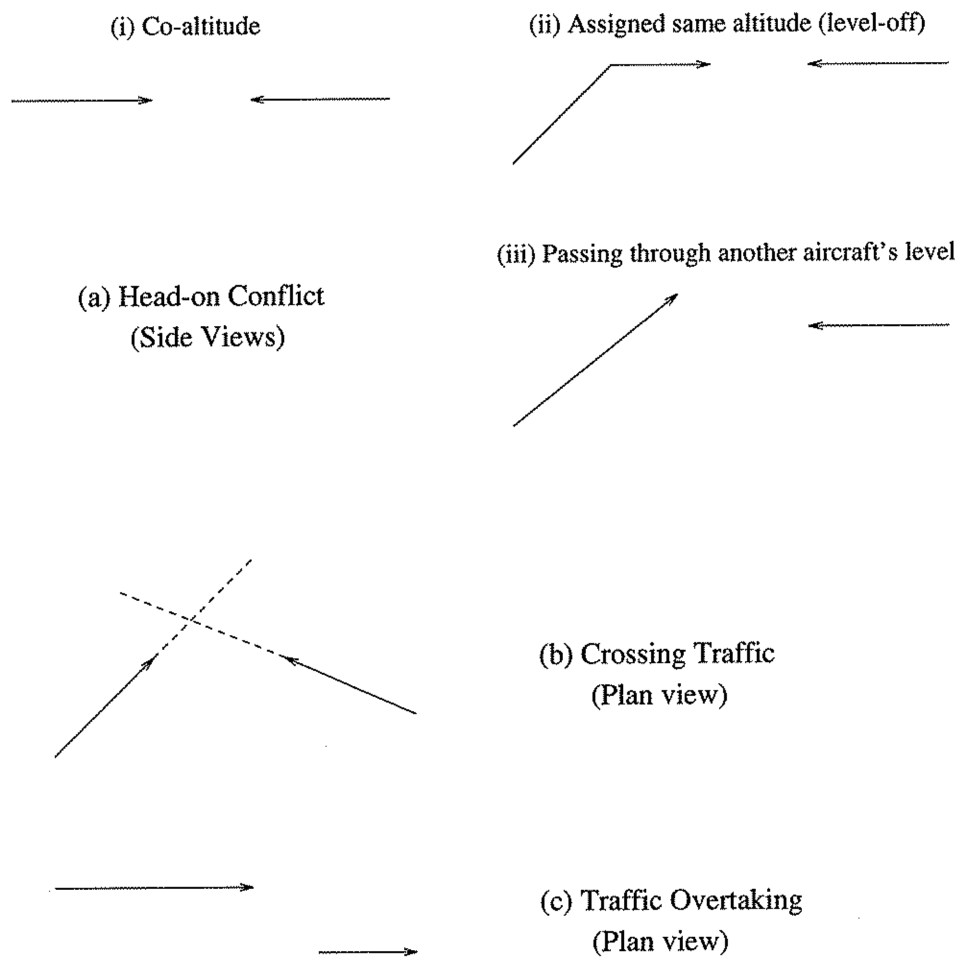


Figure 2.1: Conflict Types

2. Aircraft crossing the same geographical point (Figure 2.1(b)),

3. Aircraft overtaking on the same airway (Figure 2.1(c)).

The ATCO uses a number of tools in carrying out his/her job. The most important of these (in no particular rank order) are:

- Radio telephone (R/T) to communicate with aircraft.
- 'Land lines' to, and direct (face-to-face) contact with, other controllers.
- Flight progress strips.
- Real-time display of aircraft position (although some ATC tasks may only require an aircraft position display under certain circumstances, for example ground control). This is usually a radar-derived plan position indicator (PPI).
- Other instances of specific information (e.g. weather reports).

The flight progress strips and radar, and how they are used in control, are described briefly below.

2.2.2 Flight Progress Strips

Perhaps surprisingly to the uninitiated, the flight progress strip (or more colloquially 'flight strip' or simply 'strip') is more important in ATC than radar. (Shapiro *et al.* [SHRH94] give a detailed discussion of the rôle of the flight strip and how it fits in with the work of the controller, and Maltezos [Mal91] gives a more detailed description of the strips themselves and how they are annotated.) ATC may be carried out using flight strips and procedural separation rules with no real-time display of air traffic, as indeed was done in the past.

The flight strip comprises a piece of card, measuring approximately 1×8 inches, corresponding to an aircraft on which pertinent data are printed or written: Aircraft callsign, type, route code, speed, time expected at a navigation point, current altitude etc.[1] The ATCO has these arranged in racks near to the radar screen. The strips may be organised in the racks by, for example, sector entry point, and under each entry point by time or some other system of the ATCO's choosing. The racks

[1]Some systems are now entering service which feature 'electronic strips', using tabular electronic displays instead of physical pieces of paper. However, the principles are the same.

allow for strips to be 'cocked out', usually as an *aide mémoire* to draw attention to a strip, for example if an aircraft requires special monitoring or other action.

The strips are annotated by the ATCO as instructions are issued to the aircraft and so reflect the state of the aircraft at any given time. For example, when instructing an aircraft to descend, the ATCO will cross out the aircraft's current altitude[2] and write a descending arrow with the cleared altitude. This will be further amended when the aircraft levels off.

The strip thus represents the current state of the aircraft, and is a key element in the task of ATC, as will be seen below. It is very closely identified with the aircraft in the mind of controllers, as one ATCO related to the author *"To me, the strip is the aircraft."*

2.2.3 PPI Aircraft Position Display

Another major tool in ATC is the PPI radar display[3]; i.e. a display showing azimuth and distance of radar targets from a point (usually the radar head, although with electronic processing this can be any point within the area of radar coverage). At its most basic, this comprises a circular monochrome cathode ray tube (CRT) with a long-persistence phosphor. As the radar antenna is scanned through 360° in azimuth, the transmitter sends out radio pulses which are reflected from objects (e.g. terrain, clouds, aircraft) back to the receiver[4]. The time between a pulse being

[2]In aviation terminology, the words 'height', 'altitude' and 'flight level' are distinct: Height is defined as *the vertical distance of a level, a point or an object considered as a point measured from a specified datum*, altitude as *the vertical distance of a level, a point or an object considered as a point, measured from mean sea level* and flight level as *a surface of constant atmospheric pressure, which is related to a specified pressure datum, 1013.2 mb, and is separated from other such surfaces by specific pressure intervals* [Civ94]. Aircraft in controlled airspace usually set their altimeters to read altitude when flying below the so-called *transition level*, and to read flight level above this level. However, this thesis will not be so rigid in applying the distinctions since it is concerned with displays rather than the finer details of the ATC system.

[3]Although most PPI displays show radar-derived information, 'automatic dependent surveillance' is now undergoing trials: Accurate satellite-derived position information may be downloaded from an aircraft's navigation system via datalink into the ATC system and displayed to give a pseudo-radar PPI picture outside areas of radar coverage. The display is still a PPI, though, and so from the point of view of this research, the distinction is irrelevant.

[4]The same antenna is usually used for both transmission and receiving: The radar sends out a pulse and then listens for a while through the same antenna for echoes before transmitting the next

transmitted and an echo from a target being received gives the target's distance from the radar, and its direction from the radar is simply the azimuth of the antenna at the time. Radar returns are drawn on the CRT at the appropriate azimuth and at distances from the centre of the display depending on the time delay; thus, a plan-view picture of radar targets around the radar head is built up. Since a radar sweep can take several seconds, the long persistence phosphor of the CRT is used to retain the image over several sweeps, and this allows moving targets to be seen against a static background (since their positions change between successive sweeps of the antenna); they are shown as a bright radar return (the last 'fix') with a trail of previous returns of diminishing intensity behind them. The distance between the trailing returns enables the speed of the target to be estimated, and their direction enables the heading and any horizontal manœuvering of the target to be seen.

A raw radar display is often very cluttered, so Moving Target Indicator (MTI) processing is used to remove static targets such as terrain, cloud etc. This involves measuring the Doppler frequency shift of the echoes, which gives the component of velocity of the targets directly away from or towards the radar; static targets can thus be filtered out[5].

A *video map* showing airway structures, navigation reference points, coastlines etc. may be superimposed on the display. Concentric *range rings* may also be superimposed to enable range from the centre of the display to be estimated.

Instead of displaying MTI-processed raw returns, the position data from the radar data processing (RDP) computer can be used to show aircraft position on a *synthetic display*, using computer-generated symbology. Such information may be shown on a raster display with a short-persistence phosphor, and different colours may then be used.

The primary radar gives the azimuth and range of an aircraft from the radar head, but not its identification or height. This information is obtained using a *secondary surveillance radar* (SSR), usually located alongside the primary radar,

pulse.

[5]If the target is moving tangential to a circle centered at the radar head, its velocity component along a radial from the radar will be zero and so MTI will filter out the target so long as it continues along a tangent. Since this would involve travelling in a near-perfect circle around the radar, in practice no aircraft return is filtered out by the MTI for more than a few successive sweeps. A plot may be 'coasted' by the radar data processor for a few sweeps to counteract this.

and a *transponder* on each aircraft. The SSR sends an interrogation signal which is received by an aircraft's transponder, causing it to transmit a *mode-A* or *squawk* code (a four digit octal number) and, if the transponder is an altitude-encoding unit, a *mode-C* code which gives barometric height in hundreds of feet above the reference pressure level set by the pilot. The transponder information is received by the SSR, correlated with the primary radar position information, and displayed on the PPI next to the corresponding 'blip' as an alphanumeric *datablock* or *tag*.

Commercial aircraft generally have flight plans filed in the ATC flight data processing (FDP) computer. In this case, the aircraft's mode-A code is associated with the flight's file in the FDP system. This enables the flight number to be displayed on the PPI instead of the squawk code, and allows the FDP computer to track the progress of the flight and to issue flight strips to controllers automatically as required.

Figure 2.2 shows the radar consoles in the approach room[6] in the London Heathrow control tower. Figure 2.3 shows a time-exposure photograph of a radar tube. The static parts of the image are the video map showing the airspace structures around London Heathrow airport (at the centre of the tube), whilst the aircraft and their datablocks can be seen leaving trails as their positions are updated over the successive radar sweeps recorded during the exposure.

2.3 ATC as Flow Control: Planning

To reiterate, the ATC problem is basically one of organising the flow of air traffic through the sector. As Hopkin has observed, this is a *four*-dimensional problem [Hop92], with aircraft moving through space and time. Flow is managed by looking at the current and projected future state of the traffic and planning to deal with it. ATCOs refer to the strips and radar to build and maintain a mental 'picture' of the traffic under their control. (The picture is discussed in more detail in the next section.) Contrary to what might be expected, although the radar provides a current representation as to the disposition of traffic in the sector, it does not provide a complete picture of the situation; the 'situation at any point in time' is

[6]This is now a misnomer since Heathrow approach control is now carried out from London Air Traffic Control Centre (LATCC) at West Drayton, so these consoles are now used for 'Thames Radar' and 'London Special VFR' (Visual Flight Rules) traffic.

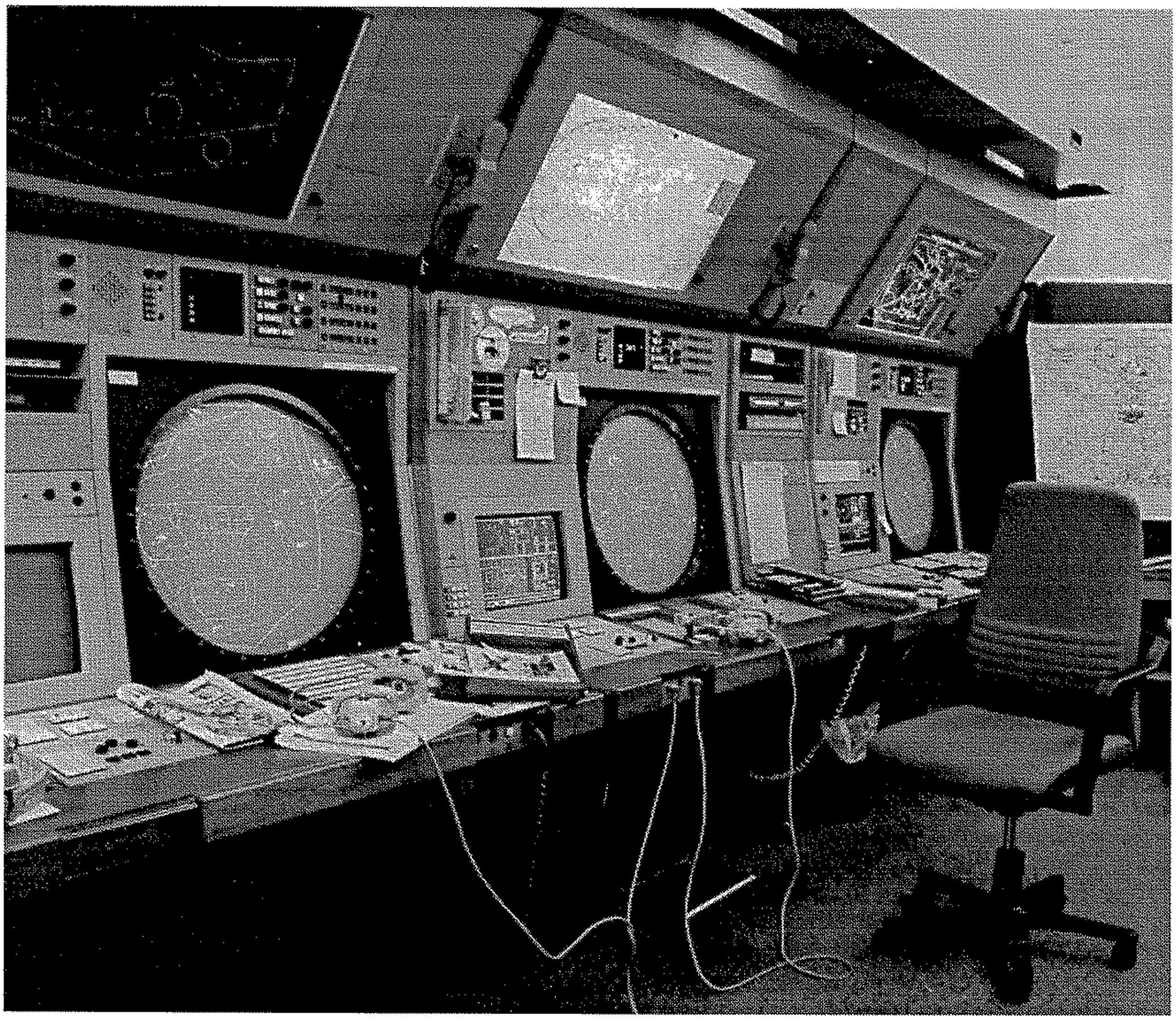


Photo: the author

Figure 2.2: Approach Control Room, Heathrow control tower

Photo: the author

Figure 2.3: Radar tube picture, showing London terminal airspace

not only what is happening 'right now' but what also might be happening in the near future [SHRH94].

The author has found it convenient to discuss the control problem in ATC in terms of 'strategic' and 'tactical' planning aspects. Although there appears to be no precedent for formally splitting the problem into these categories, this has been found to be a useful model and consistent with the author's observations of how ATCOs carry out their tasks.

Strategic planning involves looking at current and projected traffic state and formulating an appropriate strategy to deal with it. In this way, traffic peaks and other problems (such as runway changes) are anticipated and planned for in advance to ease the workload. There is usually a characteristic pattern of traffic associated with each sector and most flights follow this pattern[7]. Occasionally, however, there are aircraft whose flight plans do not conform to the pattern, for example aircraft turning between airways or crossing an airway, and the ATCO must plan accordingly.

Establishing the picture and planning involves predominantly the strips, with radar used in a supporting rôle. Receiving a strip shortly before an aircraft is due to enter his/her sector, the ATCO may note its route, altitude and time expected at a point. Checking against other strips in the pertinent racks usually reveals whether or not there is likely to be a problem. Having identified a potential problem, the ATCO can plan to deal with it, and may 'cock' any relevant strips out of the rack as a reminder to monitor certain aircraft more closely. Strips are the instrument that helps to organise, and so enable, controlling work, and are referred to and manipulated constantly whilst they are in use [SHRH94]. As one ATCO related to the author, air traffic control is a matter of 'working the strips'.

The radar tends to be used in conjunction with the strips to monitor the current state of the traffic and the execution of the plan (and to modify it as necessary), and for tactical planning, e.g. adjusting headings or levels to avoid other traffic. For example, if an aircraft must descend through the level of another crossing its path,

[7]A large part of qualifying an ATCO for controlling a particular sector appears to be familiarisation with its characteristic patterns and procedures. Once the pattern is learned, things that 'do not fit' appear to be quickly identified: The author has observed an ATCO who used to handle Heathrow departures but currently working a different sector spot a minor controller error on the radar even though he was not working the traffic concerned or looking for it, because it did not 'look right'.

the ATCO may decide initially to descend it to a level above that of the crossing traffic and to monitor the situation on radar. As the aircraft get closer, the radar picture helps the ATCO to decide whether it is safe to continue the descent through the level of the crossing traffic (if sufficient lateral separation exists) or whether to delay its descent until the aircraft have crossed.

Radar is also used for overall monitoring. Human and machine behaviour are not perfect and there may be a lapse on the part of the controller, the pilot or the aircraft's automatic flight control system which may bring the aircraft into direct conflict which will not be spotted from the flight strips. An 'altitude bust' is a good example of this—an aircraft climbing from below and converging head-on with another aircraft may be ordered to level off below it, but may 'bust' through its cleared altitude and create a conflict. For this, the only means of controller detection (apart from reports of an airmiss or a crash) is the use of the radar and automatic conflict alert systems (although these have their own limitations).

At times of heavy workload, the emphasis of usage tends to be on the radar; at other times, the emphasis appears to be more on the use of the strips [HRS92].

Some controllers have expressed to the author the concept of 'foreground' and 'background' traffic. If, for example, an aircraft is proceeding along an airway at a constant level, with no conflicting traffic and is likely to require little action on the part of the controller, it tends to get relegated to the 'background'—checked periodically as with all other traffic, but not at high priority. Aircraft which the controller is more actively involved in directing, for example which are requesting or are in the process of climb and descent or are manœuvering, tend to be more 'in the foreground' and so more attention is given to these aircraft. This behaviour was also reported by Whitfield and Jackson [WJ82].

2.4 The Picture

Contrary to what might be expected, a large part of ATC is working with the flight strips supported by the radar display to build up the 'picture' rather than just concentrated on the radar. The picture, in Harper and Hughes's view is not so much a pictorial representation of the current traffic disposition but a 'display of a set of task requirements': The ATCO looks at the information presented by the strips, radar and R/T to see what needs doing 'now', 'in a moment', 'sometime later on' and so forth [HH93]. In other words, it is the ATCO's mental model of the plan, its state of execution and the general disposition of traffic.

An ATCO coming on shift may spend a period of time watching the situation over the shoulder of his or her outgoing colleague 'to build up the picture' before taking over[8], and an important part of the task is looking ahead constantly in order to get an idea of what is to be done 'now' [SHRH94]. It should be stressed, however, that the plan is continually evolving and although in general an ATCO knows what he or she is doing, it is not known in advance in detail [SHRH94]. The 'picture' is summarised by Whitfield and Jackson as 'overall appreciation of the traffic situation for which they are responsible' [WJ82]. Harper and Hughes refer to the picture as referring, *"among other things, to the controller's capacity to 'keep it all together'; to see and give coherence and organisation to the patterns of aircraft movements under varying conditions."* [HH93]. To quote an air traffic controller:

2.4 *"The name of the game is to get your priorities right, and... background those who are not in a conflict situation at the moment. A key factor in ATC is also to plan ahead and then decide what could become a conflict and taking appropriate action."*

'Losing the picture' is one of the ATCOs worst fears—having the mental picture enables the controller to be 'on the ball'. If the formation and maintenance of the picture is disrupted, the situation rapidly degenerates into 'firefighting'—retroactive action instead of pro-active planning, with a large increase in workload and reduction in level of safety. It is therefore no exaggeration to say that maintaining the picture is central to the task of ATC. As a controller has related:

2.5 *"My own thumb-rule in deciding if I have the 'picture' is that if I dare not take my eyes off the screen and have a look at my progress-strips, I am becoming saturated and need assistance."*

Although the controller's 'picture' may be described in general terms, eliciting its nature more precisely, how it is represented in the mind of the controller as a

[8]In relatively 'quiet' sectors, the author has observed the outgoing ATCO describing the current state of affairs to his incoming colleague; which aircraft are going where, what is pending and special points of note. In busy sectors, such detailed verbal interchanges may not be possible or adequate.

'mental model', is more problematic. The studies of Hughes *et al.* suggest that strips have a key rôle in building and maintaining the picture, and Whitfield and Jackson's [WJ82] analysis of protocols during handing over aircraft from one sector to another in an attempt to elicit the 'picture' found flight strips to be a predominant element, with radar occurring much less frequently but strongly linked with flight strips. The author has not been able to identify a study looking at the rôle of the radar in the forming and maintaining the picture. However, Strutt states his assumption that:

2.6 *"...the radar display is the prime mode of control used by the controller. The display image is used to build up in the controller's mind the flights under his/her jurisdiction not as a list but as a picture so that he/she is constantly aware of each flight's spatial relationship with all the others."* [Str91b]

This is not based on a formal study, but Strutt is himself an ex-UK air traffic controller who has worked in a radar environment. Burnett (an ex-US air traffic controller) seems to concur with this view:

2.7 *"The display issues with which controllers are chiefly concerned pertain to... primarily, the time and work required to translate 2D alphanumerics into a 3D mental model of the traffic situation that continuously changes."* [Bur91]

She also refers to other studies which have shown that:

2.8 *"Controllers make from 100 to 150 decisions per minute and each decision is based almost entirely on the information presented on the PVD [plan-view display]."* [Bur91]

Although the research into how controllers work seems contradictory, therefore (possibly due to differences between the US and UK systems) one thing is clear: The spatial aspects of the controller's mental picture have not been satisfactorily explored to date. In the absence of further evidence on the importance of radar related to the controller's picture, the author postulates that although the mental model which controllers have is not merely a pictorial snapshot of the present spatial relationships of aircraft in the sky, it contains some spatial elements which are built and maintained through the use of the display, and such spatial elements must include a three-dimensional aspect. However, if there is a spatial model, it may not be an accurate *integrated* spatial picture due to the difficulty in realising such an image from a pictorial plus digital code. Because the horizontal and vertical are represented as different codes (with the vertical non-pictorially), it is difficult for ATCOs to visualise vertical manœuvres.

This may be not so much of a problem at present because although current equipment cannot display an integrated 3D image of the horizontal and vertical traffic situation, the current system has evolved with these limitations and so in some ways the plan view radar display is adequate for the ATC task because the system has been designed so. Horizontal and vertical planes tend to be treated not in combination but separately. For example, if two aircraft are separated vertically and not climbing or descending, then their precise horizontal location is not of great concern; if they are all following an airway with the same heading and speed and are separated horizontally, then altitude is not of great concern. This raises the question 'If the picture is not an integrated three-dimensional spatial image, what use is a 3D radar?'. Further 'What can a 3D radar display offer that might aid the formation and maintenance of the picture?'.

The following discussion will argue that there may be limitations of present equipment not being able to represent clearly what is happening in the vertical plane, and will touch upon other potential advantages of a 3D air traffic display. As was stated in Chapter 1, strategic planning aspects are beyond the scope of this research due to the complex nature of the study which would have to be carried out to study them. Discussions will therefore not concentrate on this issue (although it will be touched upon) but on what might be achieved on the tactical planning front.

2.5 The Vertical Dimension

A current limitation of ATC equipment, which the author has observed both from the point of view of the controller and the pilot, is the difficulty controllers seem to experience visualising manœuvres which involve a vertical element.

2.5.1 The Controller's Problem

From the controller's perspective, projecting flight paths mentally to predict, for example, which geographic point a descending aircraft will reach at current rate of descent and forward speed, say, 2 000 ft below present altitude, appears to be difficult. Usually there is no rate of altitude change information presented on the PPI apart from observing the changes in the digital height readouts and, as discussed at the beginning of this chapter, rate information is difficult to perceive from a digital display. Controllers are therefore conservative and tend to adopt the approach described above: When descending an aircraft across the path of another, clear it initially to descend to a safe level above the potentially conflicting traffic, monitor the situation on the radar and as the aircraft get closer, make a judgment as to whether it is safe to continue the descent or to defer it until the aircraft have crossed.

To assist in making this sort of prediction, some equipment can display 'prediction lines' which indicate the position at the aircraft at up to 5 minutes in the future based on extrapolation of present speed and heading. As an Oslo controller related to the author:

2.9 *"This [prediction lines] I use all the time to determine what I can do, and what I can't. You normally confirm/specify a ROC/ROD [rate of climb/rate of descent] to the aircraft and it's just a task of multiplying."*

New equipment being installed at Oslo also has the capability of displaying vertical rates in the datablock, and the facility to extrapolate and display the point of minimum separation between two selected aircraft.

Although these aids no doubt help considerably in making decisions involving the vertical dimension, the presentation is still plan-view, and there is still a workload involved in making vertical judgments ('a task of multiplying', or the process of selecting aircraft for minimum separation prediction) which might be eliminated or reduced with a three-dimensional display, since the flight paths would be seen directly in 3D space and this might simplify mental extrapolation. The author has been shown predictor lines on equipment at LATCC, but then observed that most controllers did not leave them selected 'on', possibly due to the additional clutter and the additional workload that might accrue to interpreting a cluttered display. Such penalties might be acceptable where traffic load is not high, but perhaps not where there are a lot of aircraft displayed.

2.5.2 The Pilot's Problem

From the pilot's perspective, in a discussion with a commercial pilot it was stated that ATC sometimes made difficult demands for vertical manœuvres which demonstrated a lack of appreciation of the vertical performance of aircraft. An impromptu demonstration then followed where the controller at our destination aerodrome requested a high rate of descent to reach a certain altitude by a certain time, which necessitated the pilot having to deploy speedbrakes; all the while, the primary flight display was showing that the aircraft was above the required descent profile and was having difficulty in achieving it[9]. However, a controller counters that the main problem is that vertical performance varies greatly, depending on weight, winds aloft, temperature etc. and so it isn't always easy to know what to expect. This is exacerbated by the fact that airlines operate their aircraft differently regarding descent profiles, speeds etc. Despite this, it is postulated that with a 3D presentation, controllers might get a better 'feel' for the performance ranges of particular aircraft, since vertical performance would be visualised pictorially (as a flight path angle).

2.5.3 Summary

Older radar equipment lacks predictor facilities and a display of vertical rate information, and so predicting flight paths using such equipment is difficult and this, in the opinion of the author, leads to conservative decision-making on the part of ATCOs. More modern equipment is addressing some of the limitations of previous displays by providing assistance in flight path prediction and more information about aircraft vertical rates. However, there is still no direct visualisation of the vertical dimension, which the author postulates might improve matters still further by reducing workload and giving even greater appreciation of the vertical.

If greater appreciation of vertical manœuvres could be achieved, more efficient usage of airspace might result instead of the current conservative practices: For example, narrower descent 'corridors' might be visualised and an ATCO might be better able to see if aircraft were danger of violating these. Moreover, ATCOs might better appreciate the vertical performance of aircraft and issue instructions with which aircraft are more capable of complying. As a corollary, it might also be easier for controllers to identify conflicts between aircraft where one or more of the conflicting aircraft are climbing or descending.

[9]The aircraft was a Boeing 737-500; Boeing 737s are commonly-used short–medium range jet transports, and controllers in Europe and the US handle them daily.

2.6 Three-Dimensional Displays for ATC

2.6.1 Introduction

A display may be required which imparts to the controller a better understanding of the vertical dimension and a more integrated spatial mental representation. A pictorial display may be the solution since it has the potential for displaying spatial relationships clearly.

This section reviews other research into applications of perspective 3D displays to air traffic control and related areas. It focuses on research into the applications rather than on issues associated with the displays themselves, which are discussed in the next chapter.

2.6.2 Three-Dimensional Air Traffic Displays

Some of the potential benefits of direct visualisation of the vertical dimension have already been discussed above. However, there are a number of other possible benefits which might be derived from a 3D display of aircraft position.

Three pieces of research which have addressed the application of 3D displays to ATC from a general ATC, rather than human factors, viewpoint are Strutt's implementation of orthogonal 2D and 3D displays, Burnett's experimental comparison of plan-view and perspective ATC radar displays [BB91, Bur91], and Tam's implementation of an integrated 3D ATC display [Tam93]. These are discussed briefly below. Both Strutt and Burnett have operational ATC experience and so their insights are particularly interesting.

Strutt's Four-View Implementation

Strutt implemented a display comprising three orthogonal 2D views (one plan and two side views) and an integrated 3D view to demonstrate the feasibility of a 3D radar picture. He contends that, further to his assumption about the rôle of the display image in forming the picture (quote 2.6 above), the ATCO will be further assisted by such a display. This might fix the 'picture' in the controller's mind, and may reduce mental integration workload with possible benefits in allowing more time to be spent on other control tasks, improving efficiency. He cites two reasons for non-operational implementation of 3D displays in ATC to date:

1. A PPI display can convey a 3D picture to the controller by virtue of the digital height readout.

2. When shown previously demonstrated 3D representations, controllers have become disorientated. Strutt suggests that this is due to image inversion (ambiguity caused by lack of depth information in a 2D representation of a 3D object or scene).

Burnett's Empirical Studies

Burnett's research (described in full in [Bur91] and summarised in [BB91]) analysed the effects of multiple colours, information density and traffic complexity on plan-view (2D PPI) and perspective (3D pictorial) displays, using 'full-performance' air traffic controllers[10], using a survey and experiments focusing on 'cognitive workload, task and information requirements, colour preferences and traffic workload relative to traffic density and complexity'.

Burnett asserts that the 2D presentation requires operators 'to unlearn years of 3D conditioning and, consequently, learn new responses in order to use 2D interfaces in the work environment', and so 'controllers must develop new techniques that help them translate 2D information into a 3D mental model'. Since a three-dimensional representation is more 'natural' and 'intuitive' (to use the hackneyed adjectives once more), a 3D display might well contribute to lower workload associated with interpreting the spatial information, increasing efficiency and safety, and this is especially important in the light of ever increasing volumes of traffic which controllers are required to handle.

Burnett states that the primary interests to the controllers are vertical and lateral distances and separations between aircraft, and that a perspective display would allow controllers to 'see and interpret' these. She cites other potential advantages as being able to display ground features (such as 'trees, grass, runways, and buildings') in 'real-world representation and color', and the fact that certain types of information might be interpreted more quickly in a 3D representation than on a PPI, for example, an object may be identified as 'near or far, high or low, ascending or descending'.

[10]Burnett defines full-performance controllers as those 'who perform duties in every position independent of a trainer and who require no further training'.

The experimental part of her study measured times to perform three tasks: (1) resolution of impending conflicts, (2) identify callsigns of highest and lowest aircraft in a scenario, and (3) reconstruction of air traffic situations. The experiments found that performance using the perspective display was at least equal to and in some cases better than the plan view display. Subjects expressed a preference for the 3D perspective display over the PPI display for extracting immediate spatial situational and directional information, but indicated that because of their extensive training on PPI displays, they were almost entirely dependent on datablocks for extracting speed and altitude (US equipment displays aircraft ground speed as well as altitude) and that changing the display format would necessitate extensive training.

Burnett's experimental study therefore seems to support the 3D perspective format display, both in terms of performance at the tasks (with the caveat that the tasks were not a simulation of actual working conditions) and in terms of being preferred by the subjects. She highlights further areas for future study, mostly to do with optimising viewing parameters (position of the centre of projection, geometric field of view, vertical and horizontal scaling factors necessary to minimise distortions in spatial information) and colour-coding.

Tam's 'Flight Tube' Demonstrator

Tam implemented a three-dimensional display for visualising a new air traffic control concept: 'Tubes of flight' [Tam93]. In this concept, instead of aircraft using fixed 'routes', an individual 'tube', a tunnel in the sky, is assigned to each aircraft so that each flies within its own restricted area. Under the tubes of flight concept, the rôle of the controller would change, with tasks divided up functionally rather than geographically. Two such functions proposed are for 'Strategic Flow Planning' controllers and 'Tube Generation and Assignment' controllers.

Tam's suggestion for a display console for such a system is a PPI display with a '3D interactive' screen which can render a selected portion of airspace in 3D and can be manipulated (rotated, scaled etc.) independently of the PPI. Unfortunately, the display was found to be extremely cluttered.

The 'Tubes of Flight' concept is radical in that it is a new system of control entirely different from the way in which ATC is performed today; a revolutionary rather than an evolutionary idea. Tam's work concentrates mainly on the implementation of a display for such a system, and was not evaluated. Consequently, it does not address the possibilities of applying 3D displays to the current ATC system in an evolutionary approach, and does not address human factors requirements.

2.6.3 Other Related Research

Much research work in applying 3D displays to aerospace has concentrated not on the air traffic control problem, but on cockpit displays to assist the pilot. The task a pilot performs is very different from that of an ATCO, and a display which is suitable for the pilot may well not be suitable for the ATCO. Despite this, there are many similar issues between the two fields, and so the most relevant research for the ATC problem will be described here.

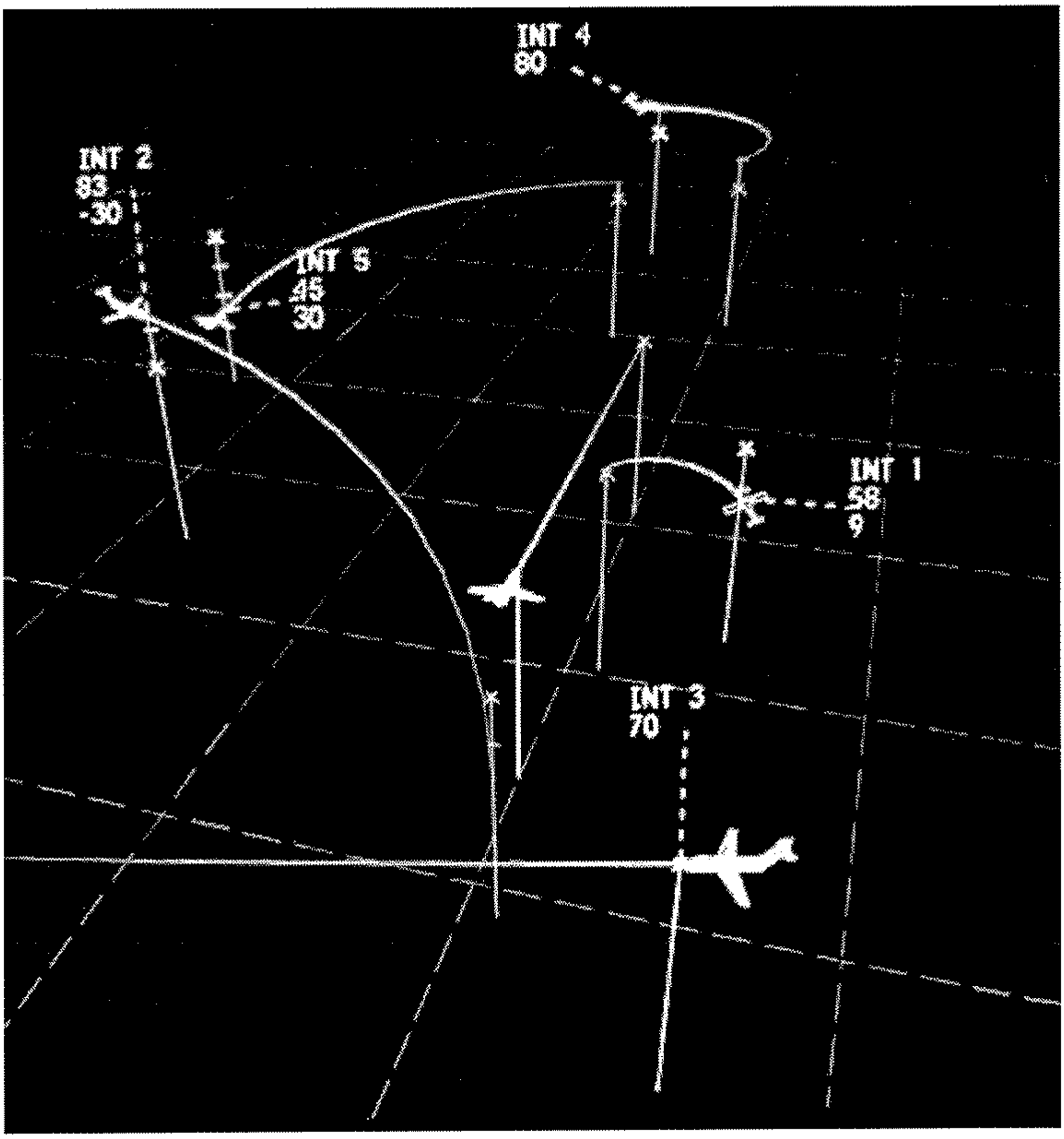
Cockpit Display of Traffic Information (CDTI)

Ellis, McGreevy and Hitchcock [EMH84] compared a perspective and plan-view display for presenting traffic separation information to pilots. The perspective display is shown in Figure 2.4. They hypothesised that because the perspective format was more closely aligned with a pilot's situational awareness in space, performance at tasks such as detecting and resolving conflicts would be improved over the use of a plan-view display with height information presented by digital readouts.

Pilots were presented with situations in which they had to carry out traffic avoidance manœuvres. It was found that such manœuvres were made 'somewhat earlier' and more frequently in the vertical dimension with the perspective presentation. Fewer unsuccessful manœuvres and fewer manœuvres producing spacing violations were found with the perspective format than with the plan-view format.

These findings are significant because they suggest that the format of presentation of vertical information influences subject behaviour; a previously observed bias towards horizontal avoidance manœuvres might be due to the poorer visualisation of the vertical dimension. Further, it shows reduced decision time, perhaps due to elimination of the need to use datablocks when reading height information.

The format of 3D display developed by Ellis *et al.* served as inspiration for those developed by Strutt, Burnett & Barfield and also for this research.



Reproduced with kind permission of S. Ellis

Figure 2.4: Perspective Cockpit Display of Traffic Information

Tactical Information Displays

Bemis, Leeds and Wiener compared plan-view and perspective display formats for a tactical display to determine the significance of vertical information on the tasks of threat detection and interceptor selection [SVBW88]. It was hypothesised that a perspective display would reduce operator information-processing load and reaction time, and that there would be no significant differences in performance between the two display formats for search and detection tasks. From experiments, Bemis *et al.* found that subjects using the perspective display made significantly fewer errors for all tasks, and unexpectedly, times taken to perform the threat detection and interception tasks were decreased. Nineteen out of the 22 subjects reported that they preferred the perspective display, with the remaining three preferring the plan view display 'because they were familiar with it'.

2.6.4 Three-Dimensional Displays and Memory

Some of the above studies have postulated that an integrated presentation of all three spatial dimensions in a pictorial format might make the spatial information easier to 'assimilate', and might assist in the formation of the air traffic controller's 'picture'. The author further postulates that presenting spatial information in a picture may also make *recall* of spatial information easier and possibly faster and the memory of spatial relationships more persistent (this might have advantages in cases such as loss of radar service, for example).

Human memory is a complex phenomenon with a number of different characteristics, often modelled as different types of memory [Bad82]. The type of memory under consideration here is known as *working memory.* As described by Wickens and Flach [WF88], working memory is employed when a person hears a number and must enter it on a keyboard, or when he/she must recall the relative positions of blips on a radar display after a brief scan. These examples illustrate two different *codes* in working memory; verbal information is normally retained using an acoustic-phonetic rehearsal, and spatial information is normally retained in working memory using a visual code. Wickens and Flach also state that there is evidence to suggest that visual codes are less easily rehearsed than verbal codes. Regarding spatial codes, evidence shows that analogue pictures are the most useful mode for storage in working memory, print the least [Lea94].

Another thing which might merit investigation is the effect of display format on what happens when the subject is 'overloaded' with too much information to remember. The number of items that can be retained in working memory is quite small, the most oft-quoted figure being 7±2 items. If more items are presented, the number of individual items which can be retained may be increased by *chunking*, consolidating items into related clusters. Working memory is more likely to treat dimensions of a single object as a single chunk than the same dimension of several objects. For example, in an ATC problem, altitude, airspeed, heading and size of two aircraft would be retained better than the altitude and airspeed of four aircraft, even though in each case eight items are to be held in working memory.

To summarise, an investigation of the effect of display format on memory of an air traffic situation display might consider the following:

1. Does presenting spatial information in an integrated 3D picture (a single visual code) rather than in two codes (2D visual picture and textual) affect:
   (a) Speed of assimilation;
   (b) Speed of recall;
   (c) Accuracy of recall;
   (d) Retention time in memory?
2. Can a person chunk information presented in different codes, such as a spatial image of aircraft position with text relating to callsign and height? If not, which codes are lost from memory first?

Burnett and Barfield's study included a task looking at the effect of display format on memory [BB91]. This compared subject performance at reconstructing a scenario (in terms of scenario reconstruction time and accuracy) between 2D plan view and 3D perspective ATC displays for two different densities of traffic. Subjects (who were all air traffic controllers) were first required to memorise flight strip information. They then watched a display of the associated scenario over 70 s of animation (in either 2D plan-view or pseudo-3D perspective formats), and were then required to reconstruct the final frame of the scenario, giving position, callsigns, altitude, speed and heading of aircraft. Two traffic densities were used: Light (7 aircraft) and heavy (17 aircraft). It was found that performance for the reconstruction task in the both traffic scenarios was about the same for both display types;

however, in the perspective display format, aircraft placement was found to be consistently 3 cm north of the actual aircraft location, whereas horizontal placement was accurate within $\frac{1}{2}$cm of actual aircraft position. Burnett and Barfield's study served as the basis for a similar investigation in this thesis.

2.6.5 Summary

The main benefit of a 3D display for ATC proposed by Strutt and Burnett, (and also by Ellis, McGreevy and Hitchcock for the cockpit display of traffic information) can be summarised by re-iterating the statement of §1.1.1: By presenting spatial information in a three-dimensional pictorial form, the characteristics of human visual perception may be exploited to yield a more 'natural' and 'intuitive' presentation of the information. Some possible implications are:

1. 3D displays might reduce the mental integration workload from that presently required by the current displays, such that more time can be spent on other vital tasks. This might improve safety and efficiency.

2. A three-dimensional display might assist the formation of the air traffic controller's 'picture', and possibly better retention of spatial information in cases of loss of radar service.

3. Spatial relationships might be more easily seen and more quickly identified.

4. The enhanced appreciation of the vertical dimension might allow controllers to make more efficient decisions, to be able to identify conflicts more quickly and to make more realistic demands on pilots.

5. Aircraft not conforming to the characteristic pattern of flow might be more easily identified. Aircraft tracks not following the normal routes may be easy to spot on a plan view display, but an irregularly high or low aircraft, or one which is climbing or descending slowly, might not so easily be observed.

In the author's opinion, Burnett's asserted advantage of a 3D display being able to present information such as 'trees, grass, runways, and buildings' in 'real-world representation and color' is dubious; the author has found no evidence, either from observing ATCOs at work or from the literature, that such information would bring any benefits. Some form of terrain visualisation might be desirable in certain applications (e.g. terminal areas in mountainous regions) to help prevent controlled flight into terrain accidents, but would add to display clutter.

Despite the apparent advantages, there are a number of issues to be addressed regarding 3D displays. These include:

1. The problem of perceptual ambiguity regarding distances along the display line of sight. Can distances and angles be interpreted with sufficient accuracy to enable the ATCO to carry out his/her tasks?

2. Choice of viewing parameters. These can introduce biases in interpreting dimensions and angles.

3. Methods of correcting or compensating biases, including compensatory distortions, training or symbolic enhancements [EH90].

4. Display clutter. A perspective projection results in increased clutter as objects move closer to the horizon. Any symbolic enhancements employed to compensate for biases may also contribute to clutter.

5. The need for additional training and re-training for operators working with 3D displays.

In addition, Strutt cited two reasons for the non-acceptance of 3D displays for operational purposes to date. One is the fact that current displays do convey 3D information by representing height numerically. While this is true, such displays do not present an integrated spatial representation; the construction into a spatial 3D representation must be done in the mind of the controller, and this leads to poor appreciation of the vertical dimension. This is supported by Ellis *et al.*'s cockpit display of traffic information research, and has drawbacks already mentioned in §2.5 above. The second reason cited was that previous displays may have led to disorientation, possibly due to image inversion caused by insufficient depth cues. One way of tackling this problem may simply be to enhance the sense of depth in the display by increasing the number of depth cues and/or employing more effective cues such as stereopsis.

Some of the disadvantages of 3D displays may be offset by taking Strutt's approach of presenting multiple orthogonal 2D views as well as an integrated 3D view.

Using purely orthogonal views has the advantages of retaining the familiar plan view (and Strutt argues that this might be more acceptable to ATCOs than dispensing with the plan view) and of not exhibiting the perceptual ambiguities associated with a 3D display, but this carries the disadvantage that the controller may have to refer potentially to up to three data sources to extract the required information. With Strutt's proposal of combination, the ATCO might be expected to refer to the 3D display for spatial awareness, then to the orthogonal views if a more precise 'check reading' of any dimension is required. However, multiple views sharing the same display device reduce the resolution of each, and multiple views on multiple display devices increase the required display area. If the integrated 3D picture alone is sufficient, then multiple views may be dispensed with.

2.7 This Research

This thesis addresses the application of integrated 3D pictorial displays to air traffic control, and tries to tackle some of the issues raised above. It builds on the research cited above, and extends Burnett's work in particular, since this was the only study found which compared performance of controllers between a 2D PPI and 3D pictorial display. From Burnett's study, the tasks of conflict detection, visual search and a variant of the memory task are adopted. This will enable this research to verify and amplify on the results previously obtained. In addition, new aspects are introduced extending the work above:

1. A task investigating the effect of display type on reading azimuth angles and distances is introduced. This is to supplement the work of Ellis *et al.* in this field, and to try to gauge the magnitude of errors to determine whether the accuracy is adequate for air traffic control.

2. A preliminary investigation was made of the effects of different types of projection on the reading of azimuth angle and relative distance.

3. An active 'chaser' task was introduced. This looked at the effect of display format on selection and direct control of the trajectory of an object space.

4. The effect of adding stereopsis to a pseudo-3D display was investigated.

5. A comparison was made between two populations; individuals untrained in air traffic control, and currently operational ATCOs. This was to assess the effect of training biases for the two-dimensional display format.

6. Evaluation of a virtual reality display.

Before this research can be examined in more detail, however, it is first necessary to introduce some of the theory of three-dimensional displays. This will be discussed in the next chapter.

Chapter 3
Three-Dimensional Displays

3.1 Introduction

This chapter introduces 3D displays: The definition of 3D displays and the types that exist, and perceptual issues relating to them.

It is beyond the scope of this thesis to discuss 3D displays in great detail. However, Wickens, Todd and Seidler of the Aviation Research Laboratory at the University of Illinois have carried out a comprehensive survey of research into perception, technologies for implementation, and applications of 3D displays [WTS89], and the reader is referred to this for a more in-depth discussion.

3.1.1 What is a Three-Dimensional Display?

Before talking about 3D displays in detail, it is necessary to define what such a display actually is. In this thesis, the definition of a 3D display is taken from the definition of Wickens, Todd and Seidler, paraphrased thus:

3.1 *A three-dimensional display is defined as one which uses any technique, whether stereoscopy or any of the cues that artists build into a perspective painting, to create a sense of depth along the line of sight into the display.*

Distance information along the LOS into the display (also referred to as the *depth-* or *z-axis* of the display) is known as *depth.*

Depth information in a scene is conveyed by *depth cues.* These may be monocular cues as employed in painting and drawing (e.g. linear and motion perspective, texture gradients, aerial perspective, relative size, shading, and interposition), muscular sensation cues from the various muscles associated with control of the visual system (accommodation and convergence), and binocular cues, which result from the use of two separated eyes (convergence and binocular disparity). Binocular disparity is of particular interest, as it gives a form of depth perception known as *stereopsis* [Way88, Way89] which, unlike monocular cues, affords the sense of 'real-world' depth.

Displays employing solely monocular cues are defined here as *pseudo-3D displays*. Displays displaying a 3D scene with compelling 'real-world' depth, such as those incorporating retinal disparity, are referred to here as *true 3D displays*.

3.1.2 Virtual Environments and Virtual Realities

A computer can represent and simulate a 'virtual environment' (VE). Examples of VEs include geometrical models (such as an architectural databases), interactive 'adventure games' and dynamic simulations such as atmospheric models and flight simulators. A display may be classified by the relationship between the virtual environment and the viewer's real environment.

Through-the-window. (TTW) (Dissociated perspective.) This refers to the type of display where the VE is viewed 'through the window' of a computer screen [SD91]. The VE and the real world remain conceptually separate; the viewer is conceptually on the outside looking in.

Immersive. (Associated perspective.) This type of display typically employs a binocular HMD to present the viewer with a stereoscopic view of a head-slaved VE. This presents a view of the VE with a compelling sense of depth which fills a substantial portion of the viewer's visual field and moves correspondingly with the movement of viewer's head. The effect is to 'immerse' the viewer in the VE, the viewer becoming conceptually part of it.

The VE in the Real World. This is the opposite of immersion, where the VE is brought into the real world of the viewer. This might be displaying an image in a self-contained area of real space (for example, vibrating mirror displays and the fictional display of the 'Death Star' in the film *Star Wars*), or 'overlaying' additional information on the real world (e.g. head-up and helmet-mounted displays in aircraft).

Most 3D displays currently extant are of the TTW type; the viewer's real world and the VE are conceptually separate and the effect is of the viewer peering into the virtual world through the 'window' of the display device as an 'outsider'. However, the immersive type of display allows the viewer to be 'brought into' the VE and to be 'present' within it. (Indeed, immersive VEs incorporate some representation of the viewer. At its simplest, this might merely be head position and orientation. A more complex representation might include a virtual 'body'. The viewer becomes not a passive outside entity but an active element of the VE itself.) The proposed advantage of such a display is that since humans are evolved to work in a three-dimensional spatial environment with information coming in through five senses, then such an environment used for display and interaction, providing appropriate sensory stimuli and reacting according to the viewer's actions in a 'natural' way, will provide more 'natural' and 'intuitive' ways of performing tasks using a computer. The VE is not merely an environment but for the viewer (or to use a more appropriate term, the *participant*) becomes a virtual 'reality'.

It could be argued that an immersive ATC display, by presenting information in a manner even more compatible with how we perceive the real world and allowing 'natural' interaction styles, might further reduce workload required to interpret the spatial relationships in the scene and to interact with it. However, the technology of VR and research into human factors such as the effects of long-term immersion are still in their infancy. Many issues would have to be addressed and the current technical state-of-the-art would have to advance substantially before such a display could be adopted for ATC even if it were found that immersion yielded significant advantages as a method of displaying position information and for interacting with aircraft. This thesis made an initial exploration of this area by showing air traffic controllers an air traffic visualisation using current immersive technology and inviting comments.

3.2 Representing a 3D Scene on a 2D Surface

3.2.1 Projection

In computer graphics, the main output device is most commonly a 2D display such as a CRT or LCD panel. In order for a 3D scene to be displayed, it must be *projected* onto the 2D image display surface. A 3D display of a scene which has been constructed through such projections is sometimes termed a *three-dimensional perspective display*. Methods for projection are well understood and so will not be described here in great depth. The reader is referred to Foley, Van Dam, Feiner and Hughes [FVFH90] for more detailed discussions of the virtual camera model used in such projections. However, the basic principles (with some simplifications) will be described here, using terminology based on the ACM CORE camera model.

A three-dimensional scene is typically represented inside a computer as 3D geometrical information using a Cartesian frame of reference known as *world coordinates* (WC). This must be transformed and projected onto the screen (or, to be more precise, onto a rectangular area of the screen known as the *viewport*).

The projection geometry is shown in Figure 3.1. Conceptually, a virtual eye-

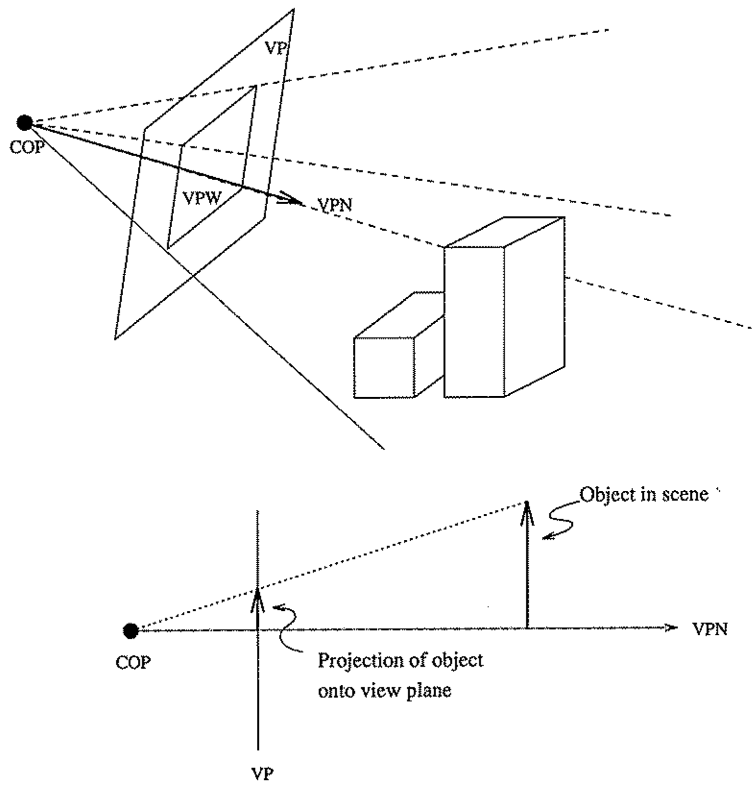


Figure 3.1: Projection Geometry

point/camera 'views' the scene from a certain position and orientation. The virtual camera is located at a point in WC space known as the *centre of projection* (COP). The line of sight into the scene is called the *View Plane Normal* and this, along with a *view up vector*, serves to orientate the camera in space. Objects in the scene are projected onto a plane called the *view plane*, which is an infinite plane orthogonal to the VPN. A rectangular area on the viewplane, called the *view plane window*, is defined to map to the viewport and so the projections of objects which fall into the view plane window are displayed on the screen.

There are two main classes of projection: Perspective and parallel (Figure 3.2). The above describes a perspective projection. Under such a projection, the angle

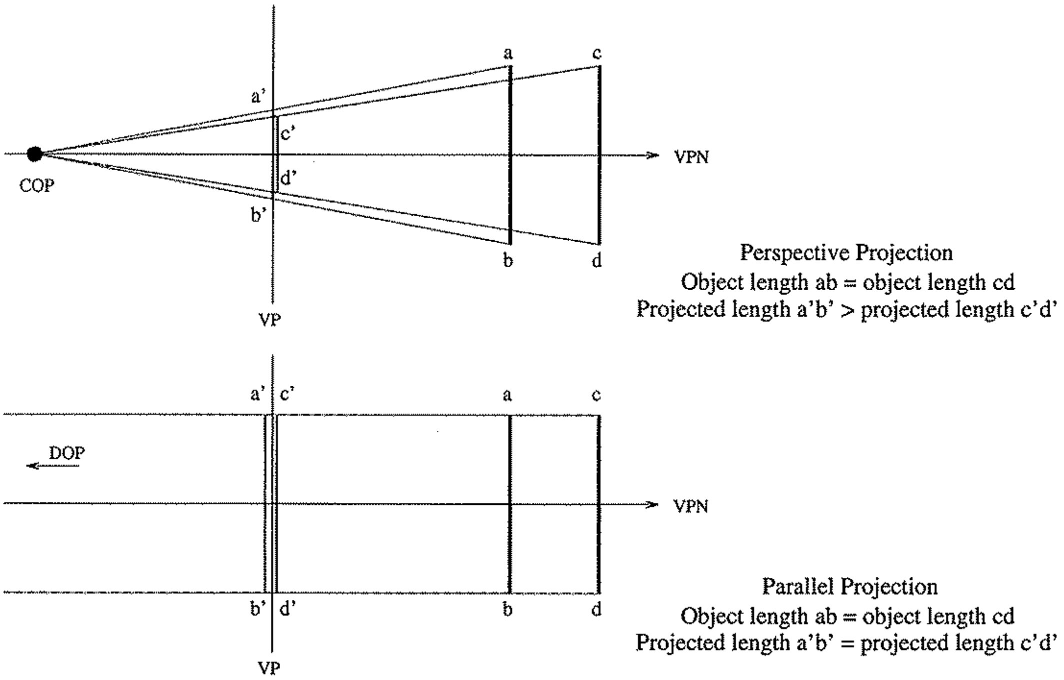


Figure 3.2: Perspective and Parallel Projection

subtended by the projection of an object on the VP will decrease with the object's increasing distance along the VPN (its z-depth); in other words, objects appear to be smaller the greater their depth, even though their size in the WC model remains the same. Perspective projections exhibit *linear perspective.* The second class of projection is the *parallel projection*, where the COP is moved back to infinity. In this class of projection, there is no linear perspective: The angle subtended by the projection of an object on the viewplane does not depend on the z-depth of the object and so its displayed size does not vary with depth. In the case of a parallel projection, it no longer makes sense to talk about a centre of projection, so the term *direction of projection* (DOP) is used instead.

Note that, confusingly, even under parallel projection this type of display is still known as a 3D perspective display in some literature, since a parallel projection is a special case of a perspective projection.

Two further parameters may be defined where the display presents an 'outside-in' perspective where exocentric judgments are required to be made relative to a reference object in the scene. The *elevation viewing angle* is the angle which the virtual camera looks down upon or up to the reference object. The *azimuth viewing angle* is the angle from which the object is viewed with respect to a reference direction.

3.2.2 Viewing

The previous section described the parameters necessary to specify the view. This section describes the geometry of perspective viewing, shown in Figure 3.3. This is

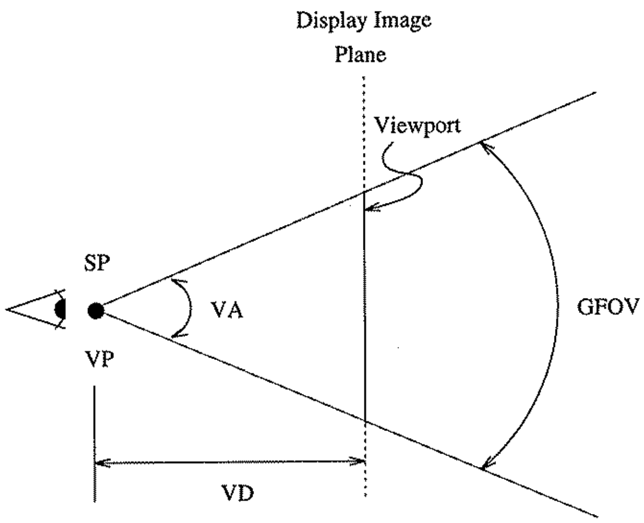


Figure 3.3: Viewing Geometry

analogous to the geometry of projection: The observer's eye, situated at the *view point* (VP), is viewing the image on the *viewport* (corresponding to the view plane window) which is a rectangular area on an infinite plane called the *display image plane* (corresponding to the view plane). The distance of the eye from the display screen is called the *viewing distance* (VD). The point in real space corresponding to the centre of projection called the *station point*[1] (SP). In Figure 3.3, the station point and the view point are coincident, but this is not necessarily the case, as will

[1]The station point is sometimes also called the 'centre of projection' in cognitive psychology. The COP in computer graphics refers to the virtual camera's position in WC space and the psychology 'centre of projection' refers to the analogous point in real space. The latter will henceforth be referred to as the station point to avoid confusion.

be described below.

Two other terms are defined related to viewing geometry. The *visual angle* (VA) of the viewport is the angle subtended by the viewport as it is observed by the viewer. The display or geometric field of view (GFOV) is the angle depicted by the display image from the station point[2].

As already stated, the station point and the view point are not necessarily coincident. If they are, then the visual angle and the field of view will be equal and objects in the world, and their depiction on the display, will be aligned. If the view point is closer to the display plane than the station point, then the display is *magnified* relative to the viewing distance, as if seen through a long focal length lens (see Figure 3.4(a): Points A' and B' are where objects A and B would be perceived to be by the observer). If the view point is further from the display plane than the station point, then the display is *minified*, as if viewed through a wide-angle lens (Figure 3.4(b)).

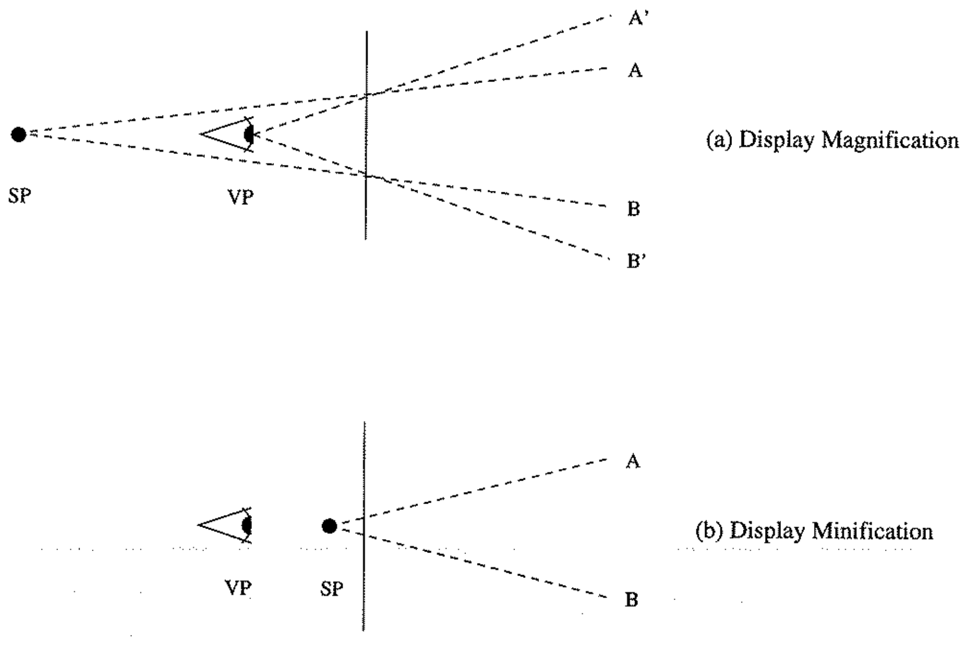


Figure 3.4: Display Magnification and Minification

[2]There is some confusion over the terminology adopted in literature: Some use the term 'field of view' to describe what is here called the visual angle; others use the same term to describe what is here called the geometric field of view.

3.2.3 Perceptual Issues

Overview

As has already been stated in Chapter 1, projecting a 3D object on a 2D image surface results in an inherent ambiguity with respect to determination of distances along the LOS into the display; dimensions lying in a plane perpendicular to the LOS are represented unambiguously whereas dimensions along the LOS are 'lost'. This may be illustrated by considering the 2D PPI as being a 3D display viewed directly from above from infinity (orthographic parallel projection); distances in the plane of the ground are represented unambiguously, whilst height information is lost altogether and so needs to be represented by some other means (in the case of an ATC display, by a digital readout). In a 3D display with an arbitrary line of sight, there will be some loss of precision in reading both horizontal and vertical distances.

The choice of viewing and projection parameters affects the appearance of the scene, and this in turn may affect how a viewer perceives it. This complicates considerably the design of such displays, especially since there are no guidelines for optimising these parameters. However, there have been empirical studies which have measured the distortions to perception of a scene when viewed from particular perspectives and these can serve as guidance.

Minification and Magnification

Generally, in a minified display, objects will be perceived closer together (or smaller) than they really are, whereas the reverse is true of magnified displays. The subjective effects of minification and magnification are well understood and exploited by photographers, who achieve the same results by varying the focal length of the lens used. Wide-angled lenses (those which give minification) are used to emphasise linear perspective as well as to give expanded coverage of a scene and depth of field. Long focal length lenses (those which give magnification) are used to compress perspective as well as to select elements in a scene and minimise depth of field.

A perceptual 'minification' has been shown to occur when viewing a 3D perspective display. Even with the viewpoint and station point coincident, objects are perceived as closer together than they really are.

Exocentric Judgments of Azimuth and Elevation Angle

Studies have shown that in perspective displays, there are errors in the exocentric judgments of the relative direction of one object from another that vary with both the true relative direction of the objects and with the perspective parameters used to generate the display [ME86, BLR90, ETGS91]. Such judgments are of the type which might be required in ATC, for example for the relative positioning of two aircraft [WTS89].

Various models have been proposed to explain and predict such judgmental errors. One is the *Viewpoint Misestimation Model*, described by Ellis, Tharp, Grunwald and Smith [ETGS91]. In this model, angles in the display are distorted because depicted azimuth angles are rotated out of and displaced from the plane of projection, the amount of rotation and displacement determining the distortion. (As a corollary, the projection will be undistorted only if viewed in plan, since the azimuth angles will then lie within the plane of projection—this is the situation with the PPI display.) It seems logical that viewers would use their perceived view direction when trying to judge azimuth, but if there is an error between the perceived viewing direction and the actual viewing direction, then there will be an error in judgment. This will occur in general when the viewer's head direction when looking at the display is different from that used when constructing the image. Ellis *et al.* further propose that another contributing factor to viewpoint misestimation, in particular causing an observed overestimation of slant angle, could be cues to the display surface.

Ellis *et al.* suggest that compensating these biases can be expected to improve perspective display design. Compensations may be by:

1. Stereo viewing or more oblique viewing to correct for the tendency to perceive the view direction too orthogonal to the ground reference (slant overestimation). (Stereo viewing helps to remove cues to the display surface.)

2. Training.

3. Symbolic enhancements, such as compass roses, at the expense of increased clutter [EH90].

Ellis and Hacisalihzade [EH90] studied exocentric azimuth judgment between two objects on a perspective display with a ground grid and a compass rose centered on the reference object. They also found obvious modulation of azimuth estimation error with target azimuth angle, estimation errors being vanishingly small at the particular azimuth angles where compass rose rays were present. Ellis and Hacisalihzade report that other studies found similar error estimation patterns from situations in which no perceptual compensation for a perspective projection was required (similar to the 2D PPI case here). These errors appeared to reflect some sort of normalisation process in which angles which were close in size to implicit standards (i.e. 0°, 90°, 180°, 270°) were erroneously judged to be closer to the standards than they in fact were.

Other Perceptual Effects

From Wickens, Todd and Seidler's summary [WTS89], the following observations have been drawn.

1. Minimum biases occur for a large GFOV matched with a large viewing angle. Biases tend to be smaller when the GFOV and viewing angle are matched.

2. Higher elevation viewing angles reduce the magnitude of azimuth errors. In the extreme case, an elevation of 90° gives a plan view and best azimuth angle judgments. Ellis and McGreevy chose an elevation angle of 30° as the optimum for their CDTI, based on pilot opinion.

3. Changing elevation angle produces a tradeoff in quality of performance between judgments of altitude and horizontal distance. Combined performance at both has been found to be better at a lower elevation angle of 15° than at a higher elevation angle of 45°.

Implications for Design

To summarise, empirical evidence suggests that there are various ambiguities and biases associated with three-dimensional perspective displays. Aside from these studies, no guidelines exist for aiding the display designer in optimising these parameters. Burnett and Barfield [BB91] suggest that further study is required regarding optimisation of viewing parameters.

The best approach for designers would therefore seem to be as follows: Select an initial set of parameters based on task requirements, 'common sense' and empirical evidence, and then trial and optimise iteratively. However, optimisation would involve extensive experimentation a lot of time, since there are numerous parameters which interact, and the designers of displays reviewed in this research took a more pragmatic approach. For example, Ellis, McGreevy and Hitchcock [EMH84] chose the elevation angle for their CDTI based on pilot opinion. This pragmatic approach was similarly adopted by the author for the displays developed in this thesis, although it is acknowledged that any operational implementation would need to consider more carefully the optimisation of viewing parameters.

3.3 Creating the 3D Effect: Depth Cues

3.3.1 Introduction

As stated in Quote 3.1 above, a 3D display is one which can represent depth along the line of sight. Depth perception may be divided into three parts: The judgment of absolute distance ('how far away is that object from my point of view'), the judgment of relative distance ('how far are those objects from each other') and the perception of the object itself as three-dimensional ('what is the true 3D shape of that object') [WTS89]. The sense of depth in an image or a scene is conveyed by one or more depth cues. For exocentric air traffic situation displays, the chief concern is the judgment of relative distance, so depth cues which convey this aspect of depth are particularly relevant.

Depth cues have been studied both singly and in combinations. The effect of adding a depth cue to a display may be either *super-additive* (the cue enhances the sense of depth) or *subtractive* (its influence diminishes the sense of depth).

Some depth cues are more salient than others—if two cues conflict, then one may dominate over another. However, which cue dominates depends on the cues employed and also on the circumstances and, to some extent, the preconceptions of the observer, which may even extend to cultural differences. As Gregory states [Gre90], no features *determine* depth or form—they only increase probabilities of seeing in particular ways. (See Chapter 10 of Gregory's fascinating and very readable *Eye and Brain* [Gre90] for a more general discussion of how we see depth.)

The studies therefore indicate a weighted additive model of depth cues: Perception of depth increases as an additive function of cues, with some cues weighted more heavily than others. Although general conclusions may be drawn as to the relative dominance of cues, this appears to be somewhat task dependent. For example, considering stereopsis: If the 3D shapes of objects are to be recognised in a dynamic environment, then stereopsis has not been found to be a highly salient cue. However, if 3D locations are to be interpreted or the display is static (or nearly so, as in the case of an air traffic situation display), then stereopsis becomes more salient.

Interactions exist between depth cues and viewing and projection parameters, and these influence the perception of distances and angles in the displays. Research into this area is incomplete, however.

3.3.2 Depth Cues

Depth cues may be represented in the objective stimulus on the retina (called *pictorial* or *world-centered* cues) and or by the state of the visual system (*observed-centered* cues). Cues may be further classified under effects of light, occlusion, object size, height in the visual field, the effects of movement, muscular sensations and binocular viewing. This section will not review depth cues exhaustively but will introduce some of the more relevant depth cues to this research.

Retinal Disparity

Due to the horizontal separation of a viewer's eyes, the retinas of each eye receive slightly different images of the visual scene. If the disparity between the images is not too great, the visual system 'assumes' that the two images fully represent the visual scene, and will fuse the images to form the perception of one visual scene with a sense of depth. However, if disparity is too great, then the visual system will no longer fuse the images, and double images may be seen.

In creating disparity artificially, a slightly offset image is presented to each eye, the degree of offset (*parallax*) being inversely proportional to the intended distance. Parallax is typically measured in minutes and seconds of visual angle.

When a viewer fixates on objects in the real world, the accommodation (focus) and convergence of the eyes are coordinated. This relationship is broken when viewing stereoscopic imagery; the viewer's eyes converge as though the images are in front of or behind the image plane but the eyes focus on the display plane. Discrepancies between accommodation and convergence may result in viewer discomfort, confusion and loss of stereopsis. Tolin [Tol87] states that these discrepancies should be minimised by using only the smallest possible amounts of parallax by placing objects at the centre of attention or objects requiring prolonged fixation at or near the display plane. He also states that the sense of depth may be greatly enhanced by the addition of monocular depth cues, particularly perspective.

In his paper on the theory, benefits and pitfalls of using retinal disparity, Tolin makes several observations about the ability of people to perceive stereo images [Tol87]. Stereoacuity amongst the general population is variable; according to estimates, 2–10% of the population fail to experience stereopsis, and perhaps a further 10% experience the effect only to a limited extent. Furthermore, stereoacuity is not constant with individuals, and can be improved with practice; experienced observers will probably perceive a stereo display differently from inexperienced ones. This complicates the evaluation of stereoscopic displays.

Object Size Cues

In perspective projections, the size of the projection of the object on the viewplane varies with its distance from the COP. In a perspective projection, object size depth cues may be useful.

Size-distance invariance. When the true size of an object is known or estimated, then the size of the visual angle subtended by its retinal image can be used to calculate the distance of the object from the observer according to the approximate relationship:

$$\text{Size} = \text{Visual Angle} \times \text{Distance} \tag{3.1}$$

This has two corollaries:

1. Objects of a greater visual angle are perceived as closer than those of a smaller angle.
2. Objects of the same visual angle are perceived as being the same distance away.

Size-distance invariance is applicable primarily to familiar objects whose true sizes are known.

Size by occlusion. Perceived objects size is supported by occlusion, with object size estimated by the number of elementary texture elements of a background surface occluded by that object. However, a relative distance cue is reliable only if the texture of the surface is uniform behind all objects.

Familiar size. A familiar object tends to maintain a constant perceived size, no matter what its objective visual angle. The perceived distance of that object then becomes a function of the visual angle it subtends. This is also known as *size constancy* or *constancy scaling.* (However, see [MVF72] for evidence that size constancy is a function of accommodation rather than of distance of the seen object.)

These cues enable the absolute distances of objects from the observer to be estimated. Further, the cues of size-distance invariance and familiar size tend to work primarily for familiar objects whose sizes are known, and aircraft representations in a 3D air traffic display are not 'familiar' objects but may be abstract representations.

Despite these limitations, these cues may still prove to be useful to the type of application under investigation here. In an exocentric air traffic display, it is the relative distances of aircraft (their separations from each other) which are important, rather than absolute distances (of each aircraft from the observer). If aircraft symbols were represented as 3D solids, a useful relative distance cue might result from the relative visual angles subtended by their projections onto the display. This was studied in the preliminary investigation described in the following chapter and is discussed in §4.2.6.

Occlusion or Interposition

The interposition of objects acts as a cue to their perceived depth. This is 'the result of perceptual organisation of the objective image by the observer' [WTS89]. When we view two objects, a near one appearing to occlude a distant one, we 'assume' that the distant object continues behind the occluding one. More familiar objects therefore increase the effectiveness of occlusion. Apparatus for upsetting this depth cue is shown in [Gre90], page 183.

Considering air traffic situation displays, it would be possible to use occlusion as a depth cue if, for example, aircraft were represented by solid shapes and were allowed to occlude each other and other symbology. This was investigated briefly in a preliminary investigation and is discussed in §4.2.6.

Height in the Visual Field

The higher an object is in the visual field, the further it appears to lie from the observer; in a typical scene, it is assumed that the foreground is low and the horizon is high (when looking down on objects; if objects are viewed from below, this cue can be misleading) [WTS89].

Texture Gradients

The texture of a surface affects the perception of its depth. Texture 'is defined by the size and spacing of elementary features of which it is composed' and gradient as 'the change of texture perceived as one looks from one's feet up to the horizon' [WTS89]. When texture gradients are used as a depth cue, the viewer assumes that elementary features are roughly the same size and are spaced approximately equally across the surface.

If texture elements are all of objectively equal size and objectively consistent density, then the change in width of an elementary feature gives a sense of *linear perspective*. Linear perspective is only exhibited where perspective projection is used to construct the scene, and is absent in parallel projections, as discussed above.

Light Effects

The effects of light can convey a sense of depth in several ways:

Luminance/brightness effects. Objects, parts of objects or regions may be perceived to be at different depths simply as a result of the differences in luminance. This phenomenon has been employed to give a '3D' effect to 'buttons' and other graphical components in graphical user interfaces, for example. However, the direction of this perception has been shown to differ between individuals; no consistent trend has been found for darker shades to signal either closer or more distant regions. In designing a '3D button' graphical object, the author has himself experienced this problem; some subjects thought a button 'in' were some perceived the button in the same state as 'out'.

Proximity Luminance Covariance. Variation in brightness can induce a sense of depth; in viewing a computer-generated, luminous wire-frame object, brighter parts appear to be closer to the observer than dimmer parts. This is also known in computer graphics as *depth cueing* ([FVFH90], p. 690).

Aerial Perspective. In real life, atmospheric effects may result in the objective colour of a distant being desaturated and taking on a bluish hue. Desaturating and/or adding the environment's ambient cue to an object's colour can therefore affect the perceived depth of an object.

Shadows and Highlights. An object's perceived depth and surface shape may be affected by the presence of shadows and highlights.

Colour. Differences in colour may effect perceived depth. For example, two juxtaposed regions of different colours may be perceived at different depths.

3.3.3 Nuisance Cues and Cue Conflicts

Depth cues may introduce undesirable 'nuisance' effects, and conflict with other depth cues. For example, cues to the display surface (screen) may indicate to the viewer that the image is flat rather than three-dimensional. An object displayed stereoscopically with negative disparity may appear in front of the screen, but if it touches the edge of the screen it may also appear to be cut off by the surround, suggesting that it is behind the display plane [Tol87]. A more compelling sense of depth may be created by removing such nuisance cues.

3.3.4 Cue Combinations

The number and type of depth cues employed is a trade-off between implementation cost and effectiveness. For example, research suggests that retinal disparity is a particularly salient cue for relatively static scenes; however stereopsis is expensive to implement compared with monocular cues, and any performance advantage which it gives must be weighed against these costs.

Theoretically, for the ATC application, where 3D location is important, stereopsis might be a highly salient depth cue. Regarding other cues, however, there is less certainty as to which cue combinations are the most effective. The theories from cognitive psychology may be consulted in the display design process, with the caveat that the research is incomplete and its conclusions may be highly task-dependent. To illustrate this, the following discussion presents some studies where stereopsis has been added to aviation displays.

Trials by Zenyuh, Reising, Walachi and Biers [ZRWB88] and Mazur and Reising [MR90] studied the effect of adding depth cues to the same format of tactical air situation display (display of positions of aircraft relative to an 'ownship'). Zenyuh *et al.* compared the addition of stereopsis to the cue of familiar object size and found improved performance (measured in terms of response speed and accuracy in a search task) where stereopsis was present, but there were no significant advantages where either was the sole cue used; i.e. the study did not imply that either cue alone was more effective. Mazur and Reising found that for the particular cue combinations they studied (all possible one, two and three cue combinations out of stereo-3D, aerial perspective and familiar object size) a law of diminishing returns appears to set in with regard to the number of cues. In a task measuring performance in determining aircraft location (response time and accuracy), there was a marked improvement in performance in going from one to two depth cues, but three cues did not seem to give a significant further advantage; further, they found that the type of depth cue did not appear to have a differential effect, that is, it made a difference if one or two depth cues were displayed, but the particular cues used had no significant influence.

These trials showed that stereopsis did have an effect, but that it was no more effective than any other depth cue tried. This would seem to contradict some of the cognitive psychology (as opposed to application-orientated) research: According to Wickens, Todd and Seidler's review [WTS89], stereopsis is more salient than aerial perspective, whereas Mazur and Reising found stereopsis and aerial perspective to be equally weighted.

Studies by Way [Way88, Way89] adding stereopsis to pictorial cockpit displays found that response time and error frequency were both reduced where disparity augmented real-world depth, but made no difference when used to make an element of an otherwise flat display more noticeable.

Steiner and Dotson [SD90] added stereopsis to a plan-view tactical air traffic display with solid aircraft symbols. They found that response times were faster and errors were fewer in the *2D* case for all traffic densities tried than for the plan view plus stereopsis. They proposed three explanations for this finding. Firstly, the ability to see stereoscopic depth on a CRT must be learned. Secondly, with a 2D display all of the information can be in focus at one time, but with a 3D display the convergence of the eyes must be varied when viewing different parts of the display. Finally, they postulated a certain fascination factor with stereoscopic displays which may cause subjects who are unfamiliar with them to gaze at the display.

In summary, these studies seem to suggest that:

1. Stereopsis enhances depth in conjunction with certain monocular cues, but may be no more effective than any other depth cue when used by itself.

2. There may be a number of depth cues beyond which depth perception does not significantly improve.

However, these must be considered with the caveat that they may only be applicable to the cue combinations and tasks in the studies from which they were derived.

As ever, cognitive theory may be applied in designs, but the final design must always be validated empirically.

3.4 Chapter Summary

This chapter has first defined what is meant by a 3D display in this thesis, and introduced a taxonomy of displays based on the relationship between the viewer and the displayed scene. Next, the way in which a 3D image is generated and viewed was introduced, and it was observed that viewing and projection may bias the way in which a viewer interprets the image. Different depth cues (cues which signal to the visual system of the viewer that an image is three-dimensional rather than flat) were then introduced.

In the next chapter, some of this theory is applied to the development of a 3D display for air traffic control.

Chapter 4
Displays for Air Traffic Control

4.1 Introduction

4.1.1 Overview

The preceding chapters have introduced air traffic control and three-dimensional displays and have cited possible benefits which a 3D display might bring to ATC. This chapter first examines some of the issues in implementing a 3D display for ATC. A preliminary display design, used to demonstrate the feasibility of such displays and as an initial development step, and its implementation on a workstation are then described and discussed. This preliminary display was a precursor for the display formats used in the pilot study, the development of which is then described.

The technology used to implement immersive displays is very different from that used for displays implemented on workstations, and has different limitations and restrictions. The development of the immersive 3D displays for demonstration in the pilot study is thus described separately at the end of the chapter.

4.1.2 Three-Dimensional Display Design Considerations

It has already been observed that 3D displays are not a panacea for the display of spatial relationships or other multidimensional data, and bad design or inappropriate application may negate any benefits over a well-designed 2D equivalent [Tol87, Ell89]. The display design must be suited to the task and must consider the capabilities and limitations of the human cognitive system [YA85].

Principles used in the design of 2D displays are generally well established since the technology is mature and well understood, and some of these principles may be applied to 3D displays. However, as has been shown in the preceding chapter, 3D displays carry additional design issues. Further, current theories of 3D perception are still not sufficiently complete to give comprehensive, concrete guidelines relating to the design of 3D displays. Some researchers (e.g. Ellis [Ell89] and Wickens and Todd [WT90]) have written observations on 3D display design based on their experimental work, and Wickens, Todd and Seidler [WTS89] have compiled a survey of research results to aid designers (and this has been referred to extensively in this research). However, these are insufficient to form a display designer's 'bible', and so it is up to the designer to apply the relevant information from the research literature as best he or she can, and then to validate the design with experiments. This is the methodology adopted in this research.

Some of the design considerations pertaining to 3D displays include:

1. Choice of depth cues. The choices are partly dependent on the chosen/available display and image generation technologies as well as on perceptual issues.

2. Choice of viewing parameters. These include geometric field of view, elevation, azimuth and vertical exaggeration. The choice of viewing parameters is important, since they have been shown directly to impart biases influencing the interpretation of spatial parameters [EMH84, ETGS91]. As Wickens *et al.* state [WTS89]:

   4.1 ... *these various factors considerably complicate the design of perspective displays, and suggest the need for principles which optimize the setting of all parameters.*

3. Symbolic enhancements, deliberately introduced distortions or training to counteract the perceptual problems associated with 3D displays [BLR90].

4. Clutter.

5. Control of viewing parameters. Is it necessary for the operator to be able to control the viewpoint and other viewing parameters? If such manipulations are permitted, what are the effects?

6. Representation of elements in the display (appearance, size, colour etc.).

7. Implementation issues. Choice of image generation and display hardware.

8. Restrictions due to working environment, human factors of operation (e.g. fatigue, screening operators for stereo acuity and colour vision), acquisition and operating costs, required changes in training and operational practice, etc.

These considerations are not isolated; there is a degree of interaction between them and also a strong dependence on the task for which the display is being designed. As with all engineering, display design is necessarily a compromise between the different factors.

4.2 Preliminary Design Investigation

4.2.1 Introduction

Relevant research material relating to ATC and 3D displays having been reviewed, a preliminary investigation was made of 3D displays for ATC resulting in the development of a basic demonstration display of a rudimentary air traffic simulation. The purpose of this was:

1. To serve as an initial step in the development of a suitable display format for a 3D air traffic display for controllers, and to gain feedback from individuals to whom it was demonstrated,

2. to develop a basic air traffic simulation,

3. to demonstrate the feasibility of the concept of a 3D visualisation of air traffic for ATC, and

4. to serve as a demonstrator to generate interest in and possibly support for the research.

The final demonstrator was presented to interested parties at CAA NATS (National Air Traffic Services).

Three display formats were developed: A 2D PPI, a pseudo-3D TTW display and an immersive 3D display. The 2D and pseudo-3D displays were implemented on a Sun *SPARCStation II* workstation, using 2D and 3D graphics libraries written by the author earlier in the research programme. In the pseudo-3D display, it was possible to specify an arbitrary viewpoint and viewing angle (i.e. the user had control of the location of the COP and the direction of the VPN). The displays showed aircraft movements over a 60×60 n.mi. area; larger than representative for approach control but smaller than for some en-route sectors. A list of requirements for information content if the displays was first compiled. These included:

1. Aircraft spatial position information:

   (a) Ground position (i.e. the position of the aircraft over the ground, neglecting its altitude).

   (b) Altitude. This must be readable to the precision provided by a mode-C readout (i.e. 100 ft accuracy).

2. Callsign/mode-A code.

3. Approximate speed information.

4. Track/heading information.

The display formats were then designed to meet these requirements.

The air traffic was driven by a basic simulation. This supported an arbitrary number of aircraft moving on preprogrammed flight paths, a flight path being a route through 3D space specified by a number of *waypoints*, each waypoint comprising a ground fix and an altitude. The simulation computed the positions of the aircraft every 0.5s, but the actual displayed positions of the aircraft were updated only every 4s to simulate the sweep time of a radar.

The following sections discuss the 2D and pseudo-3D displays developed on the Sun workstations. First, initial prototype formats and the simulation are described. A discussion of the prototype design then follows, followed by discussion of the design of the displays used in the experiments in the light of this preliminary investigation. The immersive display design and development are discussed separately in §4.4.

4.2.2 2D PPI Display

A diagram showing the format of the prototype 2D PPI is shown in Figure 4.1. Range rings were shown in grey at 10 n.mi. intervals over a black background. User-defined navigation points were represented as blue crosses (as a sort of rudimentary video map).

Aircraft were represented by a filled square centered over the radar position and a heading line, both in green. Associated with each aircraft was an alphanumeric datablock showing callsign and altitude (in hundreds of feet), attached to its associated aircraft symbol by a *leader line*. Figure 4.1 shows two aircraft: BM298 heading north at 8 000 ft, and BA045 heading east at 5 000 ft.

In order to avoid the datablocks overlapping, a simple datablock overlap avoidance algorithm was implemented. Each datablock could be displayed in one of eight

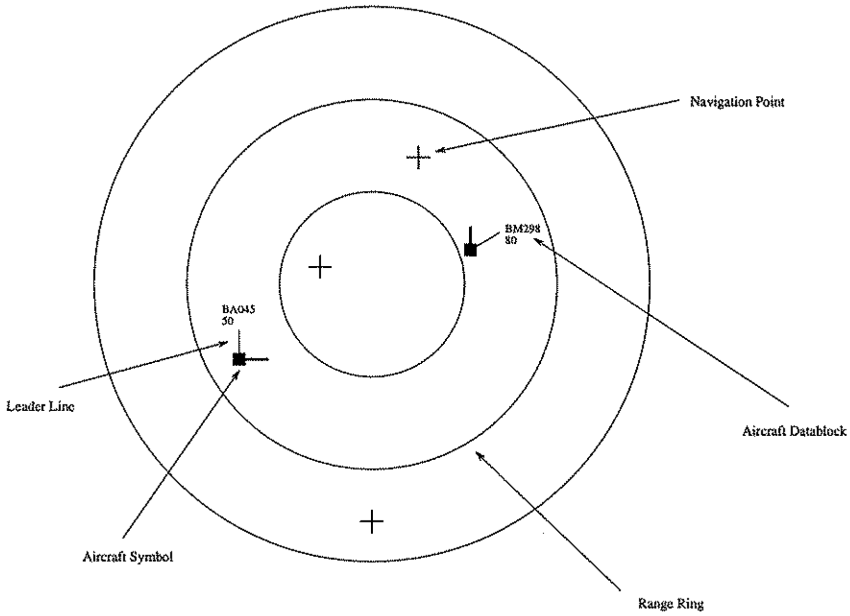


Figure 4.1: Prototype 2D PPI Display

positions around the aircraft symbol (N, NE, E, etc.). When each datablock was displayed, its bounding box (smallest bounding rectangle) was computed and tested in each of the eight positions in turn (starting from N and working clockwise) until the first position where it did not overlap another datablock's bounding box was found, where it was displayed. (If no non-overlapping position was found, the datablock was displayed in the N position.) This does not avoid datablocks overlapping the aircraft position symbols themselves, but the basic algorithm could be extended to include the aircraft symbol with the datablock in the bounding box. If this is done, a significant part of the bounding box would then be empty, so a more complex bounding polygon would have to be defined to increase display area utilisation. However, the simulations had only a small number of aircraft, so this was not a problem.

4.2.3 3D TTW Display

The 3D display format was based on Ellis, McGreevy and Hitchcock's Cockpit Display of Traffic Information (CDTI) [EMH84]. The layout of the 3D TTW display is shown in Figure 4.2. The 3D view window shows the view from a hypothetical viewpoint, the position and orientation of which may be adjusted by the user. To

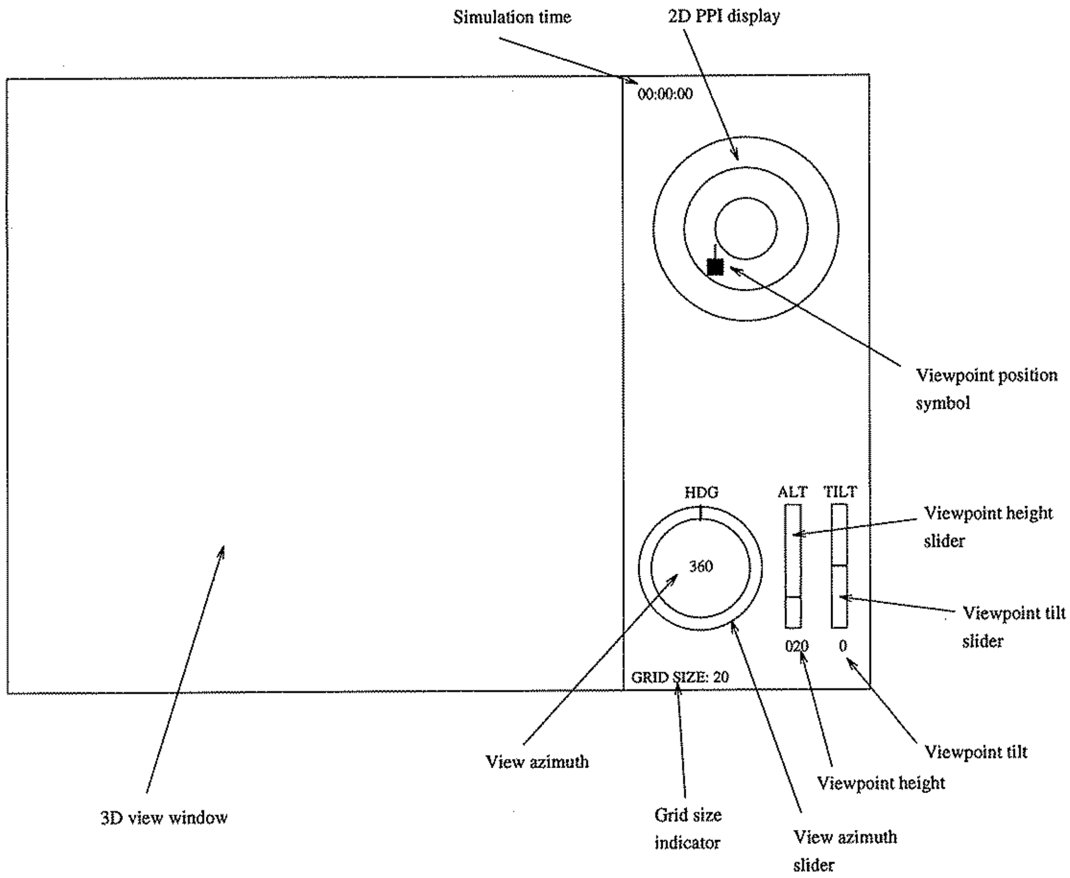


Figure 4.2: Prototype 3D TTW Display Layout

the right of this are various sub-displays and controls. The 2D PPI sub-display shows the positions of all aircraft in plan, without datablocks, the plan position of the viewpoint and its current view azimuth (represented by a red filled square and a red line projecting from it in the appropriate direction). The user could move the plan position of the camera by moving a cursor (controlled by a three-button mouse) over the desired spot and pressing the left mouse button. As long as the button was held, the viewpoint followed the movement of the cursor within the PPI, and the view in the 3D view window changed accordingly. Releasing the button released the viewpoint from the cursor at its current position. The view azimuth was controlled by a circular 'slider'; the user could 'grab' the slider bar with the mouse and move it to the desired azimuth as desired, again with the scene in the 3D view window being updated as a result. A digital readout in the centre of the slider gave the view azimuth in degrees. The viewpoint's elevation (tilt) and height were similarly controlled by two linear sliders, each with a digital readout below it. Elevation was restricted to $\pm 90°$, with $0°$ being horizontal. No viewpoint 'roll control' was provided, since it was thought that this would be of limited use and could contribute to disorientation if it were permitted.

Still earlier versions of the prototype lacked the 2D sub-display and used a 5 degree-of-freedom input device (the Desktop Bat [SD91]) for viewpoint orientation and position control. However, using this device, the author experienced great difficulty in orientating himself within the environment, and so the 2D display and control sliders were introduced to provide a simple, albeit non-integrated, control over viewpoint position and orientation.

Aircraft were represented as non-regular flat-shaded tetrahedra, stretched along one axis with the long axis aligned with the aircraft's heading. The symbol was defined in world coordinate space and a perspective projection was used with a wide geometric field of view (approx. $90°$), so the displayed size of the aircraft symbol varied with its depth, the aim being to provide object size depth cues.

The ground plane was represented by a square grid instead of range rings, following the example of Ellis *et al.*'s CDTI (§2.6.3). This provided a depth cue in the form of a texture gradient. (Moreover, this eased implementation since the author's graphics package did not support 3D circle drawing, although this could have been implemented.) The grid interval was set to 10 n.mi. As in the 2D PPI, it was displayed as grey line segments on the black background, with navigation points being displayed as blue crosses.

As with the CDTI, a vertical *drop line* was projected downward from each aircraft to the ground plane to represent the aircraft's height and to show its position over the ground. This was omitted on a trial basis, but the author found that without it there were insufficient depth cues to be able to determine the positions of the aircraft in space. It would also be possible to project lines horizontally onto two orthogonal vertical planes (NS and EW); however, this would probably lead to excessive clutter and be of little additional benefit, apart from showing relative heights more clearly.

The 3D view display was drawn in the following order: Ground plane grid, navigation points, datablocks and leader lines, drop lines, and finally aircraft symbols. Aircraft symbols were flat-shaded (Lambertian shading) and Z-buffered (in software). This meant that they obscured each other and also other symbology. The aim was to see whether occlusion could be used as a depth cue.

4.2.4 Simulation

The simulation of an air traffic scenario was specified in a human-readable ASCII script file which was read at the start of the simulation. This specified each aircraft's callsign, its initial heading, position, altitude and speed, and an optional flight path defined by two or more waypoints (given as a ground position and altitude). A new speed could also be specified at each waypoint if desired. (Without the flight path, the aircraft just proceeded according to its initial conditions.)

The behaviour of each aircraft was determined by a simple mathematical model. This was constructed in a layered manner similar to automatic flight control systems, although greatly simplified. A simple model was used to simulate the aircraft's flight dynamic characteristics. Around this was 'wrapped' a simple autopilot capable of capturing and maintaining the values of certain flight parameters (altitude, heading, speed). This in turn was driven by demands from the flight path specified in the scenario script file.

Aircraft were represented as point masses and their basic behaviour was modelled by simple dynamics equations:

$$\begin{array}{ll} \text{Position} & \begin{cases} ds = u\ dt + \frac{1}{2}\dot{u}dt^2 \\ dx = ds \sin\phi \\ dy = ds \cos\phi \end{cases} \\ \text{Altitude} & dz = -w\ dt \\ \text{Speed} & u = u + \dot{u}dt \\ \text{Heading} & d\phi = \arctan\frac{v}{u} \end{array}$$

where:

(u, v, w) are velocities along the aircraft's body-fixed **X** (longitudinal, roll), **Y** (lateral, pitch, positive right) and **Z** (vertical, yaw, positive down) axes respectively

$\dot{u}$ is the acceleration along the aircraft's **X** axis

(x, y, z) are the aircraft's position in world coordinates (x is positive east, y is positive north and z is altitude, positive up)

ϕ is the aircraft's heading (the angle between the aircraft's **X** axis and north (the world y-axis)

The aircraft being modelled as a point, its yaw, pitch and roll angles were set to zero: Therefore, the aircraft **Z** axis was always parallel with the world z axis (but in the opposite direction).

The 'autopilot' simulation worked by simple feedback, responding to demand altitude z_d, demand speed u_d and demand heading ϕ_d and controlling w, $\dot{u}$ and v respectively:

$$\begin{aligned} z_{err} &= -(z_d - z); & w &= Az_{err} & &: (-W \leq w \leq W) \\ u_{err} &= u_d - u; & \dot{u} &= Bu_{err} & &: (-\dot{U} \leq \dot{u} \leq \dot{U}) \\ \phi_{err} &= \phi_d - \phi; & v &= C\phi_{err} & &: (-V \leq v \leq V) \end{aligned}$$

where z_{err}, u_{err} and ϕ_{err} are error terms, A, B and C are fixed 'gains' and W, $\dot{U}$ and V are maxima/minima (W and $\dot{U}$ being fixed, and V (maximum turn rate) being given by the expression $V = u\tan(M\ dt)$ where M is the maximum turn rate (3° per second for a 'standard' turn). The gains and maximum rates were all fixed, so the aircraft all had the same performance. Varying these terms would allow aircraft of different performances to be modelled.

Demand terms were from the script file. Each waypoint might specify a new demand heading or altitude, for example, and then the aircraft would turn, climb or descend as the autopilot 'captured' the new parameter, which would then be maintained until the next waypoint.

The simulation was designed to be only loosely coupled to the display parts of the program, so that the same 'core' could be used for the different display formats. This approach was adopted throughout this research, and considerably eased development.

4.2.5 Implementation

The simulation and displays were written in C++ [Str91a] and implemented on a Sun *SPARCStation II* workstation running the SunOS version of the UNIX operating system. 2D and 3D graphics support were provided by a machine-independent graphics library developed by the author and based on a specification developed for another project (the QMW-developed Graphics Interface Layer of the SPIRIT European workstation project [BHRR92]). The X11 windows system provided the interface between the library and the hardware. The original motivation behind this was to allow portability.

The simulation was constructed as an event-driven object-orientated message passing system. Elements in the simulation (aircraft, the environment, navigation points etc.) were all represented by C++ objects. Each simulation object type implemented some common interface functions (in particular, the functions `Render` and `Behave` were supported by all objects). Non-simulation objects included graphical user interface (GUI) elements (e.g. the slider objects). The GUI elements were implemented by the simulation application instead of relying on toolkits (such as Motif or Xtk toolkits under the X11 windowing system) to achieve greater portability.

Objects could send messages to each other in response to events. The simulation was driven by a timer which periodically triggered events, such as simulation object behaviour (e.g. aircraft model computation), aircraft displayed position update, time display update etc.

Animation was achieved using double buffering. The normal approach to double buffering is to have two sections of video memory (frame buffers) which can be displayed in turn. While one frame buffer is being displayed (the 'front' buffer), the next image can be constructed in the buffer which is not being displayed (the 'back' buffer) and when the time comes to show it, the video hardware simply switches buffers, 'swapping' front and back buffers. This is efficient, since it requires no copying to display the next frame (the video hardware simply switches the source of the screen image from one buffer to another), but requires a lot of display memory, and may compromise spatial or colour resolution.

Unfortunately, double buffering with frame buffers could not be done because not all *SPARCStation II*s support multiple frame buffers. Therefore, the offscreen image was drawn in a pixmap in main memory and then copied to the frame buffer by a pixel block transfer operation when the screen was updated. However, this limits the size of offscreen pixmap and the maximum frame rate. If the areas which are to be pixel block transferred each frame are too large, then the machines tend to run out of memory and as displayed area increases, so does the rendering time and the memory transfer time. If the displayed areas are too small, then legibility suffers.

In light of this, it was felt that it would be very difficult to support a field-sequential time multiplexed stereoscopic display on the *SPARCStation II* since this would require quad buffering (front and back buffers for each eye). It was therefore decided to drop the author's graphics library from the research and to use the then newly introduced Silicon Graphics workstations, which supported hardware double buffering and stereo buffering, had much higher graphics performance, and had their own powerful 3D graphics library (GL), a version of which (OpenGL) exists on several other platforms (including Sun workstations) giving portability.

4.2.6 Discussion

This section introduces some of the issues raised by implementing and demonstrating the prototype displays described above. It was beyond the scope of this research to address all the issues, and some are left as questions for further investigation. However, the results of the preliminary investigation served as a useful base for further development and allowed some choices to be made.

Aircraft Symbol

This preliminary implementation explored briefly the possibility of using 3D solids for aircraft symbols (implemented as polyhedra) to see whether object size depth cues could be exploited. As discussed in §3.3, the cues of size-distance invariance and familiar size tend to allow *absolute* distances from the viewer to be estimated, and apply mainly to *familiar* objects. However in an ATC display, although the symbols are not familiar in everyday experience, it is postulated that differences in *relative* distance of two aircraft from each other might be judged from differences in visual angle subtended by their projections onto the display: If two objects of the same objective size are displayed, one being twice the distance of the other from centre of projection, then the further object will subtend half the visual angle of the nearer one.

The possibility of using a 3D model to represent the aircraft instead of a simple fixed-size 2D shape (e.g. square or cross) introduces the further possibilities of making the shape in some way representative of the aircraft type, and also of orientating the symbol with the aircraft it represents so that pitch and bank angles are shown as well as heading (if such information were available).

Aircraft come in different shapes and sizes. The ATCO needs to know the broad characteristics of the aircraft with which he or she is dealing to account for their relative performance. It might therefore initially seem a good idea if the aircraft symbols were representative of their types. However, this level of complexity may not be necessary since aircraft can be 'lumped' into similar classes in terms of performance and wake turbulence category. A better alternative may be to restrict the symbols to represent these classes. Though a representative aircraft symbol might appear attractive, one must weigh carefully the costs against the benefits. Current 2D displays use only the one symbol for aircraft and the ATCO must refer to the flight strip and memory to identify the aircraft type; however, this does not appear to be a limitation at present. Incident reports would have to be studied to determine whether or not misclassification of aircraft by ATC has been a significant factor in incidents. The costs of implementing a 3D symbol are significantly greater complexity in modelling, higher graphics performance required and possibly more clutter.

If the 3D models used to represent the aircraft were of different sizes, this might interfere with judgments of relative distance. With a large object displayed in the foreground and a small object displayed in the background, the differences in the visual angle subtended by the two objects on the viewplane would not only be due to their different depths but also to their different true sizes. If the observer fails properly to take this into account and interprets part or all of the difference in visual angle due to size alone as due to distance, this might cause him or her to think that the relative distance is greater than it actually is.

A further complication is that the cross-sectional area of an aircraft is much less than the area of its side, so two aircraft of the same size and equidistant from the observer, but with one end-on and the other side-on, would appear to be of different sizes and this might also interfere with distance judgment, especially if the orientations of the objects could not be determined easily.

In this preliminary investigation, all aircraft symbols were made uniform. The simplest closed polyhedron is the tetrahedron, so a 'stretched' tetrahedron was chosen to represent the aircraft; the long axis can be orientated with the aircraft's heading to give an indication of direction of flight.

Choosing the size of the polyhedron was not straightforward. When the symbols were scaled to the size of an 'average' aircraft, then with the viewing parameters used, the aircraft symbols were for the most part no larger than small points and were too small to discern their orientation (and certainly too small to make out any features if they were more complex representative solid models). However, when the viewpoint was placed close to an aircraft symbol, the symbol obscured a significant portion of the display.

One solution would be to use non-linear scaling: The size of aircraft symbol varies with depth over a restricted range, having maximum and minimum limits.

Datablocks and Object Occlusion

The alphanumeric characters making up the datablocks should be a constant size so as not to be illegible or confusing. Leader lines should be unambiguously attached to the corresponding aircraft symbol. However, problems were found in this display where aircraft symbol size was allowed to vary and where object occlusion occurs.

Should aircraft symbols be allowed to obscure datablocks and leader lines? If so, there could be a problem reading partially obscured datablocks, and possible problems in associating aircraft with datablocks, since if more distant aircraft's symbol were obscured but not its datablock, the distant datablock might appear to be attached to the near aircraft. If aircraft symbols are allowed to obscure each other but not datablocks, then there is a potentially confusing discrepancy between the occlusion of aircraft and non-occlusion of their associated datablocks and leader lines.

Fixed versus Movable View

This could be another important point for study: Does varying the view of a complex 3D scene reinforce the perceived spatial relationships, or does it confuse the observer by changing the displayed relationships and forcing him or her to update his mental representation?

From discussions with air traffic controllers, it appears that ATCOs do not generally want to spend their time fiddling with controls, since it takes them away from important tasks in hand. However, where something is difficult to interpret from one angle (too many aircraft near the same line of sight causing clutter) a change in viewpoint could clarify the situation.

Representing Aircraft of Unknown Height

The 3D display makes the assumption that the height of aircraft is known (by mode-C transponder). However, not all aircraft are equipped with mode-C transponders (general aviation aircraft are usually not, since they tend to be flown by private pilots under visual flight rules and so tend not to fly in airways), begging the question as to how they are to be represented.

These could be represented not as discrete points but as 'columns', but such a representation may give excess clutter and over-emphasise them with respect to other traffic. An alternative would be to allow the ATCO to enter a height determined by voice contact with the aircraft, but this would not necessarily reflect the true state of the aircraft if the pilot changed his altitude without informing ATC, and presenting false information of this sort is potentially dangerous.

4.3 Pilot Display

4.3.1 Introduction

The preliminary investigation having shown the feasibility of the 3D ATC concept and having raised some 3D display issues, further work was carried out to develop the display formats to be used in the pilot study described in the next chapter. This section describes the development of the display formats for the non-immersive 3D displays, and some of the issues which were considered. Different considerations applied to the immersive display, and these are discussed in §4.4.

4.3.2 Lessons from the Preliminary Display

The following design points were adopted in the light of the findings from using the preliminary display:

1. The geometric field of view (90°) was, in the opinions of people to whom the display was demonstrated, too wide.

2. It was evident that using shaded solid models to represent aircraft, and so allowing object size and occlusion depth cues, raised a lot of questions which would have to be explored in greater detail, and there are problems with their implementation. It was decided that the aircraft symbol should not be scaled with depth and should be of the same appearance irrespective of the type or orientation of the aircraft; the easiest way of doing this is not to use a solid model to represent aircraft but simply to draw a 2D geometric symbol directly onto the viewport at the position which the aircraft would have been projected to. This approach was adopted for the non-immersive displays.

3. Datablock text should be displayed at a fixed size and where possible, datablocks should not overlap. If the aircraft symbols are sufficiently small, then temporary obscuration of part of a datablock should not be a problem.

4. The question of fixed versus movable views (i.e. allowing manipulation of viewing parameters such as COP position and VPN direction) was not investigated further. Changing the view (either through direct control of the viewer or automatically—for example, sweeping the view azimuth angle periodically

through a small cone) might enhance the mental representation of spatial relationships, and would certainly help where spatial relationships are difficult to determine from a certain viewpoint but this might also lead to disorientation. Anecdotally, conversations with a number of ATCOs indicated that they wanted to spend as little time manipulating controls as possible due to diversion of attention from the main task. From the point of view of an experimental investigation, allowing the subject to change the view would add another uncontrolled variable. Since it is desired to minimise these, only a fixed viewpoint was adopted for later displays. For certain tasks, a changeable viewpoint should not be ruled out, particularly in non-time critical or workload critical rôles such as perhaps overall flow monitoring displays, or non-operational rôles such as post-mortem analyses of training exercises or real incidents.

In an immersive virtual environment, changes in viewing azimuth and elevation angle can be achieved by an action as trivial as looking in another direction, although large changes in viewpoint position may require more substantial actions.

5. The use of a square grid for the ground plane (in the 3D display) was rejected in favour of range rings; this is more consistent with the current 2D PPI displays.

4.3.3 Display Colour and Symbology

Following demonstration of the preliminary display to CAA NATS, an interim NATS standard for the use of colour in ATC displays was obtained [RM92] and was used as a basis for symbology and colour design. This gave recommendations based on a comprehensive literature survey into the use of colour in ATC (for example, [NH85]). The interim standard proposed a 'visual layer model', dividing information into a number of conceptual visual colour/luminance layers where static, background information appears further away and less conspicuous than active data, which is in turn further away than alerting data. Further subdivision of these layers led to a seven layer scheme, where 'higher' layers must stand out against 'lower' layers. Colour palette values were suggested for each layer and adopted. The standard also included recommendations for symbology design, such as the design of datablocks and aircraft position symbols.

The colours and symbols adopted for the pilot display are discussed in §4.3.5 below.

4.3.4 Depth Cues

A 3D display allows for a direct representation of depth. According to the additive theory of depth cue combination (§3.3), in general the more depth cues provided, the greater the sense of depth, so initially it might seem a good idea to employ as many depth cues as practically possible. However, incorporating a depth cue often imposes a computation/rendering time overhead and since depth cues are not all equally important, there may be some which give only minor increases in depth compared to others. As has been demonstrated in the research reviewed in §3.3.4, a law of diminishing returns may set in; a point may be reached beyond which adding further depth cues does not have any significant influence on the performance of a task relying on depth perception. There is also evidence that the salience of a given depth cue depends on the circumstances in which it is used. Due to its theoretically high salience, stereopsis was therefore included as an experimental variable in this investigation. It was intended to use stereo displays in the pilot study, but this was not possible due to problems with implementation.

Object size depth cues were rejected by the preliminary study, for reasons discussed in §4.2.6 above.

The adoption of the CAA NATS interim colour standard described above ruled out the possibilities of introducing the effects of proximity luminance covariance ('depth cueing'), aerial perspective and other colour effects, since this might have interfered with the standard's visual layer model. The standard was designed for a two-dimensional PPI display and conceptually, the visual layers add a three-dimensional effect to the display, allowing higher priority information to stand out against lower priority information. Whether or not such a model interacts with judgment of spatial depth was not investigated in this thesis.

It was anticipated that linear perspective would be a useful depth cue; however, discussions with air traffic controllers over 3D projections revealed concerns with the ability to judge distances and azimuth angles accurately in a 3D display. The ability to judge horizontal distances and azimuth angles with sufficient accuracy was cited by controllers as being very important, since separation minima must not be violated, and radar vectoring instructions might be required to separate traffic. Strutt [Str91b] observes that for tasks which require angles to be determined with precision, for example issuing vectors to line an aircraft up with a runway extended centreline for approach, the accuracy required may not be within the range of error associated with reading a 3D display. In a perspective projection, the visual angle subtended by a given objective distance varies depending on where that 'distance' is located in WC space. With a parallel projection, this problem is not so pronounced. This raised the question as to which type of projection would be preferable for an ATC display. Can viewers compensate for linear perspective (or lack thereof) or will they mis-read distances?

There are other implications in deciding whether to adopt a perspective or parallel projection: Due to linear perspective, objects at greater depth tend to be displayed closer together, possibly increasing clutter near to the 'back' of the display. However, the presence of linear perspective may significantly increase the perception of depth in a display.

It was decided to carry out an experiment during the pilot study to help resolve this issue (described in §5.2.1). For the pilot study, therefore, two 3D display formats were developed, one using a parallel projection and the other a perspective projection.

4.3.5 Display Format

Figure 4.3 shows the design for the 2D pilot experiment display. The 3D displays are shown in Figures 4.4 (perspective projection) and 4.5 (parallel projection).

All displays depicted an area of 100×100 n.mi. centered on the radar head at London Heathrow airport. Video map data of major airspace structures and coastlines around the London terminal area were obtained from CAA NATS. Different coloured infills were used for the background of the video map and the terminal airspaces associated with Gatwick, City and Heathrow airports. Lines delineating the airspaces and coastlines were drawn in grey. As an aid to assisting judgment of distance, range rings were superimposed on the map at 10 n.mi. intervals.

In the computer model of the airspace, the world coordinate system had the origin at the Heathrow radar head, with the world x-axis being East, the world

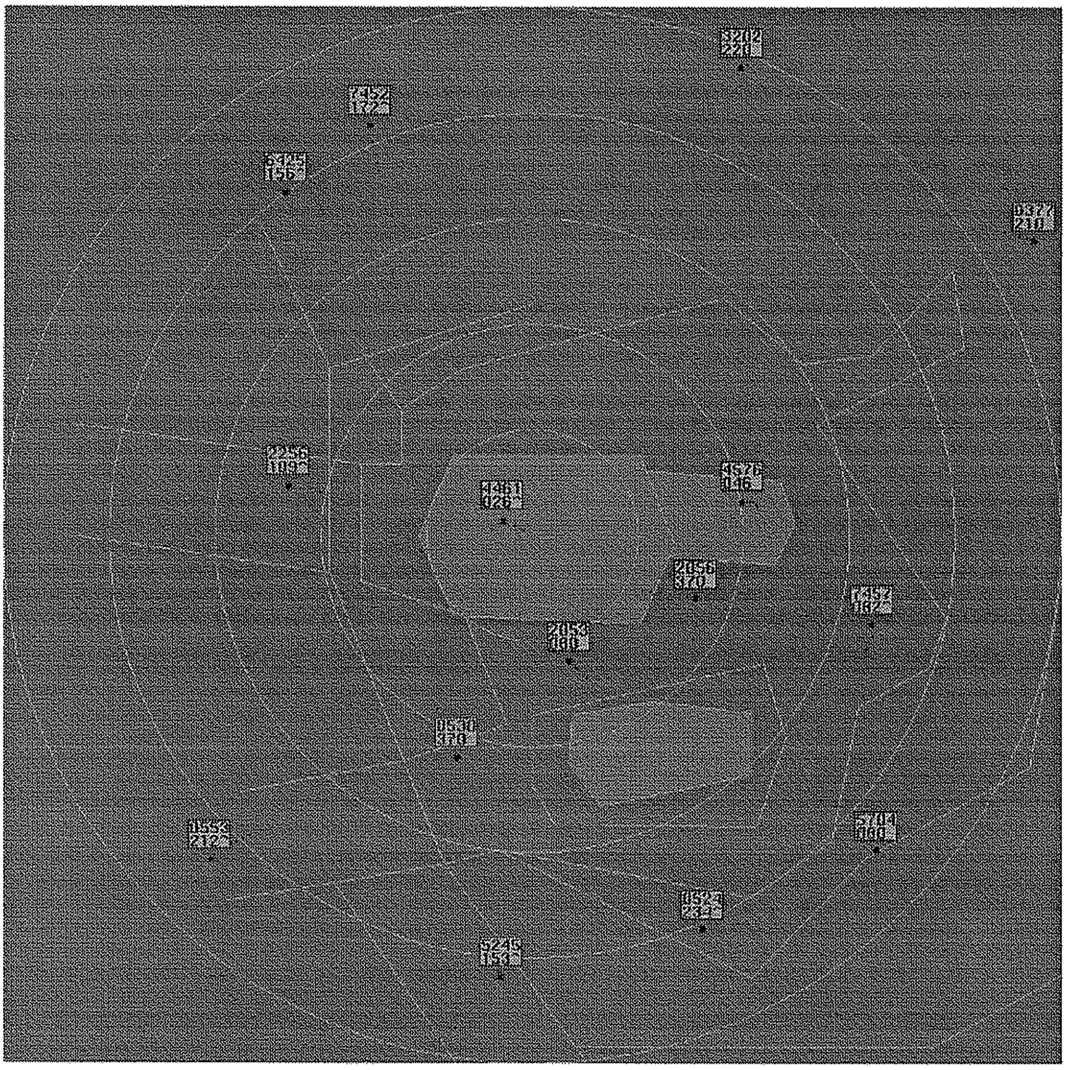


Figure 4.3: Pilot Experiment 2D Display Format

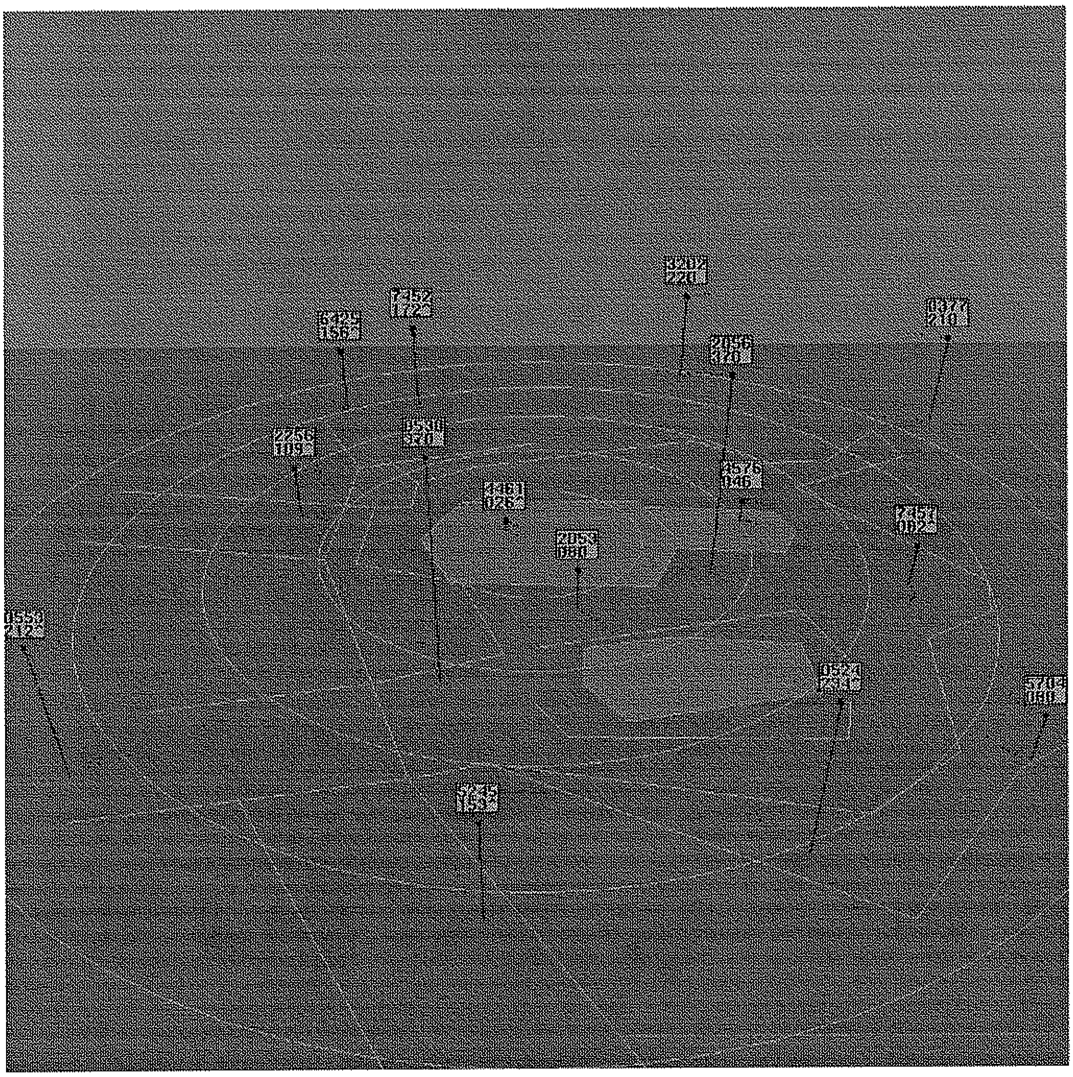


Figure 4.4: Pilot Experiment 3D Display Format: Perspective Projection

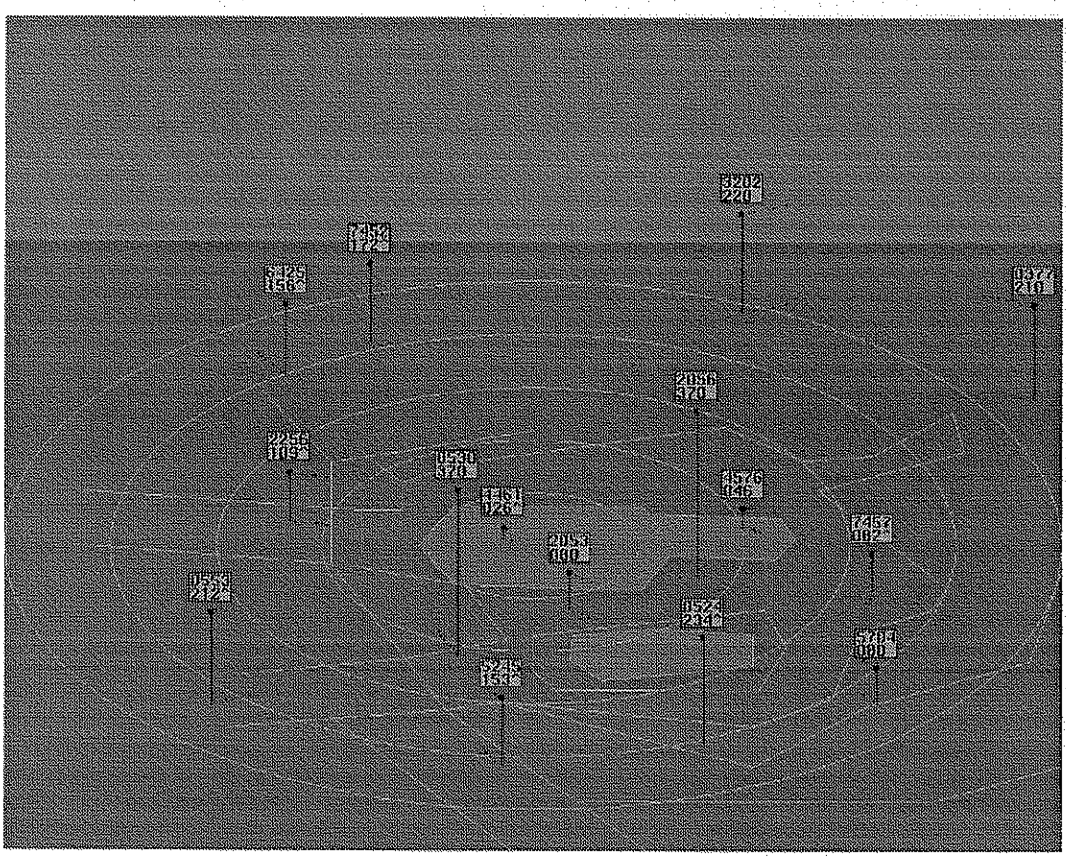


Figure 4.5: Pilot Experiment 3D Display Format: Parallel Projection

y-axis being North, and the world z-axis being vertically upwards (height). It was desired to have a 'North-up' display since this is the way that most ATC displays are orientated. For the 3D displays, the centre of projection was therefore located on the world y-axis at some distance to the south of the origin, with the horizontal component of the VPN pointing along the direction of positive-y (i.e. north) (see Figure 4.6). For the perspective projection 3D display, viewing parameters were as

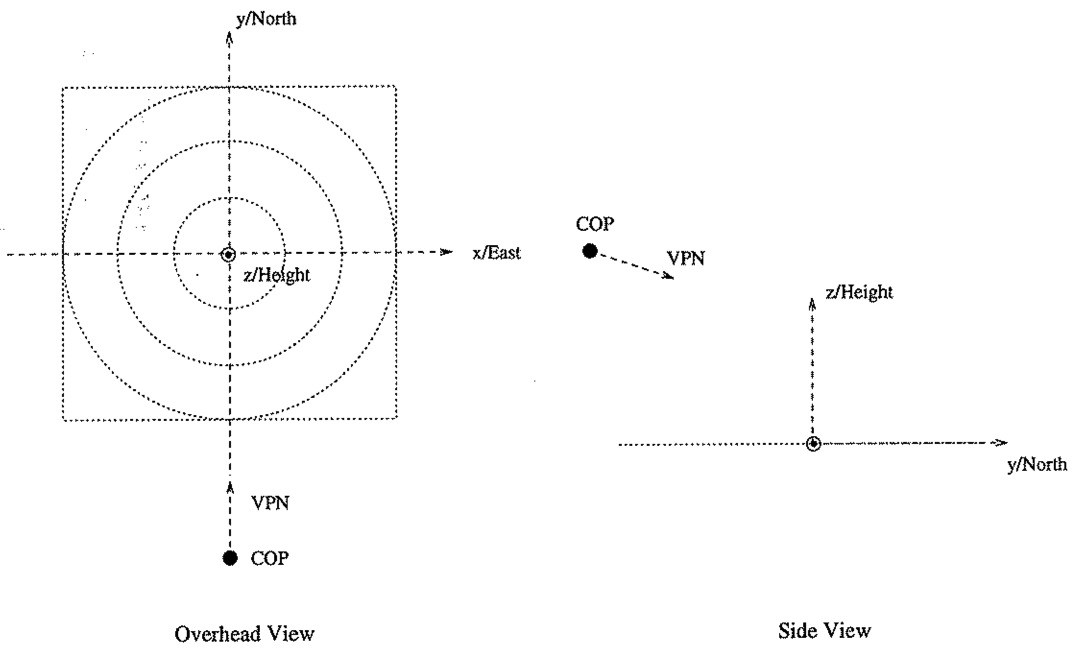


Figure 4.6: ATC Display Coordinate Systems

follows:

- Geometric field of view = 40°
- Viewing azimuth = 360° (i.e. looking from the south)
- Viewing elevation = -30° (i.e. looking down at an angle of 30°from the horizontal)
- Vertical exaggeration = ×4

The same viewing elevation, azimuth and vertical exaggeration were used in the parallel projection display.

Aircraft symbology consisted of a black filled circle for air position (this was as recommended in the NATS standard). In the 3D display, this was connected to the corresponding ground position by a 'drop line', as in the preliminary display. Each aircraft had up to eight 'trailing histories', represented by single-pixel dots. In the 3D displays, trailing histories also had ground shadows (with respect to a light source placed directly above the display at infinity). This idea was adopted from Strutt, the idea being that differences between the air histories and their ground shadows would enable vertical manœuvres to be seen more clearly. (If air histories alone are shown, it is not always clear whether a manœuvre involves any vertical component.)

As specified in the NATS colour standard, aircraft-related symbols were displayed in black on top of all other symbols (video map and datablocks) so that they were always visible. This may conflict with the depth cue of occlusion, however.

Datablocks were displayed as black text on an infilled grey background, with a black border around each. The standard font use by the GL graphics library was used, although this was far from ideal. (The characters 'V' and 'U' and the number '0' and the letter 'O' were not easily distinguishable; an alternative font was not designed due to time constraints in the implementation of the pilot study.) The datablock overlap avoidance algorithm developed for the preliminary display was used.

Datablocks were connected to the corresponding aircraft position symbol by leader lines. The nearest corner or edge of the datablock's bounding box to the aircraft's position was set at 6 pixels from the centre of the aircraft position symbol.

Datablock content varied with the task. The greatest amount of information shown in a datablock was:

- Mode-A code
- Mode-C code (altitude in hundreds of feet, or flight level)
- Climb/descent status symbol. This was derived from the current and previous mode-C code for each aircraft. A caret was used to indicate a climbing aircraft, a lower-case vee to indicate a descending aircraft.
- Cleared altitude (hundreds of feet) or flight level.

- Destination code. This refers either to a route or to an aerodrome; in the latter case, the last two characters of the aerodrome's ICAO code are used.

Figure 4.7 shows a schematic of an aircraft symbol and its associated trailing histories and datablock; the aircraft has a mode-A code of 1245, is currently at FL220 (flight level 220) and descending to its cleared level of FL110, and is bound for London Gatwick airport (ICAO code EGKK, abbreviated to KK in the destination code field). For decluttering in the 3D displays, datablocks only showed the mode-A

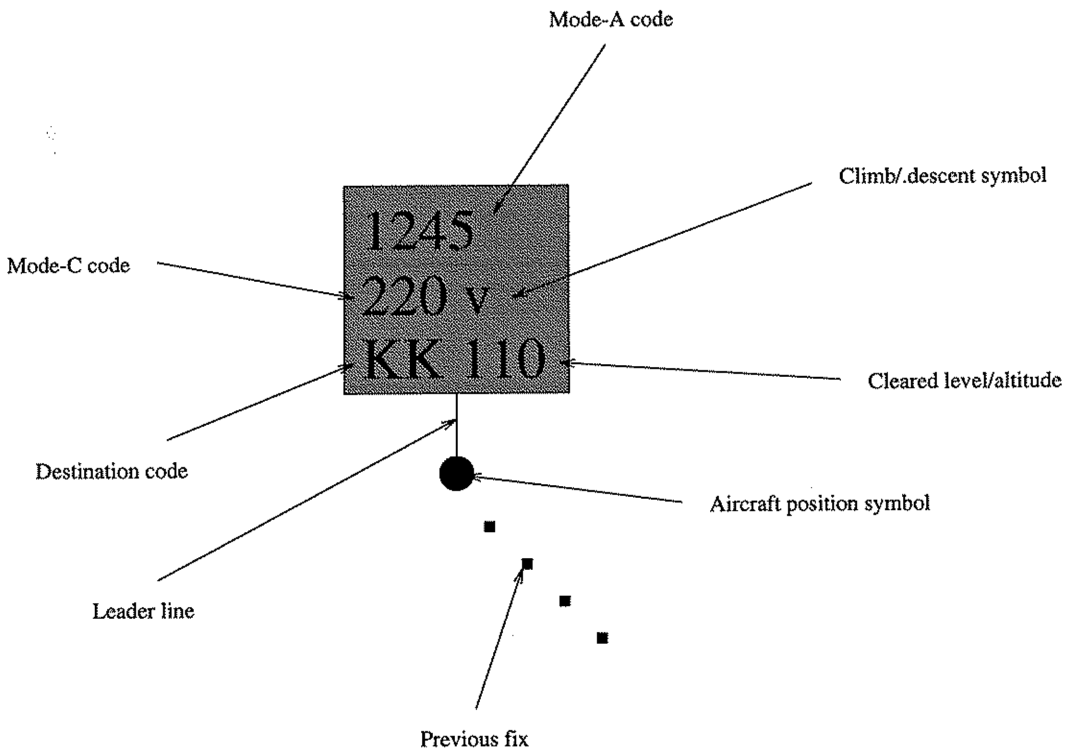


Figure 4.7: Aircraft Position Symbol and Datablock

code by default. Pressing and holding the SPACE bar on the workstation's keyboard caused the whole datablock to be displayed for all aircraft.

Aircraft positions were updated every 6 seconds (corresponding to radar sweep time). When the aircraft positions were updated, the symbols were moved in single pixel steps rather than in large jumps; this smooth updating was cited by Strutt as being less distracting and therefore preferable to a jump. However, the datablocks tended to 'jump' whenever the datablock overlap avoidance algorithm dictated a change to the datablock position relative to the aircraft symbol, since this was re-computed for every frame during the smooth movement of the position symbols.

4.3.6 Simulation

The simulation developed for the preliminary display was extended so that traffic information taken from real radar data could be mixed with simulated aircraft, to alleviate the task of designing realistic simulations of traffic movements. CAA NATS provided processed radar data taken from the Heathrow radar for the purpose of these experiments (15 minutes of data taken from about 09:00 to 09:15 on a morning in April 1994). This was in the form of an ASCII file. The implementation of software to process the radar file is described below.

4.3.7 Implementation

Platform

At around the time that the preliminary display evaluation ended, new equipment became available, in the form of Silicon Graphics *Indy* workstations. These have hardware double buffering, support for stereo buffering and time-multiplexed stereo displays, a 3D graphics library (called GL), and dedicated 3D graphics hardware (including a hardware z-buffer). These made them an ideal platform for continuing this research, and it was therefore decided to continue development on the *Indy* workstations. The software was therefore re-written in C and used GL for graphics.

Simulation and Radar Data File Processing

The simulation developed for the preliminary display was re-written to incorporate real radar data with the simulated aircraft. The same simulated aircraft mathematical model was used.

A block diagram of the simulation program is shown in Figure 4.8. The various parts of the simulation are explained below.

The ASCII radar data file was basically the processed output of the radar data processing computer. The file consisted of a number of records separated by newlines. At the start of a radar sweep, the time of the start of the sweep is written as a timestamp record. There then follow up to 256 aircraft records, which give information on the positions and identities of the aircraft which the radar 'sees'

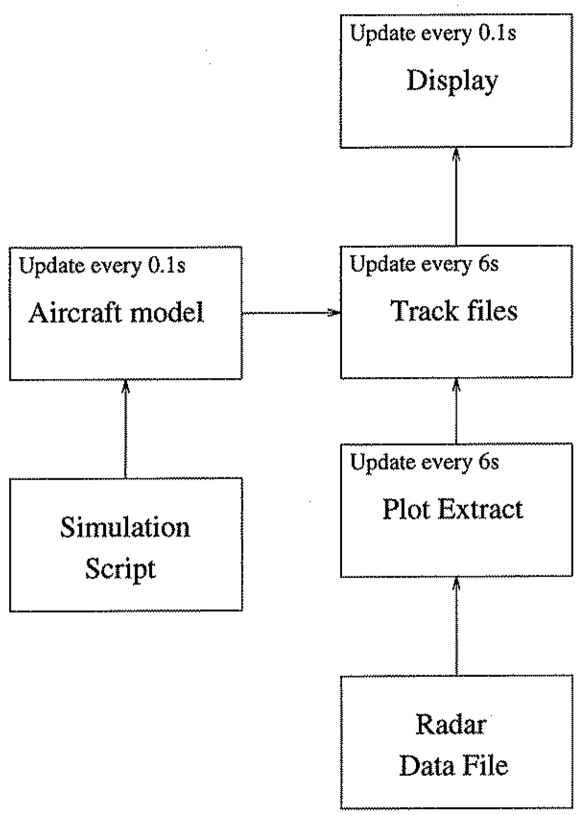


Figure 4.8: Pilot Study Aircraft Simulation: Block Diagram

during the sweep. The timestamp for the next sweep, and that sweep's aircraft information then follow, and so on. An aircraft record comprises a number of fields. The first field is a plot number in the range 0–255. Other fields are the x and y (east and north) positions relative to the radar head (in 30ths of a nautical mile), mode-A transponder code (where aircraft is transponder equipped) and mode-C transponder code (where aircraft is equipped with an altitude-encoding transponder). For a timestamp record, the first field is '-1' (to distinguish it from an aircraft record) followed by the time of the start of the sweep in the format hh:mm:ss.

The simulation program allowed the user to specify in an ASCII script file the plot numbers of the real aircraft to be displayed, and the information for the aircraft to be simulated (as in the preliminary display simulation).

Within the simulation itself, each aircraft, real and simulated, had a 'track file' associated with it. A track file was a data structure consisting of current position and altitude, and queue containing the 8 most recent positions and altitudes for the trailing histories. The displayed aircraft position data were derived directly from the track files. For each sweep of the radar, track file positions were updated either from the radar data file (for real radar targets) or from the aircraft mathematical model (for simulated radar targets). (The simulated aircraft mathematical models were actually computed every 0.1s, but the corresponding track files were only updated on every radar sweep.)

In order to maintain a track file correctly for a real aircraft, the position of the aircraft has to be correlated between successive sweeps. On the face of it, this is just an exercise in reading the aircraft positions from the same plot numbers of successive sweeps. However, there are two complicating factors. There are only 256 plot numbers available, and sometimes aircraft plot numbers may be reassigned. This tends to happen either when an aircraft leaves the radar's coverage, in which case the plot number may be used by another aircraft entering radar coverage, or when the aircraft's mode-A code changes. Further, sometimes the MTI radar processing may remove a plot, so an aircraft may 'disappear' for a few sweeps. This therefore requires an additional piece of 'plot extraction' logic.

The plot extractor initially assumed that an aircraft would have the same plot number between successive sweeps. On each new sweep, if the plot number was still extant and the position and altitude had not changed by more than a certain amount and the mode-A code was the same, then the new position in the track file would be updated from the plot numbered aircraft record in the radar data file. If there were significant changes in position and altitude (i.e. if it appeared that the plot number had been re-assigned to another aircraft) then the other radar data file records were searched to look for another aircraft at a similar position and altitude. In this case, no attempt was made to match mode-A codes since the plot number re-assignment might have been due to a mode-A change. If no matching plots were found, the track was 'coasted'; i.e. a new position was computed based on extrapolation of previous positions. An attempt to re-match the coasted plot to another plot in the radar data file was made over the next three sweeps; if no similar plot (i.e. one matching position, altitude and possibly mode-A code) was found after the plot was coasted for three sweeps, the aircraft's track file was deleted.

On a new sweep, if the aircraft's plot number was no longer extant, then the plot was coasted for up to three sweeps whilst trying to find a similar match in the radar data file. This was to compensate for MTI processing removing the plot temporarily. If a similar plot was not matched after three sweeps, the aircraft's track file was deleted.

This plot extraction logic performed satisfactorily with the radar data provided.

Stereoscopic Display Implementation

Stereoscopic versions of the 3D display formats were also implemented. This was done using time-multiplexing of the left and right eye images with liquid crystal shutter glasses (Stereographics *CrystalEyes* loaned from Division Ltd. for the purposes of the pilot study).

To present retinal disparity, a slightly disparate image must be presented to each eye (see §3.3.2). For a scene to be presented in stereo, the computer generates a pair of images (one each for the left and right eye), and these must be presented to the correct eye of the viewer. In the arrangement used in this study, the user wears a pair of 'glasses' which have liquid crystal cells in place of lenses. These liquid crystal cells are normally transparent; however, when a signal is applied to them, they turn black (opaque). The left and right eye images are presented on the computer's monitor alternately at 60 Hz. When each image is presented, the shutter glasses blank out the appropriate eye (the glasses are synchronised to the computer by means of an infra-red link, so the viewer is unencumbered by wires), i.e. when the left eye image is displayed, the right liquid crystal cell turns opaque whilst the left one remains clear, and vice versa when the right eye image is displayed. This arrangement is called *time-multiplexed* because the images are presented on the same display device, but multiplexed in time.

The model of *Indy* workstation used in the experiments did not have a hardware stereo buffer, so the XSGIStereo extensions to the X11 window system (which can work with GL) were used to implement stereo buffering. Two hardware frame buffers (front and back) are used for double buffering. Each frame buffer's memory is mapped directly to the screen: Half the memory is mapped to the top 50% of scanlines, and half to the bottom 50% of scanlines. To implement stereo, the top and bottom halves of a frame buffer are used to hold the left and right eye images, and the computer's monitor is switched to display either only the top 50% or bottom 50% of scanlines. Thus, the quad-buffering required for stereo animation can be achieved with the penalties of doubling the required rendering time (since two images must be drawn for each scene instead of one) and halving the display y-resolution.

The stereo images were computed using two virtual cameras. The positions and orientations of the virtual cameras were derived as follows (see Figure 4.9). In this

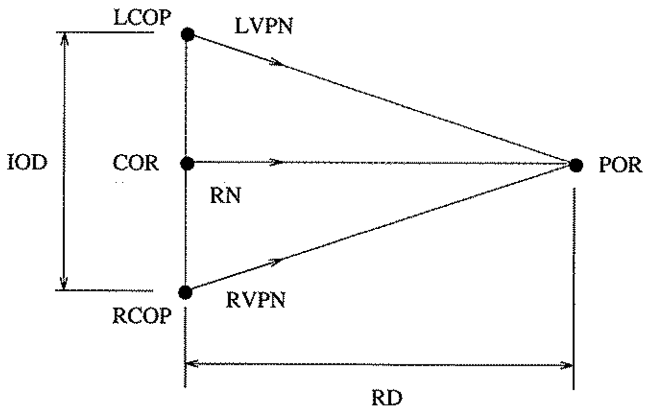


Figure 4.9: Stereo Viewing: Computing the Virtual Camera Positions

derivation, three parameters must be specified: A Point of Regard (POR), which is a point in the scene on which the two 'eyes' converge; a Centre of Regard (COR) which is the midpoint between the binocular cameras and would be where a cyclopean viewer would view the scene from; and an Interocular Distance (IOD), which is the separation of the binocular cameras. From these, the left and right centres of projection (LCOP and RCOP respectively) and the left and right view plane normals (LVPN and RVPN respectively) are computed by simple trigonometry. The Regard Distance (RD) is defined as the distance from the POR to the COR.

For the experimental stereo displays, the POR was fixed at the origin of WC space (i.e. the centre of the displayed airspace), the COR at the monocular COP, and the IOD was set to $\frac{1}{60}$th of the RD[1]. This was found to give sufficient disparity to be able to resolve the images with stereopsis, but not so much as to cause discomfort.

Stereo image generation complicated the datablock overlap avoidance algorithm, since datablocks which overlap from one eye's viewpoint might not overlap from the view of the other eye. The datablocks were therefore only drawn in the left eye's view and so were not displayed with any depth. It was originally thought that this would not present a problem, but the display could not be tested until just before the pilot study was due to commence because the CrystalEyes were not available until then.

When the display was finally tested, it was found that there was an incongruity between the datablocks being displayed with no depth and the corresponding aircraft being displayed with depth, which partially destroyed the impression of depth unless a conscious effort was made to ignore the datablocks. It was felt that this would not allow the stereo display to be evaluated effectively in the pilot study, so it was withdrawn at this stage. A solution was found for the main study (see Chapter 6).

4.4 Immersive Display Design and Implementation

4.4.1 Introduction

This section describes the design and implementation of immersive displays used for preliminary evaluation and the pilot study. Due to the differences between immersive and workstation implementations and restrictions in using the immersive technology, the immersive display was not used in the comparative empirical study; instead, an immersive air traffic visualisation was constructed to demonstrate concepts and gain subjective feedback about possible current and future applications.

As most readers will be unfamiliar with immersive displays, the equipment used

[1]This factor is a rule of thumb sometimes used in stereo photography, as determined through discussions on the Internet Usenet newsgroup dealing with stereo photography.

is first described. This is followed by a discussion of the design and development of the air traffic visualisations used in the preliminary investigation and pilot study.

4.4.2 Immersive Display Equipment

Three variants of two immersive display machines were used over the course of this research, all manufactured by Division Ltd. Division are continually developing the hardware and software of their immersive systems and as new systems were bought in, older systems were either replaced or became unsupported. Unfortunately, each system had unique sets of problems to work around, and some systems were incompatible with others such that application software had to be re-written or modified. The lack of a stable platform was one of the major problems encountered in this research. Other problems encountered were hardware reliability, lack of customer hardware support and inadequate software. (Manuals were at best terse, at worse incomplete or fictitious. Library routines sometimes did not perform as advertised and required non-obvious work-arounds.)

The principles of operation of each machine are the same. A Division immersive display machine comprises a multiprocessor computer system which carries out the management and simulation of the virtual environment (running a Virtual Environment Operating System (VEOS) known as dVS), a stereo image generator (comprising two identical channels for left and right eye image generation) feeding a stereo head-mounted display, a '3D mouse' (hand-held input device) and 3D position tracking hardware to track the positions and orientations of the head and the 3D mouse. The VE computer is hosted by another computer, either integral within the same physical unit or a separate workstation. This host computer provides operating system services to the VE computer (file I/O etc.) and a development platform for the programmer.

The VEOS is implemented using a number of processes distributed over a number of processors. A 'director' process manages communications between the processes/processors. Other processes (called 'actors') manage specialist tasks; for example, image generation (visual actor), sound generator (audio actor), and application processes. Actors communicate using a shared database called VL. Local copies of the elements of the database may be made by each actor for its own purposes, but changes must be registered to the global shared database. As an example, consider a virtual object being created and managed by an application actor. Part of the object's associated database will include a visual representation. This is first constructed locally by the application actor when the object is created. For the object to appear visually, the geometry element of the database is copied to the global database. This change is picked up by the visual actor, which then renders the object appropriately. Further changes to the object's graphical representation or position by the object's managing actor must be registered with the global database for the visual actor to update scene accordingly.

Database changes were notified in two ways. In earlier versions of dVS (e.g. 0.94), used in the preliminary display, actors were notified of changes through an event mechanism. Later versions of dVS (e.g. 2.0.6c, used for the pilot and main displays) used 'callbacks' or polling. The change from dVS 0.94 to dVS 2.0.6c resulted in a 'flag-day[2]' and so application programs had to be re-written.

The variants of immersive display machine used in this research, and the phases of the project for which they were used, are summarised in below. Details are discussed in the following sections.

1. **Preliminary Display.** Division *ProVision 200*: Sun 3 workstation external host, running SunOS Unix. Inmos Transputer-based 'core' running dVS 0.94; 'director' Transputer manages communications, with a flexible ring of Transputers running actors. The image generator was Division proprietary, with Intel i860 processors in the graphics pipeline and the Toshiba HSP chip [Tos91] for rendering (according to an engineer at Division); texture mapping was not supported. Four-button 3D mouse.

2. **Pilot Display.** Division *ProVision 100VTX*: self-hosted by an integral IBM PC clone running Concensys 4.2 Unix. 'Core' unknown. IG was Division proprietary (i860-based) with unknown rendering hardware; texture mapping was supported. VEOS was dVS 2.0.6c. Virtual Research Flight Helmet HMD. Five-button 3D mouse.

3. **Main Study Display.** *Provision 100VTX* running dVS 2.0.6c. As above, but with a PixelPlanes 2 image generator.

[2]Computer-speak for a change which is neither forwards nor backwards compatible; i.e. applications written for the old version will not work on the new version and vice versa.

Different HMDs were used from the earlier ProVision 200 and the later ProVision 100VTX, but they were of similar specifications. The Virtual Research Flight Helmet HMD (Figure 4.10), used with the ProVision 100VTX, has two colour LCDs of 360×240 pixel resolution and a horizontal field-of-view of 75°. The degree of binocular overlap is not easily adjustable and the interocular distance is fixed.

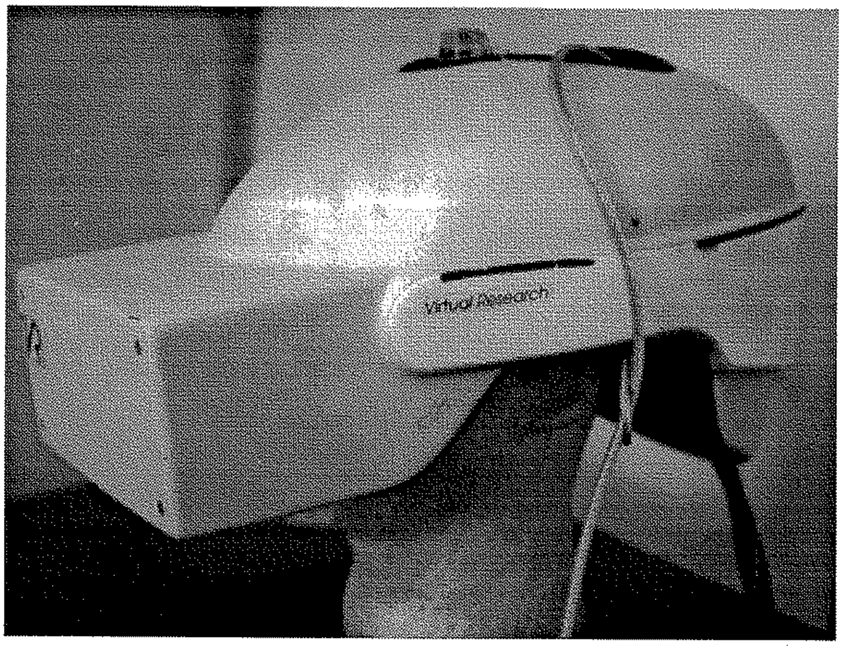


Photo: the author

Figure 4.10: VR System Head-Mounted Display

The fundamental graphics primitive is the 3D planar polygon. This is defined in world-coordinate space and may be Gouraud shaded, with optional texture mapping supported by the ProVision 100VTX. The system uses z-buffering for surface visibility computation. For rendering, polygons must be decomposed into a triangle mesh (possibly because the Toshiba HSP rending chip is optimised for triangle drawing—see [Tos91]). Software tools are provided for object construction. Drawing 2D objects (e.g. lines and pixels) directly to the display in device coordinates is not supported. The PixelPlanes II image generator additionally supports object transparency, although this was not (officially) supported by dVS 2.0.6c.

Hand and head position and orientation are sensed using two Polhemus position sensors. A single emitter radiates an alternating electromagnetic field which is detected by the two sensors (one on the HMD and one on the 3D mouse), which output their position and orientation with respect to the emitter to the ProVision system. The ProVision's world coordinate system is centered at the emitter, with the y axis vertical (positive upwards) and the xz plane being horizontal. World coordinates are scaled to real world inches.

Input is provided by a Division 3D mouse, which resembles a handgun with the barrel sawn off. Two types were used; a four-button system made from injection-moulded plastic components was used with the ProVision 200 for the preliminary display development, a more solid five-button device was used with the ProVision 100VTX for the pilot and main studies. The five-button mouse has three buttons under the thumb position, one under the index finger and one under the middle finger (Figure 4.11). Some button combinations can prove awkward for single-handed operation; for example, holding a top (thumb) button whilst trying to press the index finger button. The mouse is held in the right hand; its default representation in the virtual environment is as a disembodied half-closed right hand which appears to coincide with the position and orientation of the mouse in the real world.

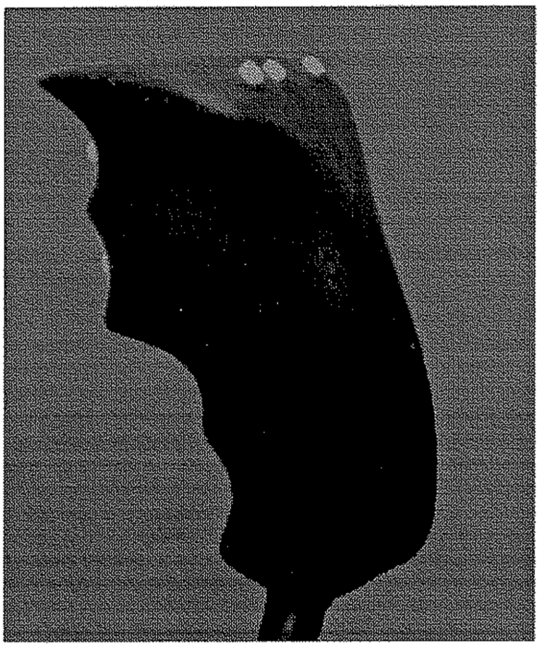


Photo: the author

Figure 4.11: Division 3D Mouse

The style of interaction with the system may be determined by the application program, but some defaults are provided. Movement in the virtual world is achieved using a 'flying' metaphor by default. The viewpoint is moved forward in the direction in which the hand is pointing along a horizontal plane in ProVision world coordinate space (i.e. along the projection of the hand's direction vector onto the world xz plane) whilst the left thumb button is pressed, and backwards in this direction whilst the right thumb button is pressed. If the index finger button is depressed simultaneously, this releases the constraint of horizontal movement.

4.4.3 Preliminary Display

As detailed in §4.2 above, a preliminary investigation was made regarding the application of 3D displays to ATC. In addition to the 2D and 3D versions described above, an immersive version was implemented and demonstrated to a small number of people, including one ex-air traffic controller.

Design and Implementation

The 'core' simulation used for the workstation displays was ported to a Division ProVision 200 running dVS 0.94; this necessitated re-writing the program in C, but the same object-orientated approach used in the workstation version was applied.

The display format was based on the 3D TTW format. A major problem was the restriction that all graphical objects must be polygons. Since the Provision 200 did not support texture mapping, the ground plane was represented by a grid composed of thin grey rectangular tubes, with navigation points as solid blue crosses. This gave the display a 'chunky' feel and caused problems where navigation point crosses intersected; as they were both displayed in the same plane and had precisely intersecting surfaces, z-buffering imprecision meant that sometimes part of a grid tube showed through a navigation point symbol, or vice versa. This was solved by 'raising' the navigation point symbols slightly so that their upper surfaces would be above the ground grid, but this looked a little strange. As with the 3D TTW display, aircraft were represented as elongated tetrahedra, with the long axis pointing in the direction in which the aircraft was heading. The 'drop line' was implemented as a thin tube.

The display of datablocks was problematic. Normally, text would be drawn as a series of character bitmaps directly onto the viewport in screen coordinate space; however, this is not possible under dVS. Each character has to be defined as a 3D planar polygon. As the characters existed in WC space, in order for the text to always face the viewer they needed to be re-orientated every time the viewpoint position and orientation change. In order to appear at a constant size (i.e. to subtend a constant visual angle on the display screens), they needed to be scaled by a factor proportional to the distance from the viewpoint. Constant-sized characters are necessary to ensure that they are always readable (allowing characters to become too large or two small renders them illegible). However, this had an incongruous effect when the whole scene was viewed from far away, the aircraft symbols being very small compared to the text, which dominated the view.

The re-scaling and re-orientation of text impacted severely on machine performance, especially if this was done every time a head-position change event occured. To alleviate this, a software timer was set up which set a flag every 0.1s. Each time the head-position change routine was invoked, the flag was examined and the text re-orientated and rescaled only if it was set (it was then cleared). Thus, text was re-orientated and rescaled at a minimum interval of 0.1s, reducing the load on the system.

Datablock overlap avoidance could not be implemented, since there was no way to determine where datablocks were projected to in screen space (the necessary graphics transformation matrices being inaccessible).

The characters themselves had to be a managed carefully. Each character exists as an 'object' under dVS in its own right, and the time required to create a new instance of an object is significant. Therefore, a 'pool' was maintained of instances of each character, and a flag was associated with each instance which indicated whether or not it was in use (the pool was implemented as a set of linked lists of instances, one list for each character type). When a 'new' character was required for display, the pool was first checked to see if there were any unused instances of that character; if there were, the instance was flagged as in use and returned. Otherwise, if no free instances existed, a new one was created, flagged as in use and added to the pool.

The fact that dVS events are asynchronous proved to be a problem. When the position of an aircraft was updated, the location of the position symbol, drop line and each text character in the datablock were updated correspondingly and the changes registered to the shared VL database as near simultaneously as possible. However, due to either the propagation delays of updates through the database, delays in the visual actor or both, the display updates by the visual actor took place at some later time such that the visual changes to the scene were non-simultaneous; the position symbol might move, followed by each datablock text character in turn in a sort of 'rippling' effect. This problem was compounded since, at the start of a radar sweep, all aircraft displayed positions are updated at once.

Discussion

People to whom the display was shown and who were inexperienced with immersive systems found the resolution poor, the equipment cumbersome (especially due to the trailing wires from the HMD and 3D mouse) and navigation within the virtual environment difficult without practice. These problems were common to nearly all applications on the immersive system, however, and were not peculiar to this implementation. Three major problems were found with the preliminary display implementation itself:

1. The number of polygons, coupled with the need to re-scale and re-orientate datablocks, was seriously affecting system performance.

2. Datablock clutter was excessive, due to two problems:

   (a) Datablock overlap avoidance could not be implemented.

   (b) Datablock text had to be fairly large to be legible, owing to the poor resolution of the display.

3. There was no effective way of implementing a video map.

The pilot study display (described below) and main experiment display addressed some of these issues. Despite these drawbacks, however, those to whom the display was demonstrated still felt that the concept was interesting and worth exploring further.

Following the preliminary investigation, it became obvious that including an immersive display in a comparison with other display formats implemented on a workstation by a series of empirical tasks would be difficult. Some of the tasks being designed measured time taken for the user to select particular aircraft, for example. On a workstation display, this could be done with a mouse regardless of the display format. In an immersive display, some alternative selection mechanism would have to be implemented which would interfere with timings. (Manually timing verbal responses was ruled out as being too imprecise.) A decision was therefore made to examine the immersive display only subjectively.

4.4.4 Pilot Display

Following the preliminary display investigation, a new Division machine, the ProVision 100VTX, was acquired. This had one major advantage over the ProVision 200: It supported texture mapping. It was hoped that this would allow for an adequate video map to be displayed and also reduce the number of polygons required (since the ground plane could now be represented by a few texture-mapped polygons instead of a grid or other representation made of significantly more polygons). Although this necessitated a switch from using the event-based dVS 0.94 to the callback-based dVS 2.0.6c, the application program was being re-written extensively anyway since the simulation 'core' was being revised to incorporate real aircraft tracks from radar data files, so this was not a significant drawback.

Design and Implementation

For the pilot display, the ground plane was represented by texture mapping the same video map as used for the workstation-based displays onto a polygon mesh (this was done simply by taking a 'screen dump' of the PPI display and using this as the texture pattern). The ProVision system uses linear interpolation for textures, and this introduces distortions under perspective projection because the texture map is not correctly foreshortened. An acceptable solution was achieved by subdividing the ground plane polygon into a regular grid of smaller polygons and mapping appropriate portions of the texture map onto each one [FVFH90, Lan91]. Although this yields only an approximate solution, in this case, subdividing the ground plane into a 4×4 grid gave acceptable results. (An exact solution would involve perspective correction during the texture interpolation process, but this is more computationally intensive and would necessitate modification of the dVS rendering code.)

This alleviated some of the performance problems encountered in the preliminary display by reducing the number of polygons required to represent the ground plane. Aircraft-related symbols were also redesigned to use fewer polygons. Aircraft were represented as white tetrahedra but with the base triangle omitted. The drop line 'poles' were made of triangular cross-section and did not have end 'caps'. However, the pilot display also added up to eight trailing histories per aircraft, both in the air and on the ground giving up to 16 additional objects per aircraft. Histories were represented as regular baseless tetrahedra, cyan coloured for air histories and yellow coloured for ground histories.

For the pilot study, it was decided to omit the datablocks altogether; these were re-introduced for the main study when satisfactory solutions to the problems of datablock clutter and system performance penalties caused by re-scaling and re-orientating the text had been found.

World scaling was $6'$ per 100 n.mi. horizontally (so the displayed area was $6' \times 6'$ and $2'$ per $10\,000'$ vertically).

4.5 Chapter Summary

This chapter described the process of design of the displays used in the pilot study. Preliminary displays were first constructed to demonstrate the feasibility of such displays and to explore design issues. The knowledge gained from implementing these displays was then applied in the construction of the displays used in the pilot experiment. A simulation of air traffic for the purposes of experiments was also developed. Finally, the chapter described the development of prototype and pilot study immersive air traffic displays.

Chapter 5
Pilot Study

5.1 Introduction

As stated in §1.2.2, this research is an exploratory empirical investigation into the use of 3D displays for applications such as ATC. It is based on comparing subject performance at part tasks over different types of displays (2D PPI, pseudo-3D and stereo-3D TTW displays), and in addition evaluating an immersive 3D display.

Chapter 4 discussed the development of the display format used in the pilot study. The aim of the pilot study was to gain further insight into the utility of these display formats, and to develop and validate the tasks and experimental techniques necessary to evaluate them. The findings of the pilot study were used in the design of the main study.

Evaluation consisted of both quantitative and subjective assessment of a number of part-task tests by a small group of pilot subjects:

1. An azimuth angle and relative distance reading task,
2. A memory recall task,
3. A conflict detection task.

The pilot study was rather more limited in scope than the main study. Nine ex-ATCOs working for the CAA in non-operational capacities were used as pilot subjects; no non-ATCOs were used in the experimental part of the pilot study. As described in §4.3.7, there were problems with the stereoscopic 3D TTW display, so this was not included in the study (although it was demonstrated to two individuals); the 2D PPI and 3D TTW display were used in the experimental investigation, and the immersive display was demonstrated.

This chapter first describes the design and methodology of the tasks. The pilot study experiment itself is then discussed, followed by the conclusions and implications for the main experiment.

5.2 Tasks: Motivation and Methodology

This section describes the part tasks used in the pilot study; the motivation for their selection, the methodology and experimental procedure used to explore them.

5.2.1 Projection, Azimuth Angle and Relative Distance

Introduction

In §4.3.4, the question as to which type of projection (perspective or parallel) would be more suitable for 3D air traffic displays was considered. A task was therefore devised to explore this question. From the results of this, it was intended to select a projection type for the main experiment and then to see how reading of angle and distance varied with display type using the chosen projection.

Methodology

More formally stated, it was desired to determine the influence of projection type λ_P on the observations of azimuth angle and horizontal relative distance between pairs of points (targets) representing aircraft positions in a 3D display.

Subjects were required to read the azimuth and distance between two targets T_1 and T_2 at the same altitude[1], each pair presented in a separate stimulus image. (Figure 5.1 shows a plan view of a stimulus.) Let the azimuth angle of T_2 from T_1 be λ_A and the distance between T_1 and T_2 be λ_D. These are the two main independent variables. The dependent variables are the observed azimuth angle y_A and the observed distance y_D. It is expected that these will depend not only on the stimulus azimuth and distance, but also upon the position of the targets in the display (especially on their depth). Define B to be a point half way along a line joining T_1 and T_2, and let the world x and y coordinates of this point be B_x and B_y respectively. B_x is a measure of left/right placement of targets within the display (if $B_x < 0$ the point B lies to the left of the centre of the display, if $B_x > 0$ then B lies to right right, if $B_x = 0$, the targets straddle the centre). B_y is related to the depth of the pair of targets (i.e. their distance from the observer) (Figure 4.6). From

[1]In a real air traffic position display, targets will generally be at different heights, in which case it will probably be necessary to provide suitable ground projections of the aircraft position to enable accurate determination of horizontal separation and azimuth angle.

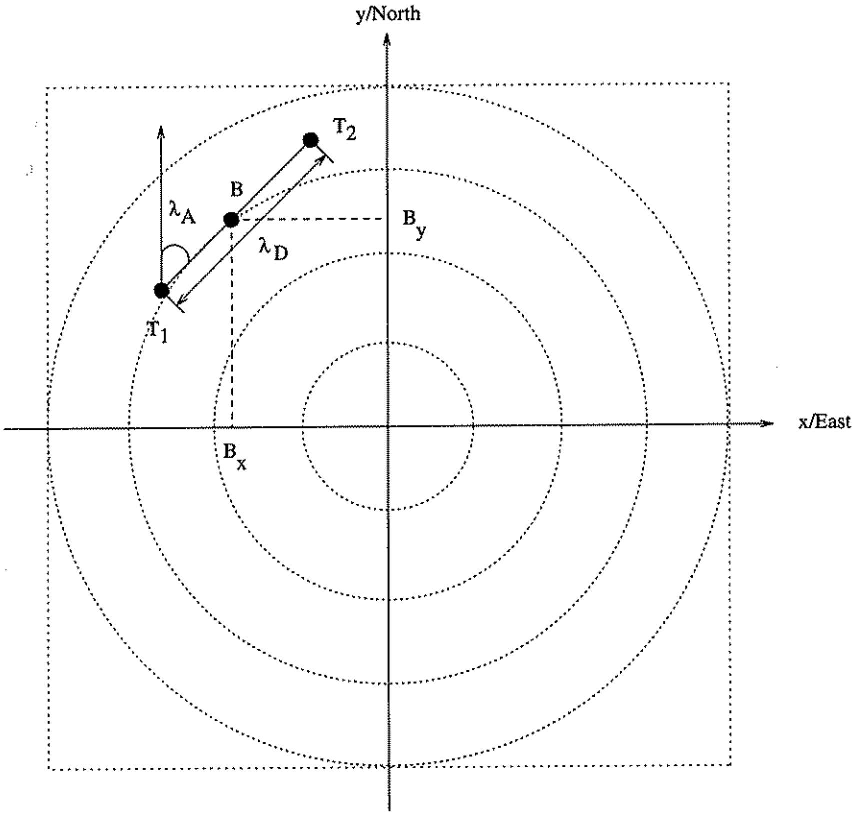


Figure 5.1: Azimuth & Distance Independent Variables

Figure 5.2, it can be seen that the relationship between depth d, camera elevation angle θ and B_y is given by:

$$\begin{aligned} d &= \tfrac{B_y - COP_y}{\cos\theta} \\ \Rightarrow \quad B_y &= d\cos\theta + COP_y \end{aligned} \tag{5.1}$$

i.e. $B_y = \alpha d + k$, a simple linear relationship; B_y increases in proportion to depth.

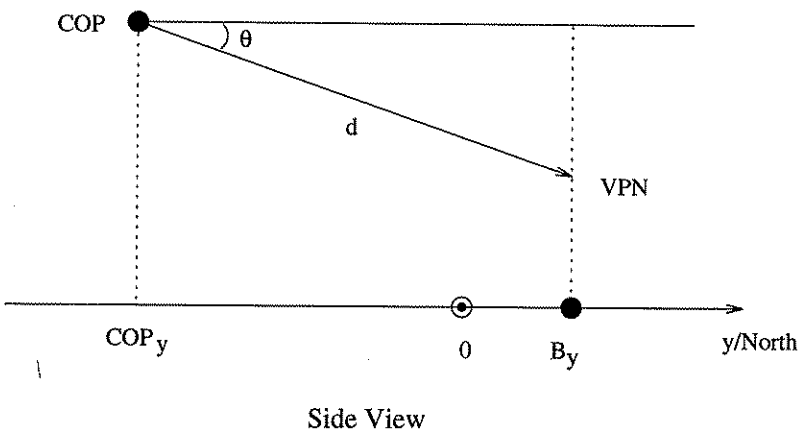


Figure 5.2: Relationship between z-Depth and B_y for 3D Displays

The dependent variables are expected to be functions of the independent variables:

$$y_A = f(\lambda_P, \lambda_A, \lambda_D, B_x, B_y) \tag{5.2}$$

$$y_D = f(\lambda_P, \lambda_A, \lambda_D, B_x, B_y) \tag{5.3}$$

In multiple regression analysis, $f()$ is assumed linear in coefficients and the *null hypothesis* is that the coefficients are all zero: i.e. that observed azimuth angle and distance are not influenced by projection type, stimulus azimuth angle, stimulus distance, B_x or B_y. The task experiment is designed to test the null hypothesis.

This experiment is really only concerned with the effect of λ_P on y_A and y_D, and in a sense the other independent variables are nuisance variables. A full investigation of the effects of these by systematic variation of each one individually is well beyond the scope of this thesis (each factor would merit a full study in its own right). The values of λ_A, λ_D, B_x and B_y were therefore generated randomly. The ranges of these variables are shown in Table 5.1 (displayed area was 100×100 nm). λ_P has two levels (perspective and parallel). It was decided to use an independent groups

−45nm	$\leq$	B_x	$\leq$	+45nm
−45nm	$\leq$	B_y	$\leq$	+45nm
−180°	$\leq$	λ_A	$\leq$	+180°
2nm	$\leq$	λ_D	$\leq$	20nm

Table 5.1: Pilot Task 1: Ranges of Randomly-Generated Variables

experiment design; the subject pool was divided randomly into two groups, with one group being allocated to the parallel projection display and the other to the perspective projection display. Twenty stimulus images were generated, and each subject was shown the same stimuli in a random order. For each stimulus, subjects recorded their responses (azimuth and distance observations) on a piece of paper.

Procedure

The task instruction sheet is shown in Appendix A, p. 254. Having read it, subjects were shown the display, which was then explained. Subjects were then given a response sheet on which to record their answers, and wrote down the first response with the supervisor watching. That having been done, the supervisor left them to complete the task on their own (although being nearby in case of any problems).

5.2.2 Memory Recall

Introduction

In §2.6.4, it was postulated that display format might have an effect on the memory of an air traffic scenario, in particular relating to speed of assimilation into memory, speed and accuracy of recall, retention time in memory and the ability to chunk information.

This experiment was designed to make a preliminary exploration of this area. The approach adopted was an adaptation of the work of Burnett and Barfield, as described in §2.6.4. Burnett and Barfield used air traffic controllers as subjects. The subjects were required to memorise flight strip information relating to a dynamic scenario and were then shown the scenario over a 70 second period. They were then required to reconstruct the last frame of the scenario. However, the author felt that this approach did not concentrate sufficiently on the effects on memory of the display alone (due to the memorisation of the strip data) and would also complicate matters when carrying out an experiment with non-air traffic controller subjects. The methodology chosen here was therefore to try to memorise a *static* traffic scenario, recalling information presented solely on the display.

The pilot task was designed to look at the influence of display format on the accuracy and speed of recall of spatial relationships. (The main experiment task extended this to look at the effect of increasing the amount of information to be remembered.) Two display formats were used: 2D PPI and pseudo-3D TTW with perspective projection.

Initial Hypotheses

In a 2D plan-view, vertical information is represented textually, whereas in the 3D display it is shown graphically. It is expected that an approximate spatial mental model will be formed faster given a 3D presentation than a 2D presentation for two reasons:

1. Pictures are more useful for storage of spatial codes in working memory than text [Lea94], and might thus be assimilated more quickly.

2. In the 2D PPI display, the integration into an internal 3D spatial model must be done mentally, whereas it is already presented in an integrated analogue form in a 3D display.

A 3D display should therefore give a lower mental workload in forming an integrated spatial mental model than a 2D PPI display and thus be quicker to assimilate, and may have a more durable memory trace (since the relationships are remembered in analogue form which is more useful to working memory than text). Regarding positional accuracy of the recalled image, however, it is expected that the recall of horizontal positional information in the 2D display will be more accurate than in a 3D display, since in a 2D PPI display, horizontal information (horizontal distances and azimuth angles between targets) is shown unambiguously, whereas in a 3D display it is subject to ambiguity unless viewed in plan.

To summarise, it is conjectured that 3D displays will be better for conveying rapidly approximate spatial relationships because these can be seen directly as an analogue picture. However, this will be at the expense of analytic detail, such as precise horizontal and vertical position.

Methodology

Define λ_P as display format (2D PPI or 3D perspective projection), time to recall a scenario y_t and accuracy of recall (a measure yet to be defined) as y_a. It is expected that y_t and y_a will be functions of λ_P. The null hypothesis for this task is therefore that display format will have no influence on the speed and accuracy of recall of static air traffic scenarios.

To test the null hypothesis, subjects should be presented with one or more stimulus images for a period of time. These are then removed and the subject required to reconstruct the image from his or her memory on a piece of paper.

Some preliminary research was conducted into how many aircraft should be in the stimulus scenarios, and how long the stimuli should be presented for. Too long a presentation with too few aircraft might remove any differences in performance between the display formats; more aircraft for a shorter a period of time might make differences in performance between the display types more apparent (since spatial relationships would have to be assimilated very quickly) but too short an exposure period with too many aircraft might overload the subjects. Moreover, different stimuli may be remembered to different degrees in different ways (for example, subjects might be able to visualise a simple geometric pattern for a small number of aircraft, but this might break down with larger numbers of aircraft). The best advice received from psychologists consulted was to try to determine the number of aircraft and the time of presentation empirically. It was decided that each scenario should contain 15 aircraft and be presented for 60 seconds. This was intended to place a deliberately heavy burden on the memory capacity of the subjects so that any differences in speed of assimilation might be shown, but it was thought that this might have a demoralising effect.

Subjects were randomly allocated to one of two groups: One used the pseudo-3D perspective display, the other used the 2D PPI display.

Procedure

The instruction sheet for the task is shown in Appendix A, p. 255. Two static scenarios were presented, the first for training. For each scenario, subjects were asked to try to memorise the ground positions and altitudes of each of the 15 aircraft shown. After 60 seconds, the scenario was removed and the subjects were given a piece of paper on which was printed the displayed area in plan view, complete with range rings and video map outlines. Subjects were required to indicate on the piece of paper the positions of the aircraft and, if possible, their heights. Heights could be recalled by writing a number and/or by drawing the length of the 'drop lines' (vertical lines connecting each aircraft to the ground, representing height) for each aircraft. (Both displays had datablocks with the mode-C code; however, only the 3D display represented height graphically.)

The reconstruction was to be timed by the supervisor, and evaluated in terms of accuracy of recall (number of aircraft, and horizontal and vertical placement error) and time taken to reconstruct the scene.

5.2.3 Conflicts

Introduction

As related in Chapter 2, the prime task in air traffic control is to 'maintain the safe, orderly and expeditious flow of traffic', and this is done by long-term (strategic) planning using flight strips, and short-term (tactical) planning and monitoring using the radar. The rôle of a 3D display in strategic planning is difficult to assess as it would involve long interactive scenarios. This research therefore concentrates on whether or not a 3D presentation is more or less effective in the purely monitoring and tactical planning rôles; more specifically, using the radar-derived position information to detect and resolve potential conflicts. Detection of conflicts may be taken as a measure of controller awareness—if controllers are more aware of the current state of the traffic they will detect more conflicts. It is postulated that because of its integrated presentation of the three spatial dimensions and reduced mental integration workload, there will be greater situational awareness of the traffic state with the 3D display than with the current PPI.

Having detected a conflict, the way in which it is *resolved* is also important. This may require split-second decisions to be made and these are critical, since aircraft must be instructed to manœuvre to avoid conflict with each other and also so that they do not come into conflict with other traffic in the vicinity as a result. As described in §2.6.3, McGreevy and Ellis's CDTI study showed that pilots made more traffic avoidance manœuvres in the vertical plane with a 3D display than with a plan-view presentation augmented with digital height readouts. If similar behaviour were demonstrated in the ATC context, this may give controllers an 'extra dimension' in which to work when resolving conflicts, although Hopkin states that whether this would be desirable in ATC is a moot point [Hop94].

Burnett and Barfield's Conflict Experiment

Burnett and Barfield's study (described in §2.6.2) included an experiment in which controllers were required to detect impending conflicts in 2D and 3D perspective displays. Traffic density (two levels: 7 and 17 aircraft) was crossed with display type (two levels: 2D plan-view and 3D perspective display). Subjects were provided with tabular flight plan data (akin to flight strip information) and shown dynamic scenarios using a time-calibrated slide-projector to present successive frames of animation at 12 s intervals, corresponding to radar sweep updates. Each scenario comprised six frames and contained three conflicts and a random number of controller instructions. Prior to presentation of each frame, subjects were informed of one of four situations:

1. A controller instruction(s) to be issued at the subject's discretion,
2. impending conflict; the experimenter informed subjects of conflict type and conflicting aircraft, and provided three possible resolutions of which only one was correct,
3. both (1) and (2),
4. no response required; subjects were advised to use the time to review the traffic situation.

Response time from presentation of a frame to verbal response was measured manually using a stopwatch. The study found that conflict resolution performance in the 7-aircraft scenarios was faster for the 3D format than for the 2D format, but little difference was observed between the formats for the 17-aircraft scenarios.

Methodology

The null hypotheses are:

1. Conflict detection performance (in terms of time to detect a conflict and number of conflicts correctly detected) is independent of display type, and

2. type of conflict resolution instruction (e.g. horizontal or vertical manœuvre) is independent of display type.

This task was designed to test these hypotheses.

For this research, it was felt that Burnett and Barfield's approach contained two main problems:

1. The need to interpret tabular flight data excluded non-ATCOs from the study unless given additional training, and

2. subjects were warned in advance as to whether or not an impending conflict existed.

A different approach was therefore adopted for this task. This task involved detecting and resolving conflicts in a dynamic scenario, but subjects were not informed as to whether or not the scenario contained a conflict, nor on the number or type of conflicts. All responses were at subject discretion. No tabular information would be available; all information would be extracted from the display. Conflict detection performance measures would be speed of identification of potential conflicts, and correct identification of conflicts.

For controllers to detect conflicts, some route information is required, and this was incorporated in the aircraft's datablock. However, it was recognised that this information would be useless to non-ATCOs.

For the experiment, subjects were randomly allocated to one of two groups: 2D PPI or pseudo-3D perspective display. Two dynamic scenarios each of 3 minutes duration were presented, and subjects were asked to identify conflicts and give resolution instructions. The scenarios contained one conflict each, and were based on real traffic samples, with one simulated aircraft to cause the conflict. The first scenario had two aircraft on the same airway travelling in the same direction, but slowly converging, with the lower climbing through the level of the upper. The second scenario contained a crossing conflict: An aircraft level at high altitude against a climbing aircraft crossing its path. This second conflict actually resulted in a merged plot on the display, and a minimum separation of less than 0.1 nm (i.e. a probable collision). The time to detect the conflict and the resolution instruction(s) issued were recorded for the second scenario, the first being used for training.

Procedure

The task instruction sheet is shown in Appendix A, p. 256. When subjects had read the instruction sheet, a real traffic sample was shown in order to familiarise them with the dynamic display, whilst the instruction sheet was read by the supervisor. The familiarisation display showed all traffic movements within the displayed area contained in the radar data file, minus general aviation traffic (i.e. all aircraft with squawk code 7000) and non-mode-C transponder-equipped aircraft (for which no height information was available). Additionally, traffic below an altitude of 3 000′ (i.e. a mode-C code of less than 30) was filtered out.

In the 3D display, the use of the SPACE key in presenting a full datablock was demonstrated. Subjects were then left to examine the display for unlimited time until they reported ready. The two scenarios were presented, the first for training purposes. The supervisor manually timed subject's responses with a stopwatch and noted any instructions which they issued.

5.3 Pilot Experiment Findings

5.3.1 Introduction

This section describes how the pilot experiment was carried out in practice, and presents the findings.

5.3.2 Overall Procedure

Nine subjects participated in the pilot study over the course of a week. These were ex-ATCOs currently working in various non-operational capacities within CAA NATS. Subjects had a variety of operational backgrounds, ranging from civil ATC without radar to military ATC. All were enthusiastic and interested in the research, and were commendably objective, despite having reservations about the applicability of 3D displays to current operational practices.

Tasks were conducted with the 2D PPI and pseudo-3D TTW displays described in §4.3.5. The stereoscopic display was demonstrated rather than used in the experiment due to development problems as explained in the previous chapter. After the tasks were complete, subjects were invited to comment on the experiments and were offered the opportunity to use the VR display (described in §4.4.4).

The experimental groups for each task were allocated before the arrival of any subjects. For the azimuth and distance reading task, subjects were allocated to use either a parallel or perspective projection 3D display. For the remainder of the tasks, each subject was allocated to use either a 2D PPI display or a 3D perspective projection display—i.e. each subject used the same display format for all tasks apart from the first. Thus, during the experiment, a single subject might use a 2D and a 3D display, or two different 3D display formats, or a single 3D display format. For a full study, it would be desirable to have subjects using one format of display throughout.

Instructions to subjects were presented by instruction sheets (Appendix A). On arrival, subjects were asked to read a background information sheet and the instruction sheet for the first task (azimuth angle and distance reading experiment) whilst it was prepared. It was explained that the experimental programme was a pilot study, and that open, constructive criticism was welcome. It was intended to have subjects write comments at the end of the whole experiment. Questionnaires were not used, the type of information required having yet to be determined (apart from demographic information, which was not important for the pilot study).

Tasks were presented in the same order to each subject:

1. Azimuth Angle and Relative Distance Reading.

2. Memory Recall.

3. Conflict Detection.

Prior to each task, its appropriate instruction sheet was given to the subject while the task was being set up. The supervisor then repeated the instruction sheet and answered any questions. The subject then performed the task, after which it was discussed before moving to the next task. (It had been originally intended to have subjects fill out a comment sheet having completed all the tasks. However, this proved to be impractical, since the subjects were commenting whilst the tasks were in progress, and if comments had been deferred until the end of the experiment, some useful feedback may have been lost.)

Having completed the quantitative tasks, subjects were given the opportunity to try the VR display prototype. Subjects were asked to read the instruction sheet, which described the VR display and its operation, before trying the display. Before donning the HMD, subjects were again briefed verbally, and the HMD components pointed out. When subjects put the HMD on, they were given the 3D mouse control device and its correspondence with the position of the computer-generated 'hand' was explained. Navigation in the virtual environment was then described and subjects were asked to perform simple navigation tasks until they were used to it. They were then allowed to explore the environment at their leisure for unlimited time. The environment showed a 15 minute scenario from a real traffic sample.

5.3.3 Azimuth Angle and Relative Distance

General

The pilot study was carried out with ex-ATCOs of varying experience. Several commented on the fact that they were 'rusty' and that current ATCOs would be able to estimate angles and distances with greater proficiency, this being done routinely in radar control work. Out of the nine subjects who took part, four were shown the parallel projection, five the perspective projection.

None of the subjects had used a 3D display previously. Most reported that subjectively, the non-uniformity of ground distance with position on the display (and allied to this, the distortion apparent in the range rings) presented a difficulty. Regarding the perspective display, one subject commented that the fact that the drop lines of the same *actual* height would appear to be *different* lengths depending on the depth at which they were displayed, and that this would negate any usefulness which they might otherwise have.

One subject commented on the lack of a north index in the display for a reference, but was able to perform the task when it was indicated that the virtual camera azimuth was 360°.

One subject reported that estimation of distance presented more of a problem than heading.

One subject who had had operational experience within London Air Traffic Control Centre reported that he used his local knowledge (knowledge of relative bearings and distances of features on the ground map) as an aid.

A couple of subjects used a pen up against the display screen to measure distances instead of 'eyeballing'. (The instructions did not prohibit this—in fact, did not mention it.) This technique is not uncommon amongst operators of plan view radar displays.

Instruction Sheet

The instruction sheet used an incorrect technical term (it talked about *heading* rather than the correct term which was *bearing*).

The instructions also failed to clarify whether to use the air plot or the ground position of aircraft position for the task (in fact, in this task it made no difference since the aircraft were co-altitude, but some subjects requested clarification).

Even though the instruction sheet pointed out that the bearing was to be taken from aircraft 1 to aircraft 2, this still caused some initial confusion with some subjects providing the opposite direction bearings for one or two responses until they recognised their mistakes and queried the supervisor. (Although the bearing of target 2 from target 1 is just the opposite direction bearing of that from target 1 to target 2, and subjects corrected their mistakes by adding (or subtracting) 180° from the first bearing they gave, requesting the bearing or that of its opposite direction may influence perception.)

All of the above would probably be rectified by clearer instructions, perhaps illustrated with examples (either on the sheet itself, or on the computer as a 'training' exercise).

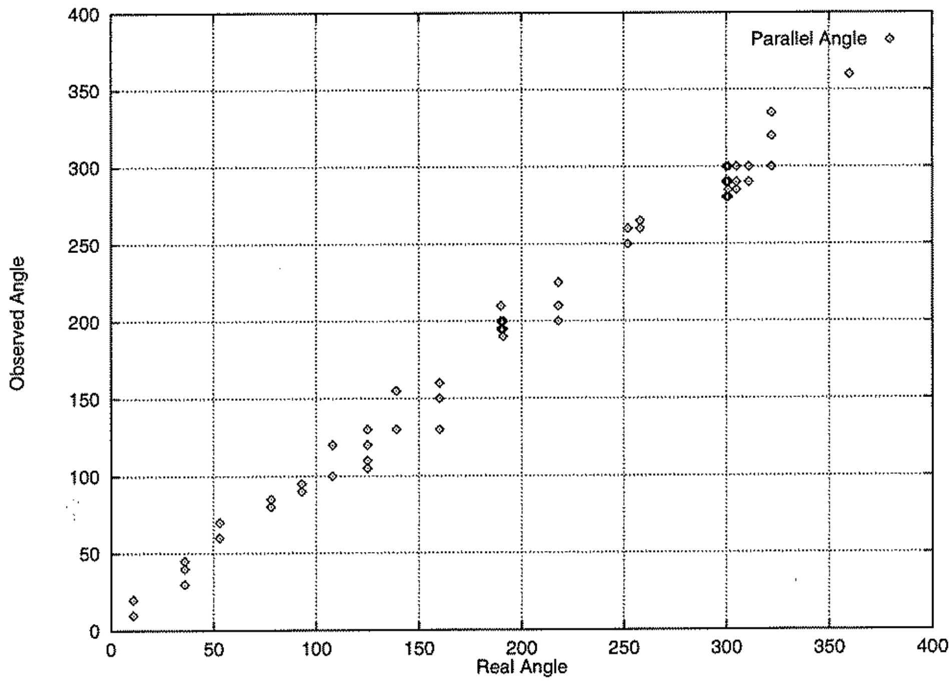
Azimuth Angle Observations

Figure 5.3 shows scatter plots of the observed azimuth angle y_A versus stimulus azimuth angle λ_A for parallel and perspective projections.

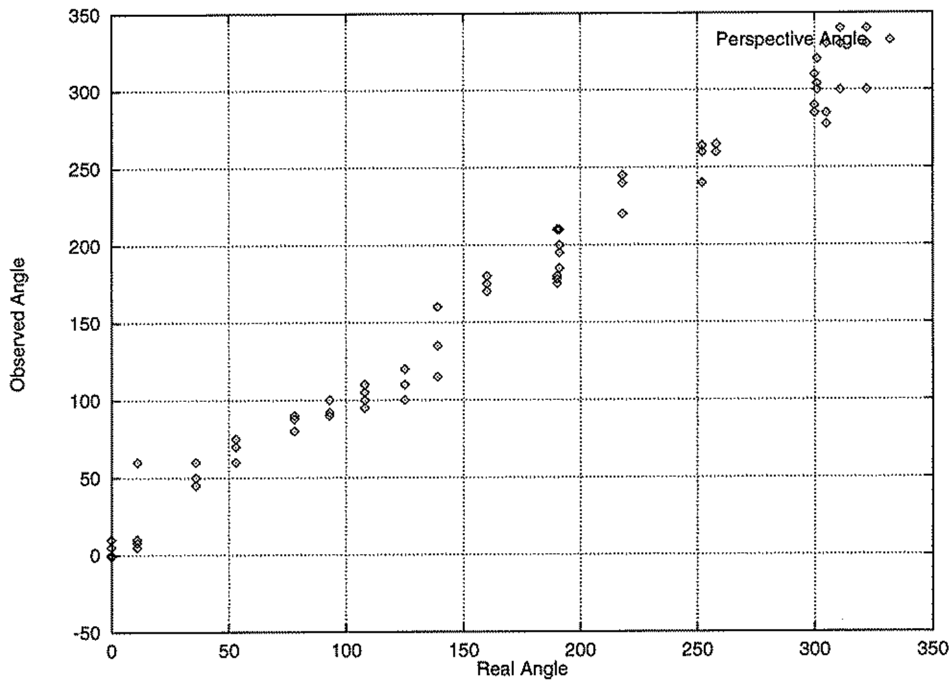
The data were analysed using linear regression analysis to fit a linear model to Equation 5.2. This gave the following:

$$\text{Parallel} \quad y_A = 6.43 + 0.14B_x + 0.96\lambda_A \quad R^2 = 0.99, p < 0.01 \tag{5.4}$$

$$\text{Perspective} \quad y_A = 26.01 + 0.82\lambda_A \quad R^2 = 0.72, p < 0.01 \tag{5.5}$$



Parallel Projection



Perspective Projection

Figure 5.3: y_A vs. λ_A Scatter Plots

This suggests that there are significant differences between observations of azimuth angle under the parallel and perspective projections used. For the parallel projection, the λ_A coefficient is closer to unity and the constant term is smaller than for the perspective projection. This suggests that observations of azimuth angle under the parallel projection will be closer to the true azimuth angle than under the perspective projection; in other words, the parallel projection affords greater accuracy *for the viewing parameters chosen in these displays, all other things being equal.*

The results suggest that there is a correlation between observed azimuth angle and B_x in the parallel projection case, but not in the perspective projection case. Equation 5.4 suggests that for the parallel projection, when targets are in the left half of the display, there is a tendency to under-estimate the true azimuth, and when the targets are in the right half of the display, there is a tendency to over-estimate the angle.

In both parallel and perspective projections, the results suggest that there is no significant correlation between observed azimuth angle and B_y.

Distance Observations

Figure 5.4 shows scatter plots of observed relative distance y_D versus stimulus relative distance λ_D.

Linear regression was not valid for distance analysis since as can be seen in the figures, the variance of λ_D is not constant with y_D. Therefore, a logarithmic fit was tried. Regression analysis gave:

$$\ln \lambda_D = 1.4 + 0.083 y_D \tag{5.6}$$

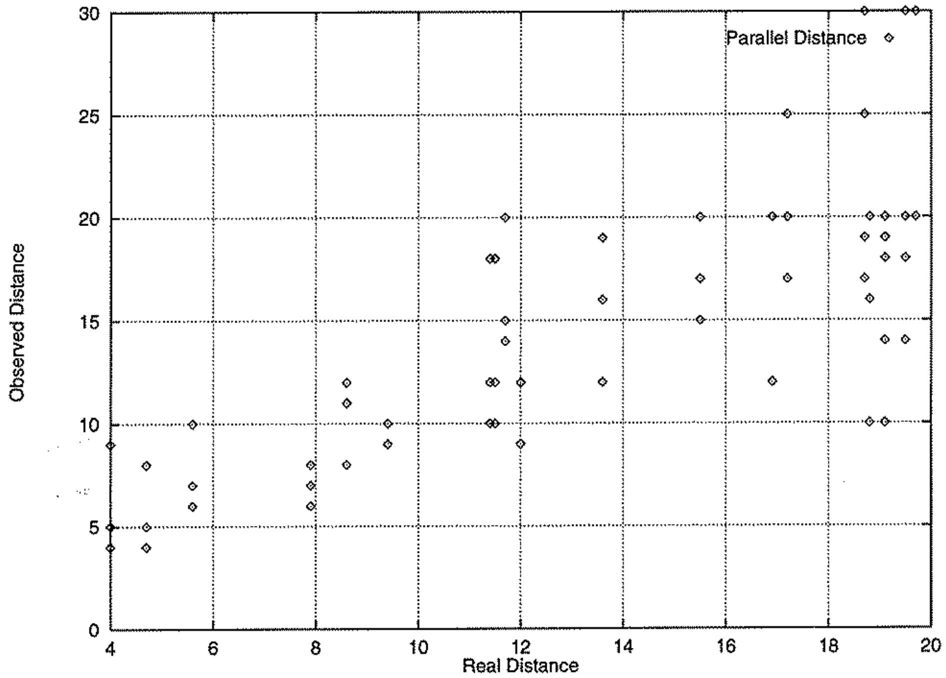
$$\Rightarrow \quad \lambda_D = 4.06 e^{0.083 y_D}, R^2 = 0.73 \tag{5.7}$$

This model was the same for both parallel and perspective projections. The observed relative distance therefore unexpectedly appears to be independent of projection type, B_x and B_y.

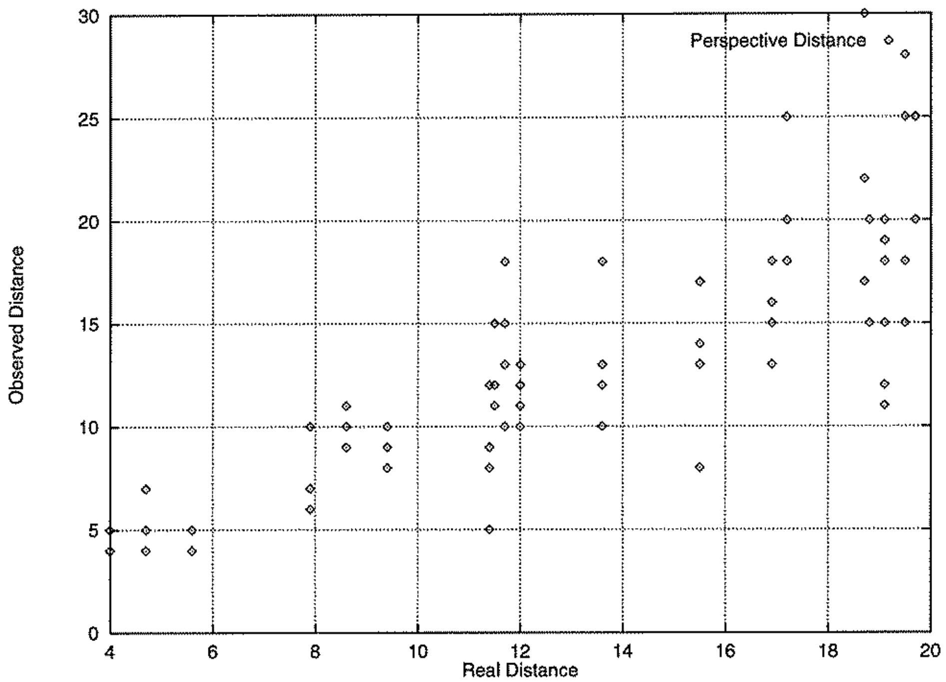
5.3.4 Memory Recall

General

This was the least effective experiment in the study, but still yielded useful information. Subjects generally found the task to be very difficult, due to the sheer amount



Parallel Projection



Perspective Projection

Figure 5.4: y_D vs. λ_D Scatter Plots

of information to be committed to memory and the short time of presentation. The results were of such poor quality (i.e. the recalled patterns bore so little relationship to the stimulus patterns) that they were not analysed here.

It was intended to use the first scenario for training and the second as a recorded run, and to present both identically. In practice, however, because of the unfamiliarity with the 3D display format in particular, the first training scenario was presented whilst the supervisor re-read the instruction sheet to the subject and explained the format of the display, and subjects were then allowed to examine the display for unlimited time until they reported that they were ready to continue. Subjects reported using this time to familiarise themselves with the format of the display (particularly in the case of the 3D presentation), and to work out a method of memorising the scenario. When they reported ready, the second scenario was presented for 60s and subjects were asked to reconstruct it. It was intended that the reconstruction be timed. However, the presence of the supervisor meant that the subjects tended to make comments whilst performing the reconstruction and so the time of finishing was difficult to determine.

Queries raised during the reconstruction soon highlighted the fact that there should have been a full training run prior to the recorded run, with the same procedures as the recorded run. This would have allowed any questions to be raised during the training run (during which time the supervisor would be on hand) instead of in the recorded run.

The reconstruction task given the 3D stimulus presented problems in that subjects were asked to draw the approximate lengths of the height poles if they could not remember the actual heights, but only one subject did so. This was not made explicit in the instructions, and should perhaps have been clarified by an illustration in the instruction sheet. One subject commented on the potential difficulty of reproducing a display presented in 3D in a plan view. This may have been a factor for more than one subject. In both display formats, subjects were reluctant to guess heights, often preferring to omit these, although whether or not this was because they genuinely couldn't remember the heights even approximately, or was due to some other factor wasn't clear.

One subject reported that the task had a demoralising effect. This may have had an impact on the following task (conflict detection). As related by one subject, ATCOs as a group generally find their job stimulating and take pride in being able to perform a challenging task well. The author postulates that the demoralising effects of this task may therefore have had a more severe psychological impact than may be expected with other groups.

One subject was accidentally shown the stimulus for 2 minutes, and seemed to fare better at the recall. This may indicate that the difficulty found in the task was due to the sheer amount of information and/or lack of time, rather than being inherent in the memorisation and recall of a scenario *per se.*

Task Performance

The most interesting aspects of this experiment came as subjects related the ways in which they were performing the task.

It was expected that subjects would remember the traffic position in terms of the overall spatial pattern, and try to recall the heights in the 2D display by reading the datablocks and in 3D display by a combination of memory of the length of the height poles as well as the datablock. In fact, the subjects did not discernibly conform to this behaviour.

In the training scenario, some subjects attempted to devise a strategy for memorising the scenario. One said that she started at the centre of the display and worked outwards, another that he tried to quarter the display area and remember the traffic in each quarter. In general, however, the method of memorising the traffic did not appear to be visually based in terms of chunks of spatial patterns, as expected, but in terms of traffic patterns and approximate levels; Some remembered traffic in terms of it being at a high, medium or low level, and tried to classify the aircraft in terms of a familiar traffic behaviour pattern, for example, Heathrow inbounds, overflights, etc.

5.3.5 Conflict Detection

Familiarisation

Before the task was presented, subjects were shown a dynamic scenario for familiarisation. Due to the large area displayed (compared to the average sector size) and the busy time of day (around 9:00 a.m.), there was a lot of traffic in the scenario and some subjects were initially overwhelmed. This may have had a slight negative influence on the subjects (however, it was pointed out that there would be far fewer aircraft in the actual task).

Most subjects examining the 3D display tried to work out a strategy for using this display at this point, and one spent over 10 minutes on the familiarisation.

Task Performance

The supervisor was on hand the whole time, and subjects were encouraged to talk about what they were doing and how they were doing it. Although this was very useful, it made it difficult precisely to time when each conflict was identified.

The use of the datablock to present the cleared level and destination code was unfamiliar to subjects, and a lot of conflicts were reported which on closer examination of the cleared levels, would have been revealed as not being conflicts at all (mostly aircraft being cleared to a level above that of lower traffic). One controller was confused when the aircraft did not obey his (verbal) instructions—it was not pointed out sufficiently well that the scenarios were fixed and not interactive. When this was explained to later subjects, this had the side effect that once aircraft were classified as conflicting (or potentially conflicting), they were ignored.

One subject used the SPACE key as a *de*clutter button in the 3D display, in the opposite sense to that intended—i.e. he held it down most of the time to display the full datablocks, and released it where he wanted the display decluttered.

All subjects identified the slowly converging overtaking traffic conflict, but weren't sure about the precise distances involved. They tended to be conservative; e.g. the tracks were converging, so controllers tended to stop the lower aircraft's climb, sort out the lateral separation problem with a course change, then resume climb when lateral separation was seen to be clear.

The crossing conflict was missed by some subjects, or identified too late in the case of one subject, especially on the 3D display. Several sources of confusion were cited as reasons:

1. The airspace sector shown was very large compared to real sectors, and the first few subjects tended to 'assign' themselves to a particular sector and ignore other traffic! Later subjects were told explicitly to consider all traffic.

2. It was expected that the drop lines would be used as a cue to height and that the 3D display would be sufficiently clear. However, the display tended to

appear very cluttered, and separation of a cluster of aircraft along the same line of sight was difficult. Two air plots could appear next together even though their ground positions were well separated due to height differences (i.e. the depth cue of 'height in the visual field' was acting as a nuisance cue), and there were insufficient depth cues to separate the two easily.

- The perspective projection exacerbated this since aircraft near to the horizon tended to be clustered together.
- Perspective projection also caused problems in that fixed-size drop lines vary in displayed length according to its depth, negating their value in the opinion of two subjects.
- With two or more aircraft along similar line of sight, it was difficult to determine heights and which air plot corresponded to which ground position.

3. Correct identification of aircraft was made more difficult by a datablock crossover problem where several air plots were shown in the same region of the display. (See §5.4.)

ATCOs tended to look at lateral and vertical separations separately. From the comments as the task was conducted, some tended to look at the lateral separations of the ground plots, then try to follow the height poles up to the air plots to find the associated labels to read the heights. This was difficult if several aircraft were along the same approximate line of sight due to clutter and possibly insufficient depth cues.

5.4 Display Evaluation

5.4.1 2D and 3D Displays

Colours

Colours were generally acceptable. Some subjects were familiar with the colour standard; only one subject said that he did not like it, but that was a criticism of the NATS standard rather than the implementation.

Area

The displayed area was generally considered to be too large compared to the size of sectors which ATCOs usually control. This was no problem in the parameter reading task (except perhaps that the separations were viewed against a much larger area than normal) but this interfered with the conflict detection experiment, to the extent that some subjects were ignoring traffic in some parts of the display as they had 'assigned' themselves to one particular sector (e.g. Heathrow approach) and were ignoring traffic outside that sector (e.g. overflights or Gatwick traffic) as non-pertinent.

Datablock Format

There was some unfamiliarity with the datablock in the conflict detection experiment. Since the rationale was to get subjects to extract information purely from the display and to dispense with flight progress strips, the cleared altitude was included in the datablock next to the destination code. (The full datablock format used is shown in Figure 4.7.) Since the presence of the cleared flight level was unfamiliar to the subjects, it tended to be ignored. The general consensus was summed up by one subject:

> Datablock layout [is] unfamiliar—and needs time to get used to.

Another subject liked the fact that the cleared level and actual level are not vertically adjacent but offset by the route code, since this helps to distinguish the two and reduce the chance of one being read for the other.

The familiarity problem could be solved by greater exposure and training before the main experimental run.

Legibility of Symbology and Datablocks

Generally, subjects found that symbology viewed in isolation to be legible; however, some commented that the trail dots were rather small in the 2D and 3D TTW versions (being single pixels on a high-resolution display) and information was generally difficult to see against a datablock, even though aircraft-related symbology is guaranteed not be obscured by datablocks. One subject commented:

> Trail data [are] difficult to read when garbled with a datablock.

This was generally supported by the other subjects. Strutt suggested the use of XOR plotting to alleviate this difficulty.

Datablock 'Cross-Over'

There was a problem in the datablock overlap avoidance algorithm in that datablocks related to targets drawn close to each other could 'cross', so that it could be difficult to associate a datablock with a given target (Figure 5.5). There were also problems

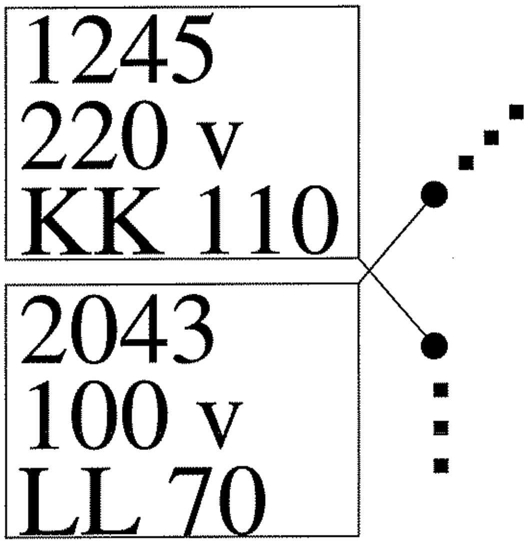


Figure 5.5: Datablock Crossing Problem

when the datablocks actually did overlap, in determining which block belonged to which aircraft. Part of the problem is related to the thinness of the leader line (only one pixel) and the subsequent difficulty in seeing it.

5.4.2 VR Subjective Trial

All but one subject tried the VR display (and the one who didn't was constrained by time rather than by lack of interest). All enjoyed the experience and found it to be novel and fun. Once navigation was mastered, none reported any problems with using the display. No problems were found with motion sickness or disorientation.

Although the technology was limited (in terms of resolution, display update rate and the cumbersome apparatus) subjects were objective and tended not to be put off by the present limitations, preferring to view the technology as developing and to see whether or not it could be made use of in an ATC context.

It was felt that the egocentric perspective afforded by an immersive display might influence the way in which subjects viewed the data, and this was indeed supported by observations. Some subjects were also pilots and they commented that the viewpoint was more that of a pilot than an ATCO, suggesting that such an egocentric display might be useful in cockpits, (presumably with a see-through HMD, or a TTW display) for gaining an awareness of the disposition of other traffic (which at the moment has to be done by monitoring radio communications and is therefore vague), but would not be useful from a controller's point of view with current working practices. Some subjects were fascinated watching the traffic patterns: Watching inbound aircraft enter a stack and then leave it to head for the extended centreline, turning onto it and landing, or watching outbounds climbing out from an aerodrome, then turning and climbing over a stack or going under it. Two subjects used the display for over 15 minutes and one for over 30 minutes (even though the radar data was only for 15 minutes and so had to be replayed). One commented that he could see the climb angle by the angle of the air trailing histories particularly clearly and so see whether or not a particular aircraft would be able to climb over the stack at present rate of ascent. (The symbols were not sufficiently clear or large to enable him to do this in the TTW display.) Subjects also tried to place themselves inside aircraft to follow them or to position themselves at an airport to gain a 'tower' view of the traffic.

Regarding potential applications, one subject thought that it might be useful in airspace planning. Here, airways are planned and tested by using simulated aircraft with representative performances to fly the routes and procedures. These are displayed in 2D plan and profile views; however, the subject felt that the pilot's perspective as afforded by the VR display might give a better feeling for how well the airway is suited to the simulated flight profiles. A couple of subjects thought of potential applications in training, for debriefing trainees after exercises.

Two applications were proposed to subjects, who were asked to comment. One was a low cost control tower visual control room simulator. Full-size simulators are constructed for training, with large visual displays providing a panoramic view, but some training could be done using a much cheaper immersive display. The other was a see-through HMD for control tower applications on which taxiways and aircraft could be shown under conditions of restricted visibility superimposed on the outside world. At present, under restricted visibility, ground movement control is done using a surface movement radar, which gives a plan view. A see-through HMD would give an overlay picture instead, and would not be restricted to aircraft on the ground (aircraft could appear as a radar target box, perhaps with mode-A and mode-C information appended). Most subjects thought both applications feasible and potentially useful.

5.5 Discussion: Implications for Main Study

5.5.1 Introduction

The pilot study helped to further develop the display formats, hypotheses and ideas of how to test them. Much was also learned about how individuals experienced in air traffic control reacted and behaved when performing the tasks, and how they carried out the task of ATC. This section reviews the overall lessons and implications of the pilot experiment, what it revealed about the major hypotheses and better ways or ideas for examining them, and the implications for the main experiment design.

5.5.2 Subject Behaviour

One of the things that emerged was the extent to which previous training and experience influenced subject behaviour, and this validates the use of a control group of non-ATCOs for the main study. For example, when presented with the conflict detection scenarios, subjects reported that they mentally 'filtered' traffic to concentrate only on aircraft of interest, even though in the conflict detection scenarios presented, all aircraft were nominally of interest. For example, controllers reported filtering aircraft by height (high, medium, low) as well as geographical location. An approach controller is not interested in overflying traffic; a controller managing an upper airway is not concerned about traffic in a terminal area beneath him. Some controllers also concentrated on areas where they had had previous operational experience, rejecting traffic in other areas (for example, concentrating on London Heathrow traffic and ignoring Gatwick traffic).

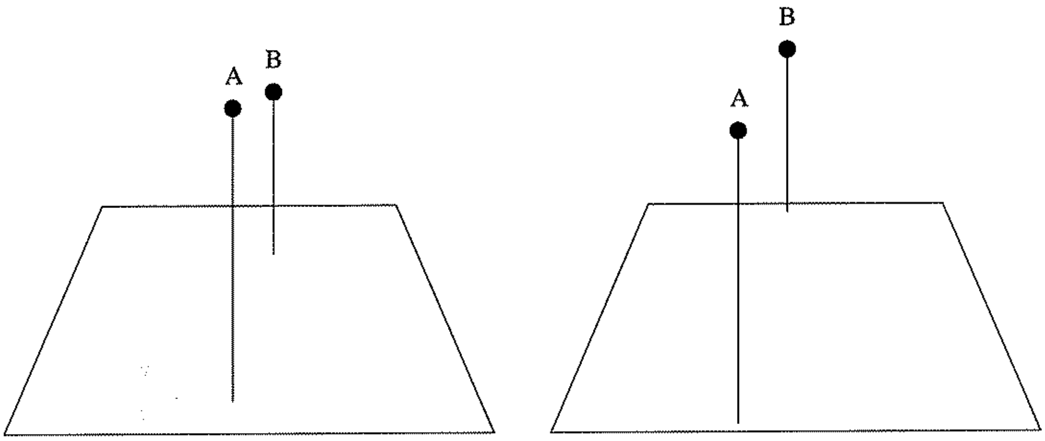
It was at this stage that the tendency of controllers to treat horizontal and vertical dimensions 'separately', as related in Chapter 2, first came to light. As controllers were describing how they carried out the conflict detection task (by a 'running commentary'), some said that with the 3D display, they were looking at ground positions and tracks, then following the drop lines up to the datablocks (which were attached to the air plots) to find the heights. (Controllers have individual styles, however, so this should not be taken as representative without further research.) If aircraft ground positions and tracks did not reveal a conflict, there was no need to look at the air plots to find vertical information.

The filtering of traffic by height and the emphasis on horizontal separation led directly to the development of a task for the main study; investigating how speed of extracting height and horizontal proximity information would vary with display format.

5.5.3 Position Ambiguity in the 3D Display

The 3D display was found to be very cluttered, partly as a result of displaying such a wide area and so tending to put aircraft likely to be in conflict close together. In some ways, it presented too much information for conflict detection with the behaviour described above, the air plots being largely redundant. The author speculates that if the datablocks were attached to the ground position instead of the air position (as in Burnett and Barfield's display), there would have been a tendency to ignore the air plots altogether. This might be partly due to trained behaviour with the 2D PPI display, and partly due to the fact that the air plots could be confusing; two air plots might appear close together but their ground plots might in reality be far apart, as shown in Figure 5.6(a). Also, from the verbal evidence of the controllers and observations of their behaviour, the author speculates that height in the visual field may have been a nuisance cue; an air plot towards the top of the display may have been interpreted as 'high', even if the actual altitude of the aircraft was not, when the length of the drop line is not properly considered, as shown in Figure 5.6(b). These problems are exacerbated where clutter is high (for example, with multiple aircraft along a similar line of sight) since it may not be easy to determine which ground position corresponds to which air plot. This is particularly problematic in the case of the perspective projection since linear perspective tends to show aircraft deeper into the display closer together.

It was postulated that a contributing factor to these problems was an inadequate



(a) Ambiguity of Position
Air plots of A and B may be close whilst their ground positions may be far apart.

(b) Height in the Visual Field as a nuisance cue.
Failure to consider ground position may result in B being perceived as at a greater altitude than A.

Figure 5.6: 3D Display Problems

sense of depth in the display. For example, in Figure 5.6(a), two air plots may be displayed in close proximity where they are actually far apart spatially. With insufficient depth cues, they might be mistaken as close together spatially, but more depth cues might better convey their separation. One subject who was shown the stereoscopic display had previous experience at viewing stereoscopic photographs (having earlier been an RAF photographic interpreter) and related that the display 'jumped straight in' when viewed in stereo. It was therefore anticipated that adding stereopsis to the 3D TTW display would give significant differences in task performance.

5.5.4 Selection of 3D Projection

One of the objectives of the pilot study was to select a projection type (parallel or perspective) for the 3D display formats in the main study. A perspective projection 3D image was found to have a number of problems:

1. The clutter problem described above,
2. subjects reported that the variation in length of drop line with depth negated its value in making relative height judgments,

3. readings of azimuth angle were found to be significantly less accurate with the perspective projection display than with the parallel projection display.

Item (2) above was only from verbal evidence and not investigated experimentally; however, the other two problems alone were sufficient reason to reject the perspective projection for the main study, and to use parallel projections for the 3D TTW display formats. (It is not possible to vary the projection type in the immersive 3D display; a parallel projection in such a display might appear to be very strange to the viewer.)

5.6 Chapter Summary

A pilot study was conducted to examine further the utility of the developed display formats, and to develop and validate experiments to evaluate them. First, the design of the experiment tasks, the motivation for adopting them and the anticipated results were introduced. The findings of the experiments, and the results of subjects' evaluation of the displays were then presented, and finally the implications for the main study were discussed.

Chapter 6
Main Study

6.1 Introduction

This chapter describes the design and implementation of the main experimental study of this research. The results are described in the next chapter. The development of the displays used in the main study is first described. This is then followed by the design of the experiment tasks.

6.2 Display Implementation

6.2.1 Introduction

The display formats used in the pilot study were further developed for the main study in light of feedback from the pilot experiment.

6.2.2 Changes to PPI and TTW Displays

The display formats adopted for the main study are shown in Figures 6.1 and 6.2. A number of changes were made to the 3D TTW display formats (see Figure 6.2):

- As discussed in §5.5.4, a parallel 3D projection was found to be preferable over a perspective 3D projection, and was therefore adopted for the 3D TTW displays used in the main study.
- In order to make aircraft ground position more clear, this was represented by a white filled circle.

For all displays, the size of the displayed area was reduced to 50 × 50 nm.

Due to of the datablock crossover problem, it was decided to remove the automatic datablock overlap avoidance and to display all datablocks in the 12 o'clock position with respect to the appropriate aircraft plots. This then created the problem of how to view overlapping datablocks. The solution adopted for the 2D PPI was that if one datablock (partially or totally) obscured another, if the user 'clicked'

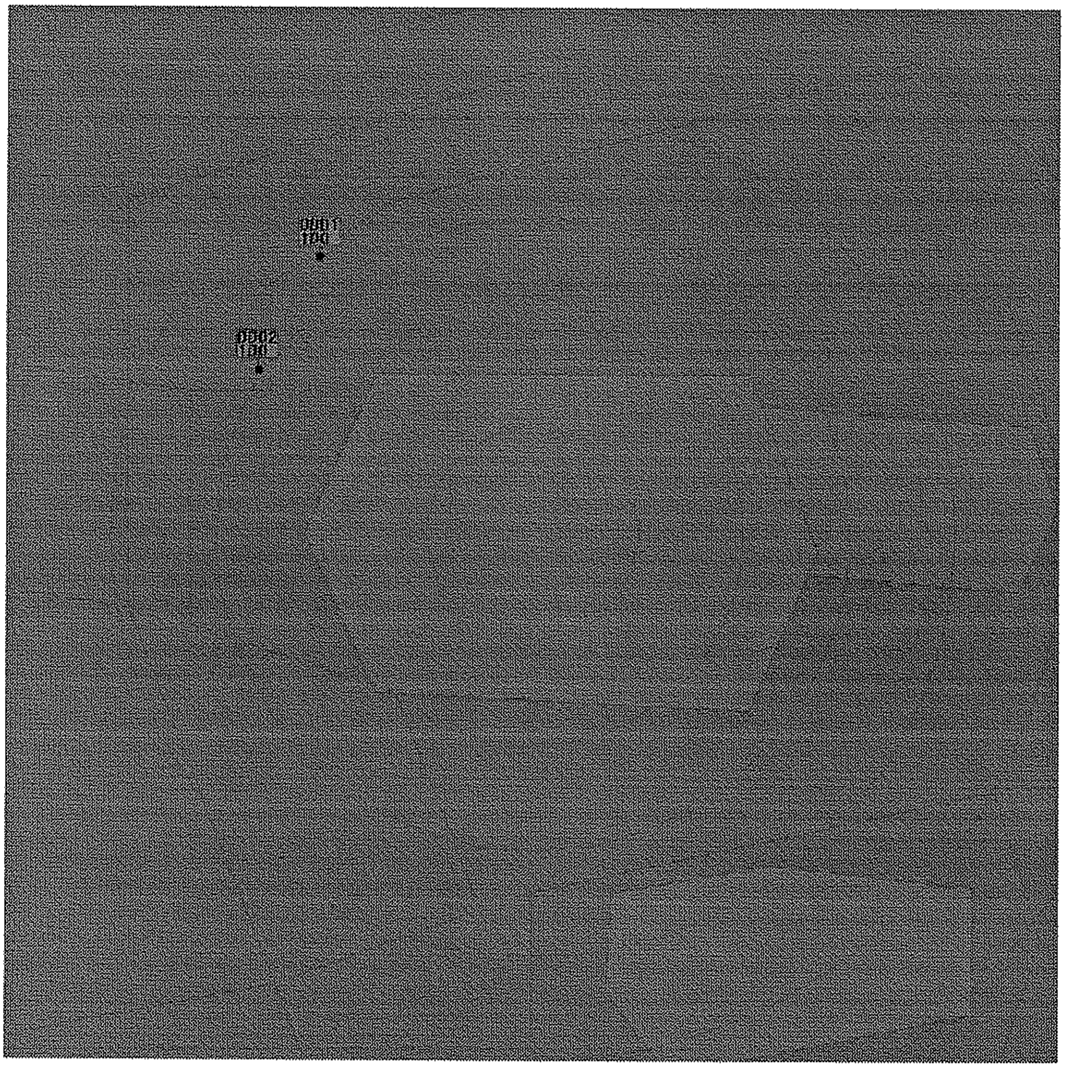


Figure 6.1: Main Experiment 2D Display Format

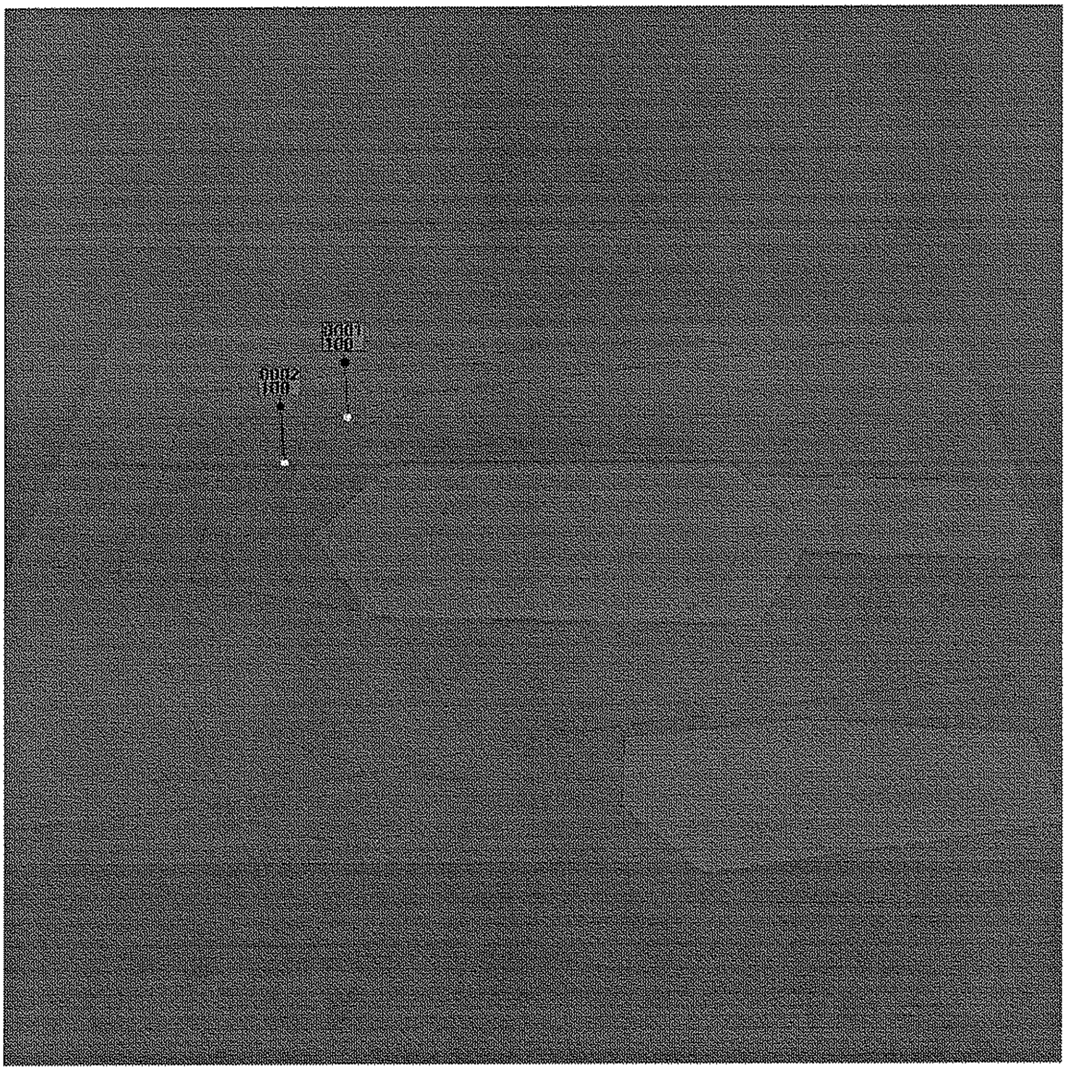


Figure 6.2: Main Experiment 3D Display Format

on the obscured datablock using the left mouse button (i.e. moved the tip of the arrow-shaped cursor to the obscured datablock and pressed and released the left mouse button) then the datablocks were 'swapped'. For 3D displays, however, datablocks had to be displayed such that near datablocks obscured far datablocks in order to avoid violating the occlusion depth cue; i.e. a datablock should not occlude another which is at less depth. For the 3D displays, therefore, a modified solution was adopted; clicking on a datablock which partially or completely obscured another caused it to become transparent, leaving just its frame and so revealing the datablock behind.

6.2.3 Changes to Immersive Display

The main problem in the preliminary display had been the excessive demand on system performance from the requirement to re-orientate and re-scale the datablocks continually in response to changes in head position and orientation. Clutter had also been found to be a major problem. This led to datablocks being omitted from the pilot study display, but it was important to try to find a solution to these problems which could be implemented in the main experiment display.

One approach to reduce both clutter and system load was to reduce the number of datablocks displayed at any one time, in which case the problem becomes one of how to select datablocks of interest. Two solutions to this problem were considered:

1. Manually select aircraft for display of datablocks.

2. Automatically display datablocks for aircraft lying within an area directly in front of the subject's head.

Solution (1) envisaged a visible narrow 'beam' coming out of the 'hand' and being intersected with aircraft in order to display their datablocks. This was suggested to some of the subjects in the pilot experiments, although it was not demonstrated; it was rejected as an idea since subjects 'don't want to spend their time continually having to select aircraft'.

Solution (2) was investigated for the main experiment. This is based on the hypothesis that the viewer will tend to move his/her head to look directly at objects of interest to fixate on them. Although the display is head-tracked but not eye-tracked, and only eye movement is required to foveate on any aircraft currently displayed on the HMD, it is speculated that for objects outside an area directly in front of the viewer, the viewer will tend to turn his/her head towards the object of interest to view it anyway. With this in mind, solution (2) envisaged only displaying datablocks for aircraft within a certain area directly in front of the viewer. It is very similar to solution (1), except that the conceptual 'beam' (invisible in this case) would be larger and head-slaved instead of hand-slaved. However, with solution (2), the hypothesis is that because it attempts to work with the way in which we view the world naturally, it would be less intrusive.

The implementation of solution (2) was a relatively complicated. The problem is to find quickly the objects contained in a given 'selection volume'. The choice of selection volume shape is important, both for consequences on computational load and for the effects presented to the viewer.

One implementation investigated using the dVS operating system's collision callback mechanism. Callbacks can be registered for collisions between the bounding boxes of objects. These bounding boxes are formed simply by taking the maxima and minima of the x, y and z coordinates of an object's vertices, and have all the usual problems of axis-aligned bounding boxes; the fit between the bounding box and the object's boundary can be poor, and for long, thin objects especially, the bounding box can be potentially a lot larger than the object. However, the bounding box mechanism does provide a rapid method for finding which objects are *potentially* in collision, and more detailed collision computation can then be carried out on this reduced set by an object boundary intersection test. The issue is then to implement this test efficiently. Unfortunately, a major flaw was found with this approach. The way in which dVS detects and handles collisions is very limited, and in the case of multiple simultaneous bounding box collisions, only *one* collision actually results in a callback to the application. Detecting the case where objects 'decollide' (are no longer in collision) was also very difficult.

It was therefore decided to abandon this approach and use a brute force method of periodically testing all aircraft objects against the selection volume. The ease of such an approach depends on the shape of the selection volume. If it is assumed that all polyhedra in the scene are closed and convex, then a simple intersection test is as follows: An object a intersects or lies within object b if any of a's vertices are on or within b's boundary area. (This does not cover the cases of edges only intersecting, or the case where the polyhedra are precisely the same shape and are perfectly aligned. However, since the aircraft polyhedra are small in comparison to the selection volume, the latter issue does not apply, and since the head position is constantly updating the former restriction should not pose a problem.) One of the easiest volumes to test for intersection against is a unit cube at the origin aligned with the world coordinate axes. Whilst this is obviously impractical as a selection volume as it is, a shape may be selected which can be transformed to this unit cube easily. The same transformation can then be applied to the object under test, so the test may be done trivially within this transformed space.

Two shapes of selection volume suggested themselves: A simple cuboid and a truncated pyramid. Transformation of a simple cuboid is trivial: Indeed, if the geometry for the viewing volume displayed by dVS was originally a unit cube to which transformations were subsequently applied, then the transformation matrix may be obtained directly from dVS. In the case of the truncated pyramid, the transformation matrix to a unit cube is well known as the camera transformation in computer graphics. For simplicity, a cuboid selection volume was chosen.

6.3 Experiment Design

6.3.1 Introduction

The experiment design extended that of the pilot study to include the stereoscopic 3D display and to introduce non-air traffic control subjects as well as ATCOs. The major hypotheses of the main study were therefore that subject performance at experimental tasks would depend (a) on the type of display and (b) on the type of subject.

The immersive 3D display was demonstrated to subjects but was not incorporated in the tasks.

6.3.2 Overall Experiment Design

Display type D has three levels: (1) 2D PPI, (2) pseudo-3D TTW and (3) stereo-3D TTW. Subject group G has two levels: (N) novices (non-ATCOs) and (X) experts (ATCOs). This gave a 3×2 study with 6 cells, shown in Table 6.1.

The tasks built on the lessons learned from the pilot experiment:

		Display D		
		1	2	3
Group	N	N1	N2	N3
G	X	X1	X2	X3

Table 6.1: Main Experiment Cells

1. Precision of Observation of Azimuth Angle and Relative Distance.
2. Information Extraction Speed.
   (a) Height information.
   (b) Horizontal separation.
3. Memory of a Scene.
4. Conflict Detection.
5. Interception.

These are described in the following sections.

As in the pilot study, subjects were given instruction sheets for each task (see Appendix B). Each task incorporated one or more demonstration/training scenarios which allowed the supervisor to explain the instruction sheet details, allowed the subjects to practice the task and allowed the supervisor to ensure that the task would be carried out properly.

At the start of the experiment, each subject was given an introduction sheet (§B.1) and required to complete a questionnaire (§B.2). They then performed each task in turn, for which measures were taken. After each task, subjects were required to complete a further questionnaire which asked for subjective difficulty ratings (on a scale of 1–7) and for any comments. The reason for taking subjective ratings was to try to ascertain whether subjects perceived any difficulty in performing certain tasks.

After carrying out the experiment, subjects were given a commercially-developed spatial ability test [SW88]. It was hypothesised that air traffic controllers might have higher than average spatial reasoning ability and this might explain any observed differences between the two subject groups. The spatial ability test showed 20 shapes presented as unfolded 'nets' with markings on the outside. There were four questions per shape, each question showing a solid. For each question, subjects had to indicate whether or not the solid could be made by folding the associated net. Subjects had 20 minutes to answer the 80 questions.

6.3.3 Task 1: Azimuth Angle and Relative Distance

Introduction

The aim of this task was to examine the influence of display type and subject type on the reading of azimuth angle and relative distance. As stated in §4.3.4, the ability to make sufficiently accurate judgments of azimuth angle and relative distance has been stated by air traffic controllers as being very important.

As related in §3.2.3, 3D displays have the problems of perceptual ambiguities in reading distances and biases in reading of azimuth and elevation angles, whereas in a plan-view display, horizontal distances and azimuth angles are presented without ambiguity. As one air traffic controller has related to the author:

6.1 *...the advantage of the 2D/mode-C concept [is] the accuracy. Use the vector and read the numbers...*

The pilot experiment indicated that with the display camera parameters used, a parallel projection gave smaller errors in reading azimuth angle than a perspective projection (although in the parallel projection, observed angle was found to be significantly dependent on B_x), and observed relative distance errors were not significantly different between the two displays. The main experiment seeks to compare the accuracy of angle and distance observations between different display types and subject groups.

Methodology and Hypotheses

The design of this task is based on that used in the pilot study (§5.2.1) and adopts the same nomenclature. In addition, the angle between the line joining T_1 and T_2 and a radial line from the centre of the display passing through point B is defined as ϕ.

The task is designed to test the influence of subject group G and display type D on observed azimuth angle y_A and observed relative distance y_D between the two aircraft. The task was carried out by presenting subjects with 20 stimuli in turn and asking them to note the angle and distance between the two aircraft in each stimulus. Since this task concentrated on exploring the influence of D and G on the dependent variables and was not designed to systematically explore the relationships between the dependent variables and the other variables, the values of y_A, y_D, B_x and B_y were generated randomly. The ranges of these variables are shown in Table 6.2.

$$
\begin{array}{rcccl}
-20\text{nm} & \leq & B_x & \leq & +20\text{nm} \\
-20\text{nm} & \leq & B_y & \leq & +20\text{nm} \\
-180^\circ & \leq & \lambda_A & \leq & +180^\circ \\
1\text{nm} & \leq & \lambda_D & \leq & 10\text{nm}
\end{array}
$$

Table 6.2: Main Task 1: Ranges of Randomly-Generated Variables

Regarding observations of azimuth angle, it is expected that the 2D PPI display will result in the greatest accuracy, with the stereo-3D TTW display the next most accurate and the pseudo-3D display the least accurate. A stereoscopic display is expected to yield higher accuracy of azimuth observations than a non-stereo display since Ellis suggests that stereo viewing removes one of the sources of viewpoint misestimation azimuth judgment bias (see §3.2.3). ATCOs are expected to read angles with greater accuracy on the 2D PPI than novices due to their training and experience, but whether there is any difference between the groups on the 3D displays (due to the ATCOs' training with 2D displays) is not known.

Regarding observations of relative distance, it is expected that the 2D PPI display will give higher accuracy than either of the 3D display formats. It is also expected that distance reading accuracy will depend on the angle between the line joining T_1 and T_2 and a radial from the centre of the display passing through point B (define this angle to be ϕ). This is because of the fact that range rings are provided to judge distance, and it is expected that it will be easier to judge distances if the line between T_1 and T_2 is perpendicular to the range rings rather than if it is tangential to them.

Procedure

Two training scenarios were shown to subjects. The first was displayed while the subject read the task instruction sheet (§B.3), so that he or she could refer to the display. (Being the first task instruction sheet, it described the format of the display in addition to the task itself.) Subjects were additionally given a response sheet (§B.4) on which to record the angle and distance between the aircraft in the training scenarios and each of the 20 experiment stimuli.

Having read the instruction sheet, the supervisor again talked through the points on the sheet, particularly explaining the features of the display (for example, the representation of aircraft and how to reveal obscured datablocks by clicking on them with the mouse) and covering any points of which subjects were unsure (for example, the use of the mouse, although most subjects were familiar with these). The subjects carried out the task for the two training scenarios under supervision, after which any further questions were answered. The subjects then carried out the task itself without further interference from the supervisor (although he remained close at hand in case of difficulty).

The same 20 stimuli were presented in a random order to each subject. Subjects could work at their own pace. After each stimulus, subjects were instructed to press the middle mouse button to proceed to the next stimulus. Having completed the 20 stimuli, subjects were presented with a questionnaire (§B.5) which asked them to record subjective accuracy and difficulty in reading distance and angle, any particular problems with reading distance or angle, and any further comments.

6.3.4 Task 2: Information Extraction

Introduction

As discussed in §5.5.2, it was found in the pilot experiment that air traffic controllers mentally 'filter' non-pertinent traffic, including sorting by height, and also tend to consider the horizontal and vertical dimensions separately. This task was aimed at investigating how speed of extracting horizontal and vertical information depends on display format and subject group.

Methodology and Hypotheses

The task was divided into two parts. The first part concerned the extraction of altitude information, the second part concerned the extraction of horizontal proximity information.

For each part of the task, subjects were presented with a succession of stimulus images each containing 10 aircraft, and for each stimulus were required to select two aircraft according to given criteria (highest and lowest, or the two aircraft with the smallest horizontal separation). For each stimulus, the time from presentation of the stimulus to selection of the two aircraft was recorded. Let time of altitude extraction for a stimulus be denoted by y_{AT} and time of horizontal proximity extraction be denoted by y_{HT}. The null hypotheses for this task are that y_{AT} and y_{HT} are both independent of display type D and subject group G.

Burnett and Barfield's study comparing a PPI display against a 3D perspective display using air traffic controllers (§2.6.2) included a task in which subjects were required to identify the callsigns of the highest and lowest aircraft in a scenario. Not surprisingly, they found that subjects performed the task faster using the perspective display than with the plan-view display for both moderate (7 aircraft) and heavy (17 aircraft) traffic densities. The current task had only one traffic density level but similar behaviour was expected. To determine height on a 2D display requires the subject to read all the datablocks, whereas in a 3D display the heights of aircraft are visualised directly in a pictorial image, and it is anticipated that the direct visualisation of height will be faster to interpret than reading all the datablocks.

For the horizontal proximity extraction sub-task, the 2D PPI display presents horizontal distance information unambiguously, whereas the 3D display is subject to ambiguity regarding components of distances along the display LOS (§3.2.3). It was hypothesised that subjects might be slower at extracting information from a 3D presentation than a 2D presentation for scenarios where the ambiguity in the 3D presentation might be problematic.

Procedure

The instruction sheet for this task is given in §B.6. It was decided to carry out the timing by computer. When each scenario was displayed, a timer was started. Subjects then had to select two aircraft from the scenario by clicking on their position symbols (the air position symbol only for the 3D display formats) with the left mouse button. Doing so caused an aircraft's position symbol to turn from black to red. Clicking on a red symbol caused it to turn back to black, which allowed subjects to change their minds in case they made a mistake. When satisfied with their selection of two aircraft, subjects were required to press the middle mouse button, which then caused the elapsed time since the start of the scenario to be recorded and the next scenario to be displayed. Subjects were asked to work as quickly as possible whilst maintaining accuracy.

At the start of the task, subjects were asked to read through the entire instruction sheet. The sub-tasks were then carried out separately, the altitude extraction sub-task first. In both sub-tasks, subjects were first given a briefing (verbal explanation of the instruction sheet) and then presented with three training scenarios (to get the subjects used to selecting aircraft with the mouse, and to counter any 'habituation' from the first sub-task affecting the second sub-task), followed by ten scenarios for which times were recorded. The same ten scenarios were used for each sub-task, but were presented in different (random) orders, both from sub-task to sub-task and from subject to subject.

After finishing the entire task, subjects were required to fill in a questionnaire (§B.7) which asked them to rate difficulty of each sub-task and to make comments on any particular problems with each sub-task and any other general comments.

6.3.5 Task 3: Memory Recall

Introduction

This extended the pilot study memory task (§5.2.2) which was designed to investigate the effect of display format on the memory of an air traffic scenario: Speed of assimilation into memory, speed and accuracy of recall, retention time in memory and the ability to chunk information.

Several lessons were learned from the pilot study, the main one being that the pilot scenario (containing 15 aircraft presented for 60 seconds) severely overloaded the subjects' memories. It also emerged that subjects were using their knowledge of traffic patterns in the Heathrow area to remember the positions and altitudes of aircraft (the pilot scenarios were drawn from real traffic samples). It was therefore decided to introduce two more variables into the main experiment task: Varying the number of aircraft to examine the effects of greater and greater load on memory, and showing random traffic patterns as well as ones selected from real traffic data.

Methodology

This task had four independent variables: D (display type) and G (subject group), A (number of aircraft) and R (scenario randomness). The overall experiment design consisted of independent groups of subjects, each being in one of the cells in Table 6.1. Each subject was presented with a number of scenarios which varied the number of aircraft and whether the traffic pattern was random or drawn from a real traffic sample.

For real traffic samples, callsign information was not available (since the radar data file from which the samples were taken only contained the mode-A codes from the RDP computer and not the associated callsigns) and so had to be made up, as in the random traffic samples. Callsigns consisted of two letters (representing an airline code) and three numbers (representing the airline's flight number). (At present real callsigns use three letters and three numbers, but the author was unaware of this at the time.) The numeric codes were chosen at random. Some airline codes were suggestive of airlines (e.g. BA for British Airways, PA for Pan American) whereas some were simply made up.

Number of aircraft (A) had four levels: 3, 5, 7 and 9 aircraft. It was expected that 3 aircraft would present a modest load on memory, 9 aircraft a very heavy load on memory. Scenarios were presented in order of increasing number of aircraft, and subjects were not informed about the number of aircraft. The training scenario contained three aircraft. Scenario randomness had two levels: Random pattern (R) or real traffic data (T).

Two stimulus scenarios were generated for each level of A, one random and one taken from a traffic sample. This gave eight stimuli, summarised in Table 6.3. Stimuli were presented as follows. First, the training stimulus, followed by the 3, 5, 7 and 9 aircraft stimuli. Within each pair of stimuli with the same number of aircraft, the order of presentation of random and traffic scenarios was randomised. For example, the presentation order for one subject might be (A, B, D, C, E, F, G, H) whereas another might be (B, A, C, D, E, F, H, G). Subjects were not informed in advance that the number of aircraft varied from stimulus to stimulus, nor that some were random and some were based on traffic samples.

Stimulus	Aircraft	Sample
train	3	R
A	3	R
B	3	T
C	5	R
D	5	T
E	7	R
F	7	T
G	9	R
H	9	T

Table 6.3: Task 3 Stimuli

Each scenario was presented to the subject for 90s. The subject was then required to reconstruct the scenario on a piece of paper which showed the displayed area in plan, complete with video map (§B.9), recalling aircraft position, height and callsign information. Time to recall y_{RT} and accuracy of position y_{PA}, altitude y_{HA} and callsign y_{CA} were all dependent variables.

Hypotheses

The null hypotheses were that recall time and accuracy of recalled position, altitude and callsign would be independent of scenario randomness, number of aircraft, display type or subject type.

As in the pilot study, it was postulated that the 3D display formats would result in position and approximate height being faster to memorise and recall. Display format was not expected to influence recall of callsign information.

It was expected that the numeric component of callsigns would be less easily remembered than the alphabetic component, especially where the alphabetic component bore some resemblance to an airline code (although it was not known whether subject type would influence memory of the alphabetic component).

For the expert subject group (G=X) it was anticipated that positions in real traffic samples would be more recalled more accurately than for random traffic patterns. No such behaviour was anticipated for the novice subject group (G=N).

Procedure

Subjects were asked to read the instruction sheet prior to the task (§B.3). The training scenario was then presented. Before each scenario, the response sheet was presented face down such that it had to be turned over left-to-right each time (this was explained to the subject and demonstrated). Each scenario started by showing a blank screen. When the subject was ready, he or she pressed the middle mouse button, which displayed the scenario for 90s. After this time, the screen was blanked, and subjects were asked to turn over the piece of paper, note down what they could remember and then to hand the paper to the supervisor when they had finished. The task was manually timed with a stopwatch from when the subject turned over the sheet to when he or she handed the sheet to the supervisor.

Subjects were instructed that they could rest for as long as they liked between scenarios, but were not told how many aircraft would be in the scenarios or that some scenarios were random.

Following the task, subjects were presented with a questionnaire (§B.10) which asked them to rate the overall difficulty of the task, and the difficulty of recalling position, height and callsign. (It was intended that the latter three should be relative to each other.) The questionnaire also asked subjects to describe how they carried out the task, and invited any further comments.

6.3.6 Task 4: Conflict Detection Task

Methodology and Hypotheses

This task was designed to extend the pilot conflict detection task to test the null hypotheses:

1. Conflict detection performance (in terms of time to detect a conflict and number of conflicts correctly detected) is independent of display type.

2. No more vertical conflict resolution instructions will be given using the 3D display types than in the 2D display.

3. Conflict detection performance will not depend on subject type.

Much of the rationale behind this task has already been described in §5.2.3 and so will not be repeated here. The main experiment task extended the pilot task to look

Stimulus	Conflict
A	No conflict.
B	Altitude bust (failure to level off). Heathrow outbound against holding stack.
C	Aircraft climbing westbound out of Gatwick against level southbound traffic. High clutter.
D	No conflict.
E	Two aircraft westbound in same airway, the east-most aircraft overtaking and converging with west-most aircraft.
F	Two converging head-on east-west, one climbing through the other's level.

Table 6.4: Main Experiment Conflict Detection Task Stimuli

at differences between expert and non-expert groups of subjects.

Subjects were presented with two training scenarios plus six scenarios in random order for which results were recorded. Each was a dynamic traffic scenario comprising 90s of animation (one frame every 6 seconds) and contained either a conflict or no conflict. Scenarios were based on real traffic data with a simulated aircraft introduced in scenarios containing a conflict. Performance measures were: Time to identify a conflict (if one existed), correct identification of conflicting aircraft and type of instruction issued.

The scenarios and types of conflicts they contained are summarised in Table 6.4 below.

Procedure

The instruction sheet (§B.11) had to explain the conflict detection task in sufficient detail for a novice to be able to carry out the task adequately. The supervisor also gave a verbal explanation and 'talked through' the first training scenario to ensure that the concepts were understood. Subjects then monitored the second scenario unassisted, but could ask questions of the supervisor.

For the task itself, the six scenarios were presented in random order. Subjects were not told whether or not each scenario contained a conflict. Before each scenario, a blank screen was presented. Subjects then pressed the middle mouse button to start the scenario, which also started a timer. The scenario was then displayed for a maximum of 90s. Time elapsed was displayed at the top right of the screen.

If subjects did not detect a conflict, they were required merely to let the scenario run to completion. If a conflict was detected, subjects were required to stop the scenario by pressing the middle mouse button, then to select the two conflicting aircraft by clicking with the left mouse button on their position symbols which caused selected aircraft to turn red, in the same manner as task 2. When the desired aircraft had been selected, subjects were required to press the middle mouse button again to confirm the selection. The time elapsed between the start of the scenario and its interruption by the subject and the two aircraft selected were recorded automatically. In addition, subjects were asked to give instructions to the conflicting aircraft to resolve the conflict, and these were noted by the supervisor.

Following the task, subjects were presented with a questionnaire (§B.12) in which they were asked to rate the difficulty of the task, to describe how they carried it out and were invited to make any further comments.

6.3.7 Task 5: Interception Task

Introduction

This task was not carried out in the pilot experiment but was a result of later considerations.

The above part-task tests are passive, in that they merely require subjects to read information from the display, and then compare the variation in speed or accuracy of the reading over the display types. Here an active synthetic task was introduced to see if the results of the part-task tests considered in isolation could apply to a task which was the combination of some of these elements.

Methodology and Hypotheses

The task devised involved controlling an interceptor (chaser) aircraft to catch a manœuvering target aircraft. The task was divided into two parts. The first part was a selection task: Given a static scenario with several aircraft, the subject was tasked with picking the aircraft which was the closest absolute distance from the origin of the world coordinate space (i.e. the point on the ground at the centre of the display range rings). The next part was the interception task: All targets apart from the selected one were removed and a chaser aircraft appeared on the ground at the middle of the screen and started moving. The subject's task was to guide the chaser to within a specified distance of the target, which manœuvered on a preprogrammed flight path.

The interceptor and target moved at constant speeds (the interceptor's speed was fixed at 1.5 times that of the target). To allow results to be compared if the subject failed to select the correct target, interception time and target manœuvres were both normalised. Interception time was normalised by dividing it by the initial absolute distance between target and chaser. The flight path of the target was specified by a sequence of triples of time, relative turn (azimuth) angle and pitch angle (flight path angle): e.g. at time t_1, turn left 40°and set pitch angle to $-5°$; at time t_2, turn right 60°and set pitch angle to 0°; etc. The target's manœuvres were normalised by setting its initial heading relative to its initial bearing from the centre of the display, and normalising the times in the flight path by the initial absolute distance from the centre of the display.

Subjects controlled the interceptor by two 'slider' bars, as detailed in the instruction sheet (§B.13). One slider controlled target azimuth angle, the other controlled target pitch angle.

The selection phase of the task is akin to the study performed by Bemis, Leeds and Wiener (§2.6.3, p. 51), comparing plan-view versus perspective format displays for threat detection and interceptor selection tasks. This study found that subjects using the perspective display made significantly fewer errors for all tasks, and that times taken to perform the threat detection and interceptor selection were lower. For this task, it was hypothesised that since this required visualisation of absolute distance instead of purely horizontal and vertical components, selection time would be lower and accuracy would be greater with the 3D display formats than with the 2D display formats, since distance between two points may be seen directly in a 3D display. (The 2D PPI display only shows horizontal distances in pictorial format.)

For the interception phase of the task, it was again hypothesised that interception time would be lower for the 3D display formats than for the 2D display.

Procedure

Having read the instruction sheet (§B.13), subjects were presented with two training scenarios for practice. After this, five scenarios were presented in the same order for each subject. The variables of time to select, aircraft selected and time to intercept were recorded automatically. After the task, subjects were presented with a questionnaire (§B.14) which required them to rate the difficulty of the task and invited any comments.

6.4 Immersive Display Demonstration

After the main experiment, expert subjects were given the opportunity to try using the immersive display on a voluntary basis. The aim was to introduce the concept of virtual reality, and to solicit opinions. Subjects were first required to read an introduction sheet (§B.15). They were then verbally briefed and the equipment was demonstrated. Subjects were then shown two scenarios, which they could spend as much time viewing as they liked. A possible application of immersive displays was then described to elicit comments as to feasibility/desirability. Finally, subjects were invited to write down any comments.

Two scenarios were shown: An architectural visualisation, in which subjects explored a room (a kitchen) and could interact with objects (e.g. pick up a teapot, open cupboards, dismantle a cooker), and an air traffic control radar visualisation with a scenario taken from real traffic data.

The architectural visualisation was shown to demonstrate the potential of immersive displays for simulating work areas. The expert subjects were from Heathrow tower, mostly working in the Visual Control Room (VCR). As part of the proposed future development of Heathrow airport, a new control tower is to be constructed, and controllers are being consulted in the design process. The architectural visualisation was introduced to give controllers an idea of the potential for simulating a VCR. (This idea was raised by an ATCO just before the main experiment was scheduled to start, so there was insufficient time to model a VCR; the kitchen demonstration was therefore substituted!)

The air traffic visualisation was to show controllers the potential for immersive radar displays, and to solicit opinions. This was a continuation of the demonstration carried out in the pilot study.

The proposed ATC application was an idea developed by the author and described verbally to subjects as a potential future application for immersive display equipment when the enabling technology becomes sufficiently mature. The use of a see-through lightweight HMD was suggested for use in the VCR, onto which aircraft and other symbology (ground vehicle positions, runway outlines etc.) derived from surface movement, primary and secondary radars and other sources, would be projected. This would allow controllers to see information superimposed on the outside world, and to stay 'head-up' in situations of reduced visibility (e.g. at night, or in fog or low cloud[1]), and to maintain contact if an aircraft being monitored visually disappears into cloud. Further, additional symbology such as projected extended runway centrelines and glideslopes could be superimposed so that controllers could see immediately whether an approaching aircraft was deviating from a safe trajectory without having to go 'heads down' to consult a supplementary display.

Ground controllers are used to carrying out their task visually the majority of the time, and such a display would enable them to do so under conditions of poor visibility. Designers of VCR simulators have related to the author that a difficulty experienced by trainees is correlating blips on a radar screen with aircraft in the outside world. This suggests that there may be a mental workload involved between relating an exocentric display such as radar to a situation where the controller is usually viewing the situation from a first-person (egocentric) perspective. Such a display would allow an egocentric perspective to be maintained and so might afford a lower workload than using a radar display. (Similarly, a workload is also experienced by a pilot transitioning from instruments to visual contact for landing, time being required to re-focus the eyes and to orientate oneself with respect to the outside world. A head-up display of flight symbology eliminates this.)

6.5 Chapter Summary

This chapter described the design and implementation of the main experiment study. First, the final design of the displays incorporating the lessons learned from the pilot study was described. Next, the design of the experiments was presented. Finally, a demonstration of the immersive display to be shown to ATCO subjects after they had carried out the experiment tasks was described.

[1]The new proposed Heathrow control tower is so tall that the VCR is expected occasionally to be in cloud!

Chapter 7
Results

"Go on, Mr. Pratt," says Mrs. Sampson. "Them ideas is so original and soothing. I think statistics are just as lovely as they can be."

O. Henry *The Handbook of Hymen*

7.1 Introduction

This chapter presents the results of the main experiment. For each task, a summary of the raw data, its statistical analysis, and a discussion are presented. The results of the immersive display demonstration are also presented.

Tabulated statistical data and analyses are given in Chapter C. Unless otherwise indicated, the analyses were carried out at a significance level of 0.05.

7.2 Demographic Data

A total of 49 subjects took part in the main experiment. Of these, the first 10 subjects were used to pilot the experiment, and debug procedures and instruction sheets. One of the remaining subjects was allocated to use the 3D stereoscopic display, but was found to be unable to see stereo (the subject reported the inability to see any difference in the displayed image with and without the stereo glasses, and further reported the inability to see random-dot or anaglyph stereograms); this subject's objective data therefore had to be removed from the study, although his subjective opinion data was incorporated. This left 38 subjects in the main study.

The main study pilot group was drawn largely from research students in the Computer Science Department at QMW, plus two students and a member of staff from other departments within the college. Pilot group subjects were paid £5 for participating in the experiment.

The main experimental factors were Group G and Display type D. The subjects were categorised into two groups; novices (non-ATCOs) (group G =N) and experts

(ATCOs) (group G =X). 16 subjects were novices, 22 were experts (with one additional ATCO's data being inadmissible due to lack of stereo vision, as explained above).

All ATCOs were employees of CAA NATS working in the control tower at London Heathrow airport, and were each paid one day's overtime for attending the experiment. However, only a few of these subjects were current radar operators, Heathrow tower operations consisting mostly of non-radar VCR tasks, with Thames Radar and London Special VFR radar being operated only by a few (although all radar operators at Heathrow also work in the VCR). All had been trained on radar, however, and all but one had previous radar operational experience.

Most novice subjects were student volunteers from QMW, although two members of staff also participated in the study. As a consequence, nearly all novice subjects were 'computer literate'. Novice subjects were each paid £5 for participating in the experiment.

Each subject was allocated to one of the three display types, and so into one of the six main experiment cells (Table 6.1). The number of subjects it was planned to allocate to each cell (based on projections of how many ATCOs might be expected to volunteer for the study), and the number of subjects actually in each cell, are shown in Table 7.1. The distribution is uneven due to "no shows" and reschedulings of subject attendance.

	Display			
Group	1	2	3	Total
N	6	6	6	18
X	6	6	6	18
Total	12	12	12	36

(a) Ideal

	Display			
Group	1	2	3	Total
N	6	5	5	16
X	7	9	6	22
Total	13	14	11	38

(b) Actual

Table 7.1: Subject Cell Allocation

Age distributions of the novice and expert subject groups are shown in Figure 7.1. The expert group subject's ages are fairly evenly distributed between age ranges 21–30, 31–40 and 41–50, whereas the novice group subject's ages are predominantly in the age range 21–30, reflecting the fact that most of this population were university students.

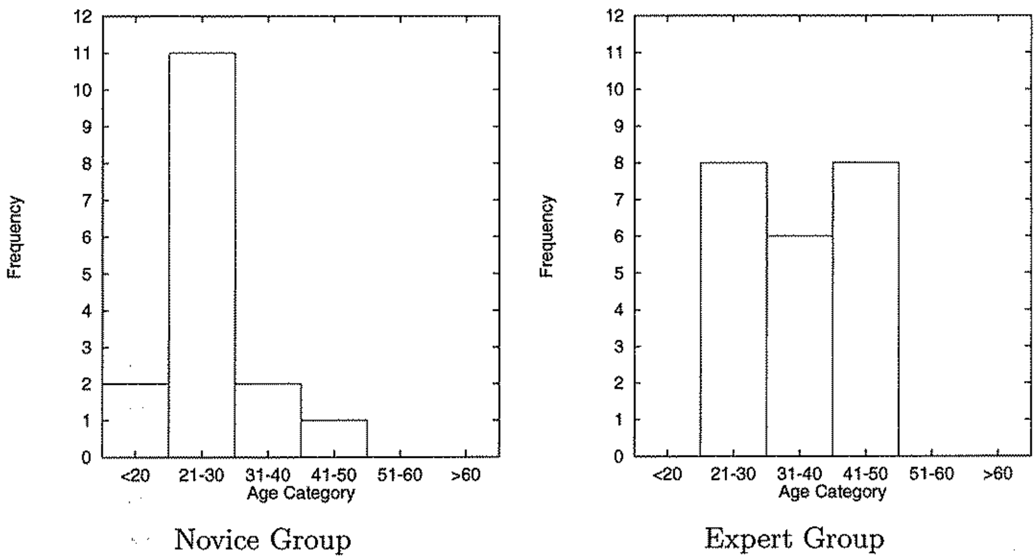


Figure 7.1: Subject Age Distribution

Breakdown by gender is shown in Table 7.2. Both expert and novice groups were predominantly male.

Group	Male	Female
N	14	2
X	18	4

Table 7.2: Subject Gender

7.3 Spatial Ability Test

A histogram of spatial ability scores for both groups is shown in Figure 7.2 (a score of 80 is the highest possible). Descriptive statistics (mean and standard deviation) are given in Table C.1. Spatial ability scores for two novice subjects were not available. A Student t-test was carried out to test the hypothesis that the means of the spatial ability scores of the two groups were different. This found no significant difference in spatial ability between the two groups (t = 0.672, $t_{34,0.05}$ = 2.032, $p = 0.506$).

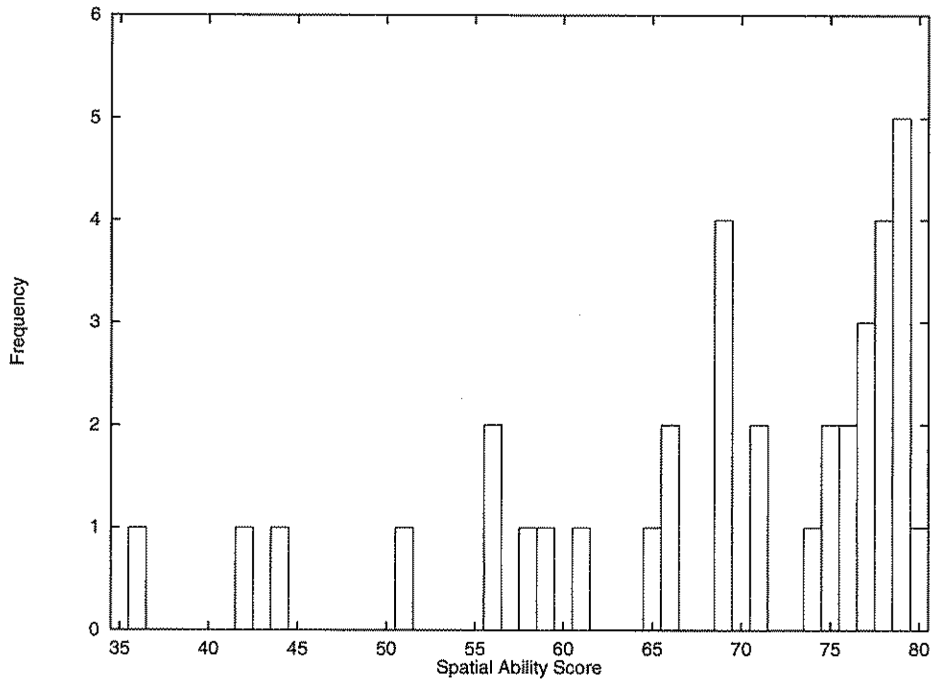


Figure 7.2: Spatial Ability Test Score Histogram: All Subjects

7.4 Task 1: Azimuth Angle and Relative Distance

7.4.1 Azimuth

Azimuth Observations

Subjects were instructed to read the azimuth of an object target T_2 from a reference target T_1. Occasionally, however, they wrote the bearing of T_1 from T_2 instead. This was a relatively rare occurrence (less than 10 observations out of the total of 760) and was corrected simply by adding (or subtracting) 180° to the observed azimuth where it was obvious that this had occurred.

The scatter plots of observed azimuth angle y_A versus stimulus azimuth angle λ_A with a superimposed line $y_A = \lambda_A$ for the novice and expert groups over all three display types are shown in Figure 7.3. The observations of the two groups appear to be very similar. An analysis of covariance incorporating the factors D and G and the independent variables B_x, B_y, and λ_A in a standard linear model resulted in the following models:

$$\mathrm{D}=1: \quad y_A = \quad -0.7 + 1.0\lambda_A - 0.1B_x + 0.2B_y \tag{7.1}$$

$$\mathrm{D}=2: \quad y_A = \quad -2.8 + 1.0\lambda_A + 0.2B_x - 0.3B_y \tag{7.2}$$

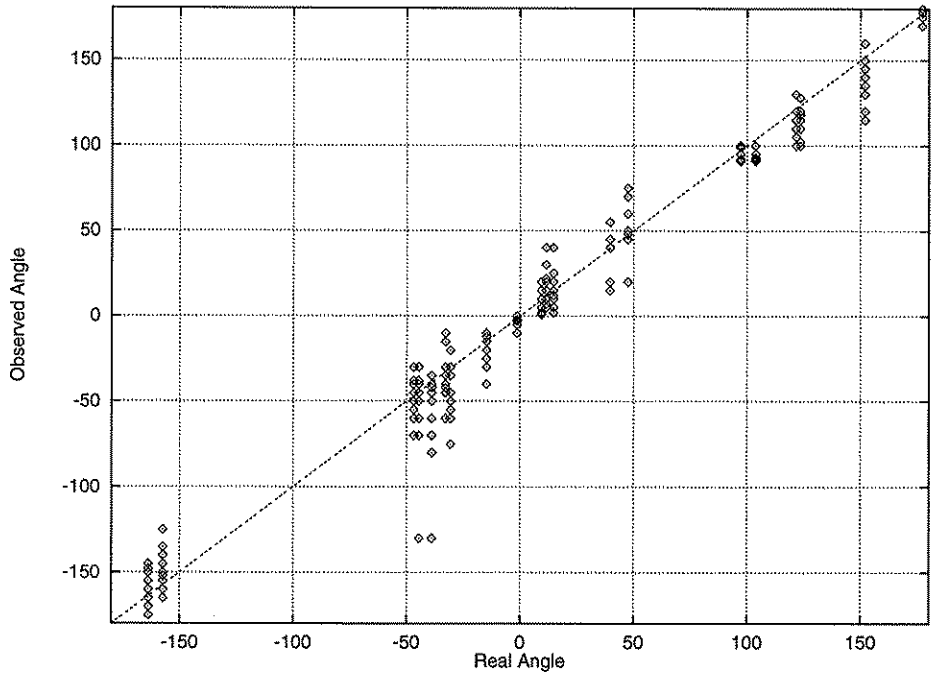
$$\mathrm{D}=3: \quad y_A = \quad -2.4 + 1.0\lambda_A + 0.1B_x - 0.2B_y \tag{7.3}$$

The model's estimates and standard errors for the significant variables are shown in Table C.2.

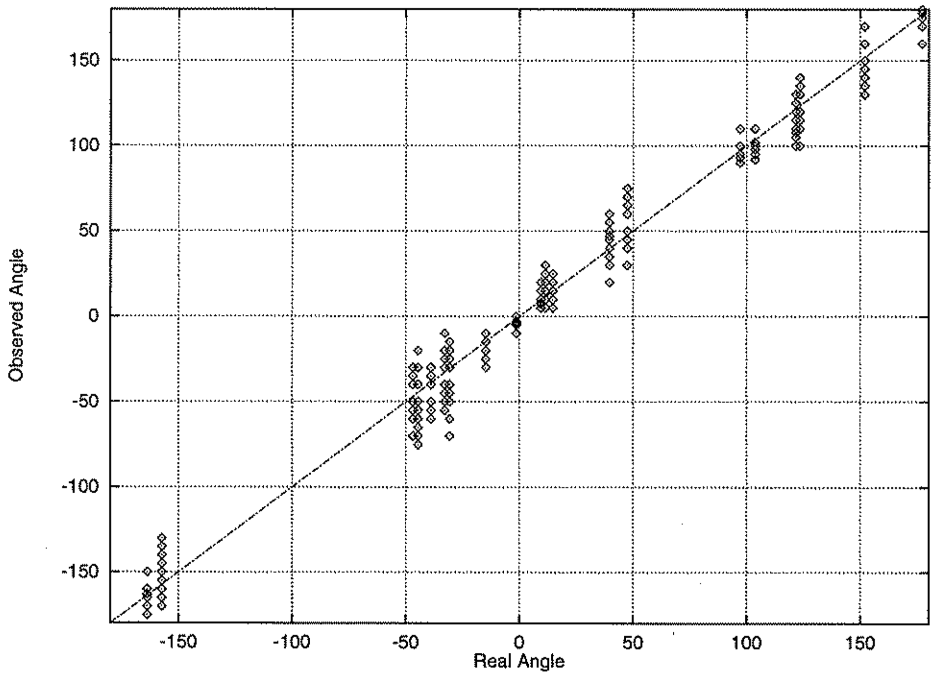
Subject group G was found not to have a significant influence on y_A: i.e. contrary to expectation, there were no significant differences between the novice and expert groups.

The constant terms in the above models take account of some of the random error and overall bias in reading azimuth between the displays, all other things being equal. For the 2D display ($D = 1$), the result is much as expected: There is no appreciable constant bias on azimuth. The 3D displays each show a negative bias (i.e. the subjects tended to underread the azimuth angles), the stereo-3D display having a slightly smaller bias than the pseudo-3D display.

B_x (the left/right position placement of targets on the display) was found to have a slight but significant influence on azimuth angle observations. This influence is greater for the pseudo-3D display than for the stereo-3D display.



Novice Group



Expert Group

Figure 7.3: y_A vs. λ_A Scatter Plots with $y_A = \lambda_A$ line superimposed

There was a significant but slight influence of B_y which was different in each of the three display types. This was at odds with the pilot study, which found no significant influence of B_y. Recall that B_y is the world y-coordinate of the point B which is the midpoint of a line joining the two targets (Figure 5.1). For the 2D PPI display ($D = 1$) B_y indicates whether the target pair is towards the top or bottom of the display window. In the case of the 3D display types, B_y is proportional to target depth (§5.2.1). The result for the PPI display (B_y coefficient of 0.2) was surprising since it indicates a tendency to overread the azimuth as the targets appear more towards the top of the screen, and a tendency to underread the azimuth as they appear more towards the bottom. For the 3D display types, the models suggest that increasing the z-depth of the target pairs imparts a negative bias on the reading of azimuth angle, the bias being greater for the pseudo-3D display than for the stereo-3D display. This is again at odds with the findings of the pilot study.

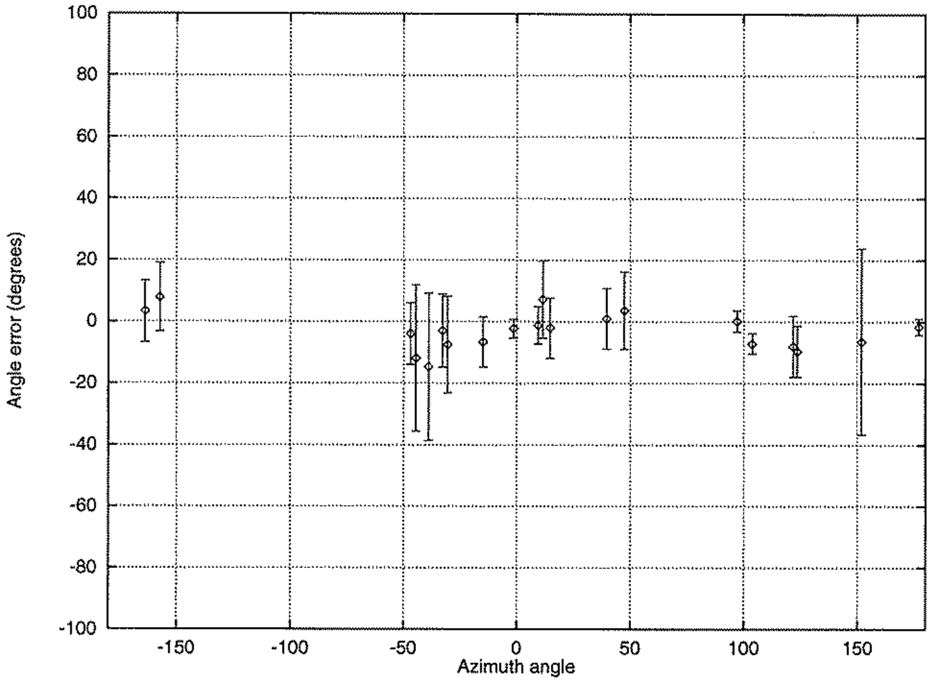
Define y_e to be the azimuth estimation error:

$$y_e = y_A - \lambda_A \tag{7.4}$$

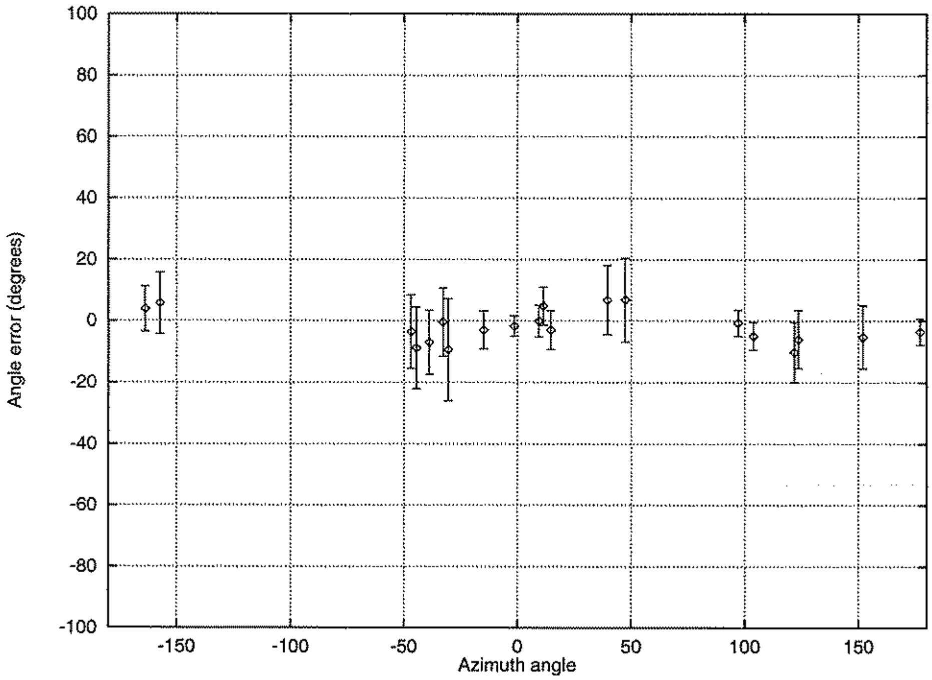
Mean and standard deviation plots of y_e against λ_A are shown in Figure 7.4. Descriptive statistics are given in Table C.3. The azimuth error plots are consistent with the Ellis and Hacisalihzade's reports that azimuth error is modulated with target azimuth angle, the errors being smallest close to implicit standards (0°, 90°, 180°, 270°) (§3.2.3, p. 65). There were insufficient azimuth stimuli in this experiment to show that this was actually the case; however, it is highly probable that similar error estimation patterns apply. In support of this, one subject commented that '*Judging the intermediate angles (i.e. not right angles) was more difficult.*' (Quote D.54).

Subjective Accuracy

Subjects were asked to estimate their margin of error by drawing a mark on a line which was marked off in steps of 5° from 0° to 45° (see §B.5, p. 272). The means and standard deviations of the subjective azimuth accuracy are shown in Figure 7.5. (Descriptive statistics are shown in Table C.4.) Cell N1 has two plots; **a** includes an outlying data point, **b** excludes it. The outlying point (an estimate of 40° of azimuth error) drastically affects the mean and variance of the sample ($\overline{X} = 13.7$,



Novice Group



Expert Group

Figure 7.4: Mean and Standard Deviation Plots of y_e vs. λ_A

$s = 13.5$ with the outlying point, compared to $\overline{X} = 8.4$, $s = 4.5$ without it). It is possible that this could have been due to misunderstanding the questionnaire. The questionnaire asked for the estimated margin of error and gave the example '*if you think you estimated angles to ±5°, please put a mark on the '5' position on the line*'. However, it is possible that the subject may have thought that the questionnaire was asking for the size of the angle range around the nominal value in which his answers might lie; a ±5° error margin means that the value might lie in a range spanning 10° centered on the nominal value, so in indicating 40°, he may have meant ±20°. It seems unlikely that one would have so low an opinion of his angle estimation ability as to rate this as ±40°; however, one cannot be certain. The data were analysed

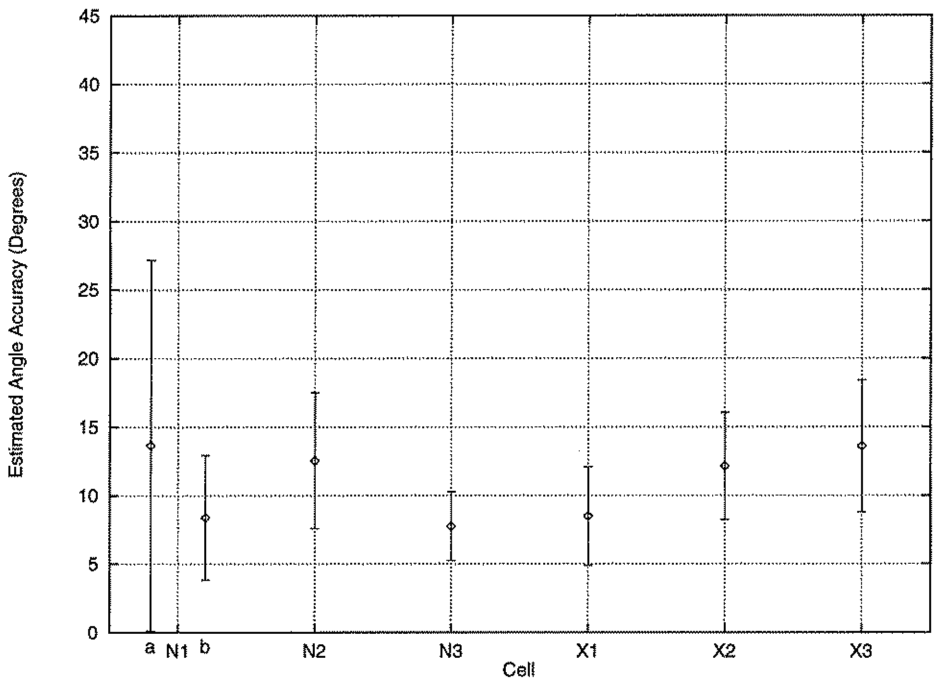


Figure 7.5: Subjective Azimuth Accuracy: Mean and Standard Deviation Plots

with and without the outlying point. Two-factor ANOVA analyses (Tables C.5 (all data) and C.6 (without outlying point)) showed no significant influence due to either display or group in either case.

It may also be interesting to see if there is a relationship between subjective azimuth reading accuracy y_{S_A} and actual azimuth reading accuracy. This was done by comparing the each subject's mean absolute azimuth error $\overline{\lambda_{S_A}}$ over all observations with his or her estimated azimuth accuracy. $\overline{\lambda_{S_A}}$ is defined as:

$$\overline{\lambda_{S_A}} = \frac{\sum_{i=1}^{n} |y_{e_i}|}{n} \tag{7.5}$$

where y_{e_i} is the azimuth error for observation i and n is the number of observations per subject (20). Linear regression analysis revealed no relationship between subjective accuracy and $\overline{\lambda_{S_A}}$, display or subject group. Thus, no evidence was found of a relationship between subjective accuracy and true accuracy.

Subjective Difficulty

Subjects were asked to rate the difficulty of reading azimuth angle on a scale from 1 (very easy) to 7 (very hard). Histograms are shown in Figure 7.6. Descriptive statistics are given in Table C.7.

Since these ratings are on a ordinal scale, rather than an interval or a rational scale, they are non-additive and thus strictly speaking cannot be analysed conventionally. However, in the spirit of data exploration, it may be useful to analyse the data, with the above caveat and the proviso that any evidence so gained may only be used in support of a hypotheses.

A two-factor ANOVA analysis (Table C.8) suggests no influence of display or group alone, but an influence by an interaction between the two. Single-factor ANOVA analysis within groups between display types and within display types between groups showed that for the 2D PPI display, there was a significant difference between experts and novices (cells N1, X1: $F = 10.756$, $F_{1,11} = 4.844$, $p < 0.01$), with novices reporting a higher level of difficulty overall than experts (novice median difficulty 5, expert median difficulty 2), but no other significant differences within groups or displays, which is consistent with the histograms.

7.4.2 Distance

Distance Observations

Scatter plots showing observed distance y_D versus stimulus distance λ_D for novice and expert subjects over all display types are shown in Figure 7.7. As with the pilot study, these plots show that variance of y_D is not constant with λ_D (variance

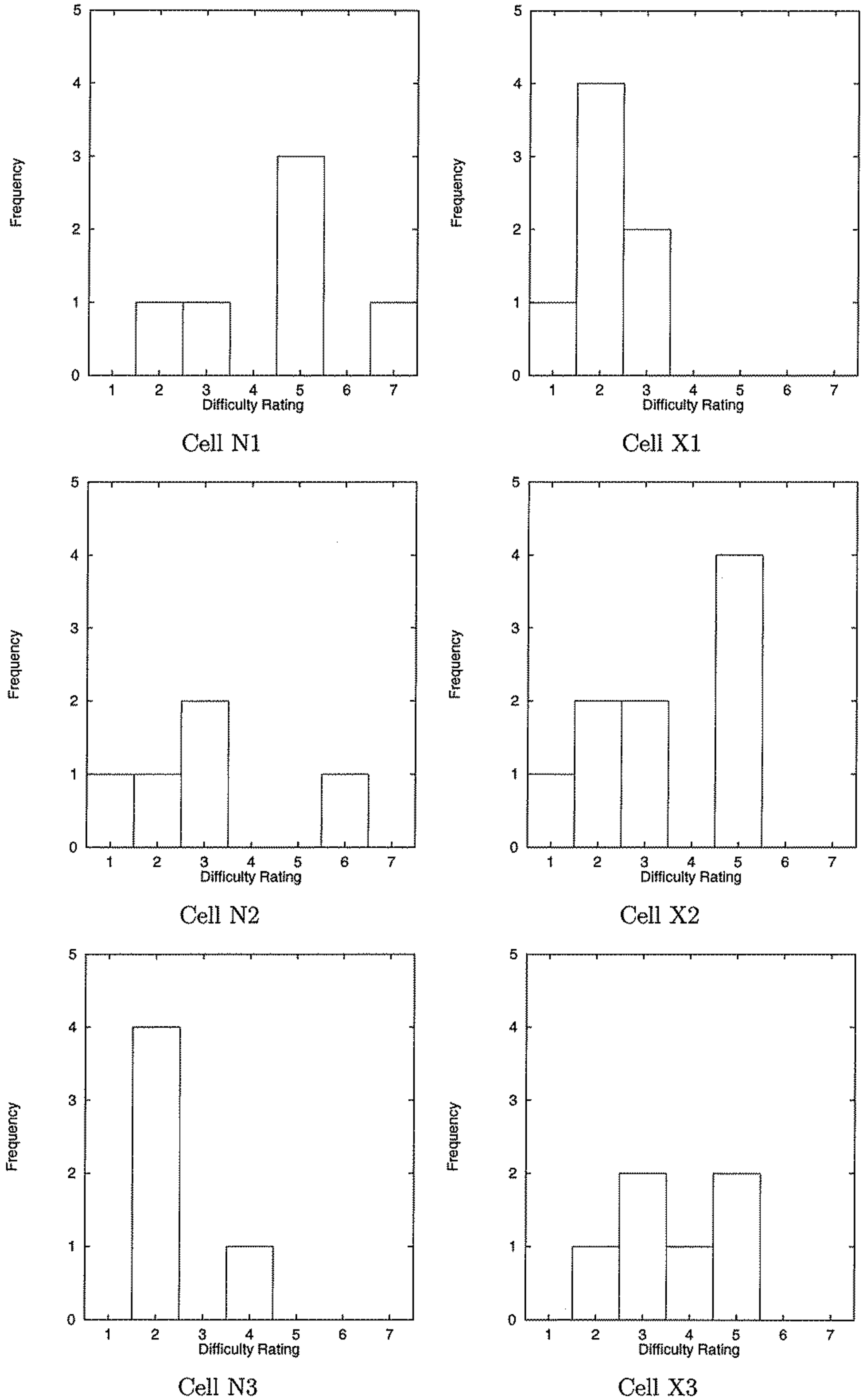


Figure 7.6: Subjective Azimuth Reading Difficulty

increases with true distance). This is supported by comments of four novice subjects, who noted that difficulty increased with increasing distance between the targets. The variation in variance invalidates a straightforward linear analysis (which assumes variance is constant). The pilot study examined $\ln y_D$ against λ_D, since this did not appear to show increasing variation with λ_D. However, this produced a curvilinear scatter plot with the main experiment data. Instead, $\ln y_D$ was plotted against $\ln \lambda_D$ (Figure 7.8), which gave a linear relationship. The following linear models were derived by analysis of covariance:

$$\text{D} = 1: \quad \ln y_D = -0.009 + 1.000 \ln \lambda_D - 0.001 B_x \tag{7.6}$$

$$\text{D} = 2: \quad \ln y_D = -0.053 + 1.000 \ln \lambda_D + 0.004 B_x \tag{7.7}$$

$$\text{D} = 3: \quad \ln y_D = -0.046 + 1.000 \ln \lambda_D + 0.001 B_x \tag{7.8}$$

which may be re-expressed as:

$$\text{D} = 1: \quad y_D = 0.99 \lambda_D e^{-0.001 B_x} \tag{7.9}$$

$$\text{D} = 2: \quad y_D = 0.95 \lambda_D e^{0.004 B_x} \tag{7.10}$$

$$\text{D} = 3: \quad y_D = 0.96 \lambda_D e^{0.001 B_x} \tag{7.11}$$

The model's estimates and standard errors for the significant variables are shown in Table C.9. Group was not found to be a significant factor, nor were ϕ and B_y.

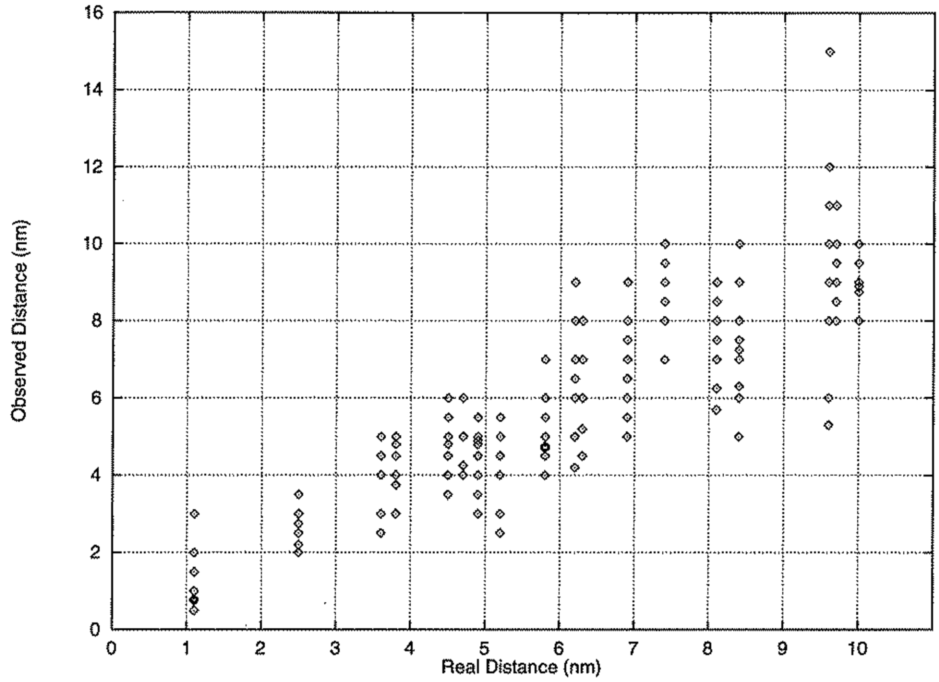
The error between observed and stimulus distance was also explored. Define the distance reading error y_{e_D} to be the difference between the observed distance y_D and the stimulus distance λ_D:

$$y_{e_D} = y_D - \lambda_D \tag{7.12}$$

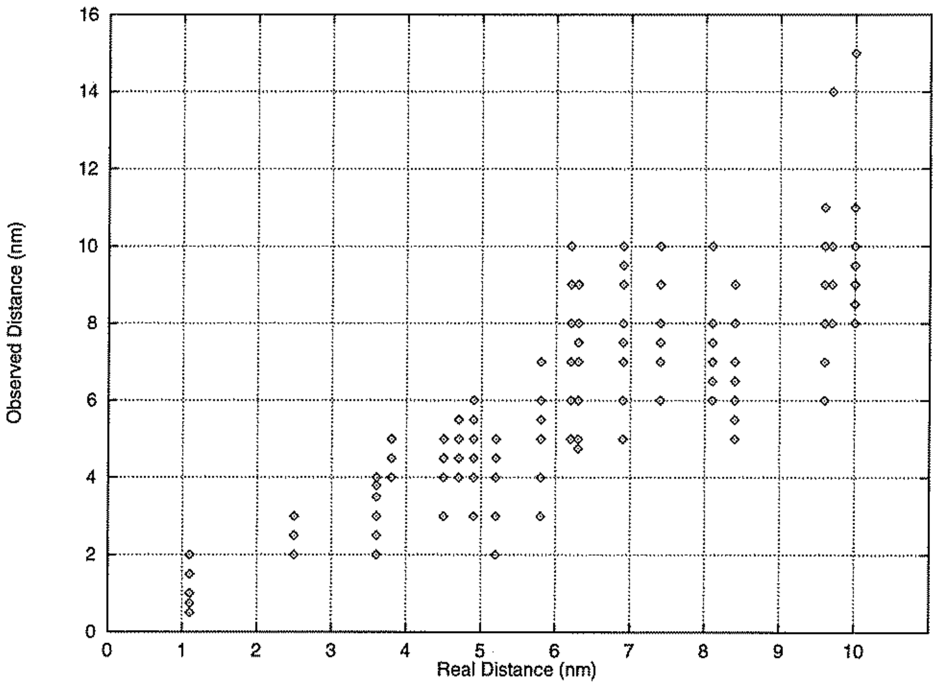
This is not itself easily amenable to analysis because y_{e_D} is itself it a function of λ_D (shown by the increasing variation of y_D with λ_D). Therefore, a percentage error term y_{E_D} is defined as the error as a percentage of the true distance:

$$\begin{aligned} y_{E_D} &= \frac{y_D - \lambda_D}{\lambda_D} \\ &= \frac{y_{e_D}}{\lambda_D} \end{aligned} \tag{7.13}$$

Distance percentage errors for each cell are shown in Figure 7.9. The variance is particularly large for the lowest stimulus (1.1 nm). A two-way ANOVA analysis



Novice Group



Expert Group

Figure 7.7: y_D vs λ_D Scatter Plots

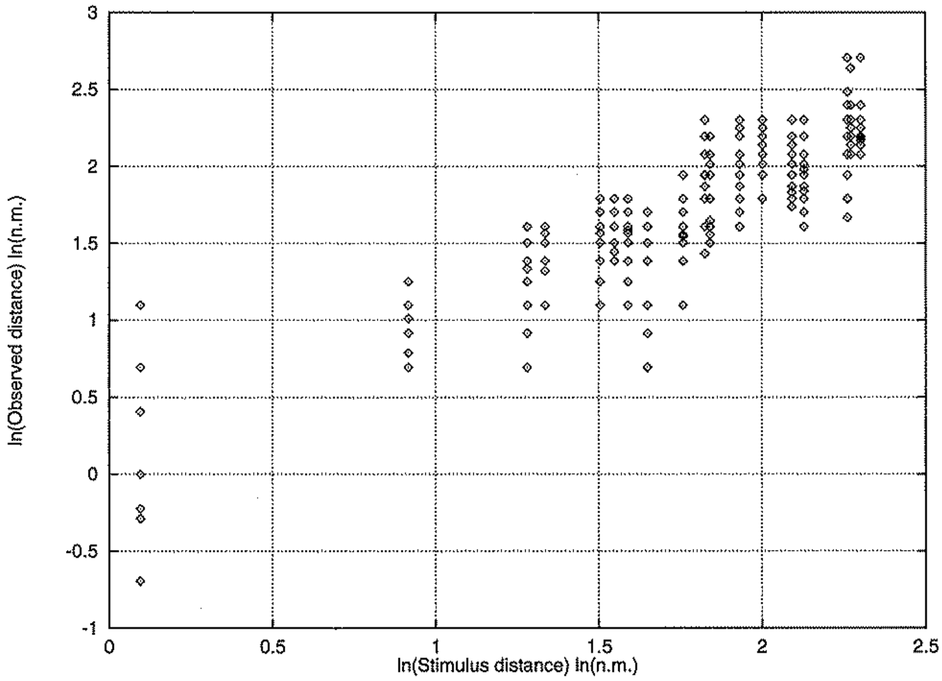


Figure 7.8: $\ln y_D$ vs. $\ln \lambda_D$ Scatter Plot: All Subjects

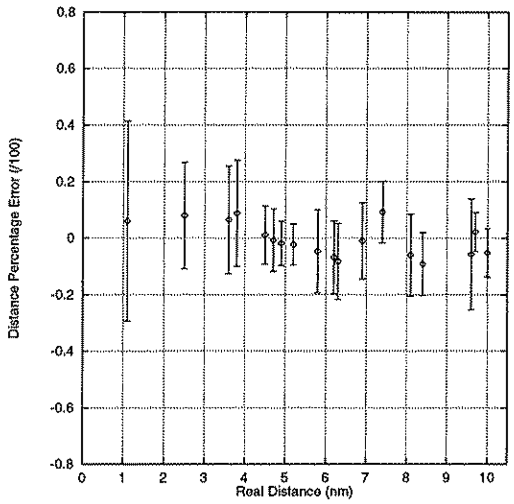
provided supporting evidence that display type had a significant effect ($F = 3.644$, $F_{2,\infty} = 3.01$ at $\alpha = 0.05$) and that group and the interaction between display and group did not have any significant influence.

Subjective Accuracy

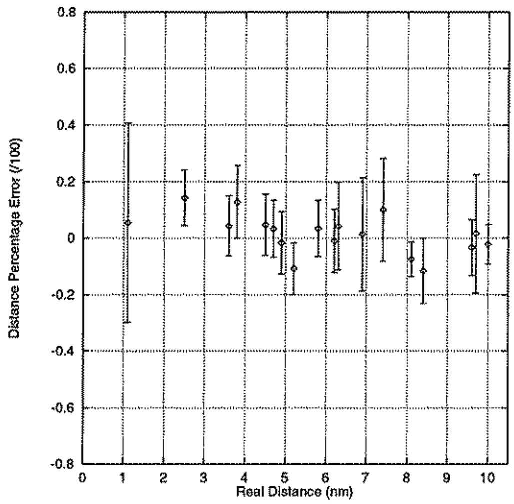
Subjects were asked to estimate their margin of error by drawing a mark on a line which was marked off in steps of 1 nm from 0–10 nm (see §B.5, p. 273). Subjective accuracy mean and standard deviations are shown in Figure 7.10, with descriptive statistics in Table C.10. Analysis by two-factor ANOVA (Table C.11) found no significant influence on subjective distance reading accuracy of display, group or the interaction between them.

Consider subjective distance reading accuracy against true distance reading accuracy. The mean absolute distance reading error for a subject is defined as follows:

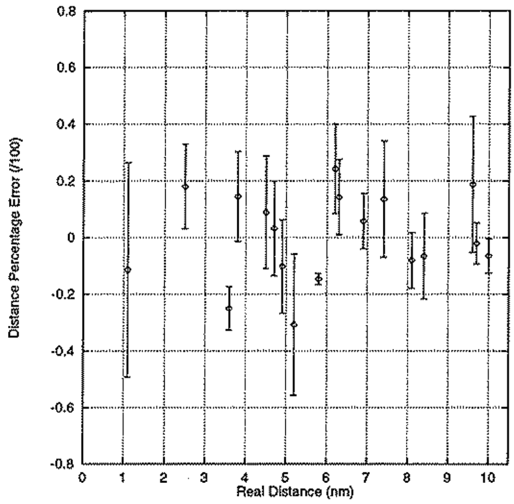
$$\overline{\lambda_{S_D}} = \frac{\sum_{i=1}^{n} \left| y_{e_{D_i}} \right|}{n} \qquad (7.14)$$



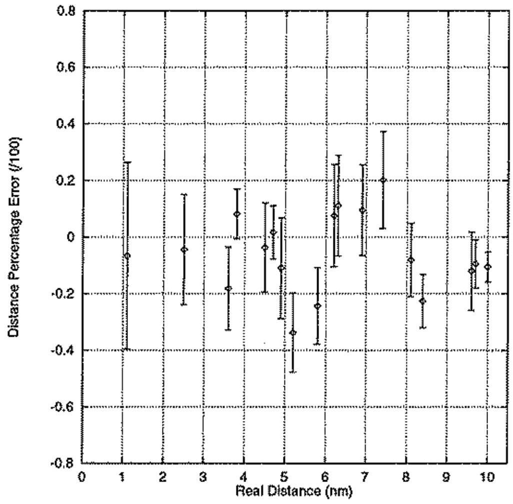
Cell N1



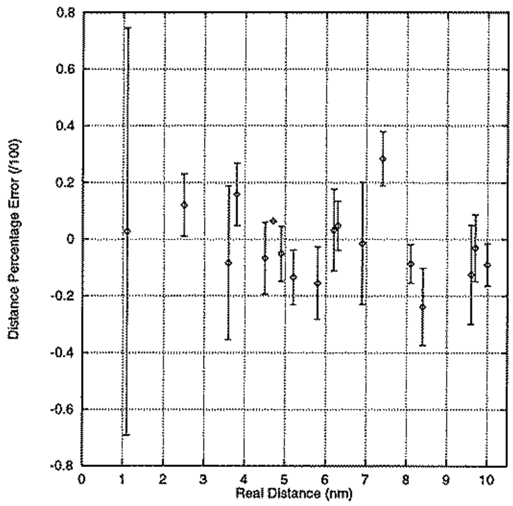
Cell X1



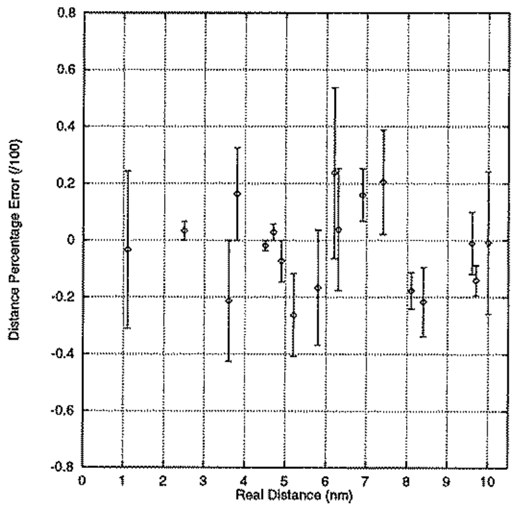
Cell N2



Cell X2



Cell N3



Cell X3

Figure 7.9: y_{E_D} vs. λ_D: all cells

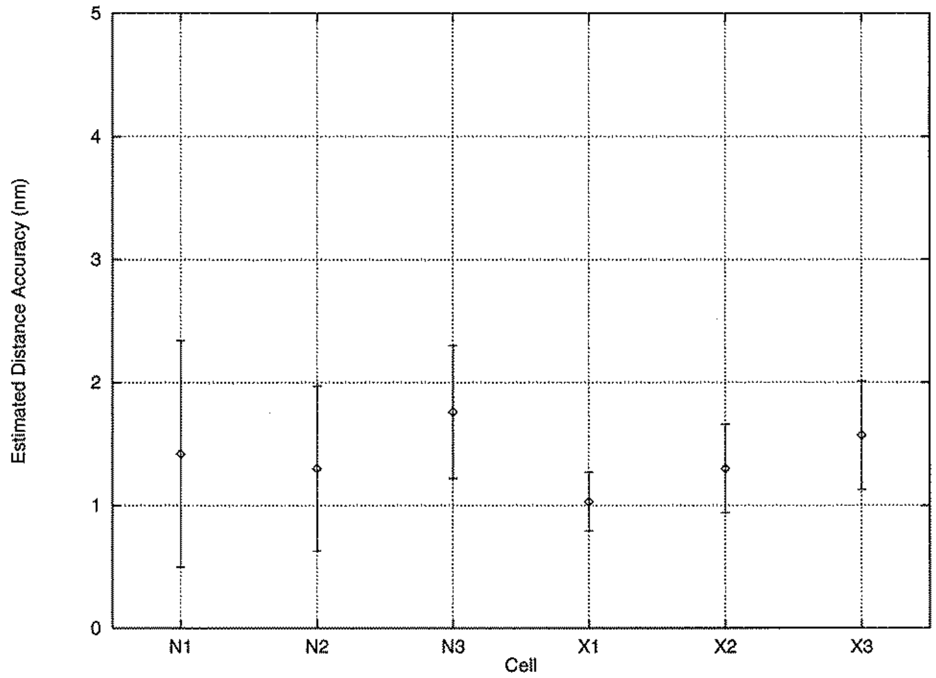


Figure 7.10: Subjective Distance Reading Accuracy: Mean and Standard Deviation Plots

where $y_{e_{D_i}}$ is the distance error for observation i and n is the number of observations per subject (20). Linear regression analysis of subjective accuracy versus $\overline{\lambda_{S_D}}$ failed to find a relationship between subjective distance reading accuracy and mean absolute distance error, display or subject group. Thus, no evidence was found of a relationship between subjective accuracy and true accuracy.

Subjective Difficulty

Subjects were asked to rate the difficulty of reading distance on a discrete scale from 1 (very easy) to 7 (very hard). The subjective difficulty is shown in the histograms in Figure 7.11. Descriptive statistics are shown in Table C.12. Analysis by two-factor ANOVA revealed no influence of display type, group or the interaction between the two on subjective distance reading difficulty.

7.4.3 Discussion

The Effect of Subject Group

Since radar operators are used to reading angles and distances from a display, it was speculated that this training would have led to experts being more accurate than novices. Contrary to expectation, however, subject group was not found to have a significant influence on the accuracy of reading azimuth angle or relative distance.

One explanation for the failure to find a difference may have been that the behaviour of the two groups was different. For both azimuth and distance observations, some subjects asked questions of the type "How accurate do you want it?" to which the reply "As accurate as you think you can make it." was given. From informal observation by the experimenter, the novices tended to take more time over their readings and attempted to achieve higher accuracy than the experts.

Because ATCOs seldom have to read azimuth angles with precision of greater than 5°, they may have just read angles to this sort of accuracy because they thought there was little point in trying for higher accuracy. Expert subjects had a higher proportion of observations that were multiples of 5° (428 out of 440 observations = 97%) than novices (283 out of 320 observations = 88%), which is consistent with this hypothesis. Similarly, novices may have attempted to read distances to a higher precision. Distance observations which were whole numbers were a higher

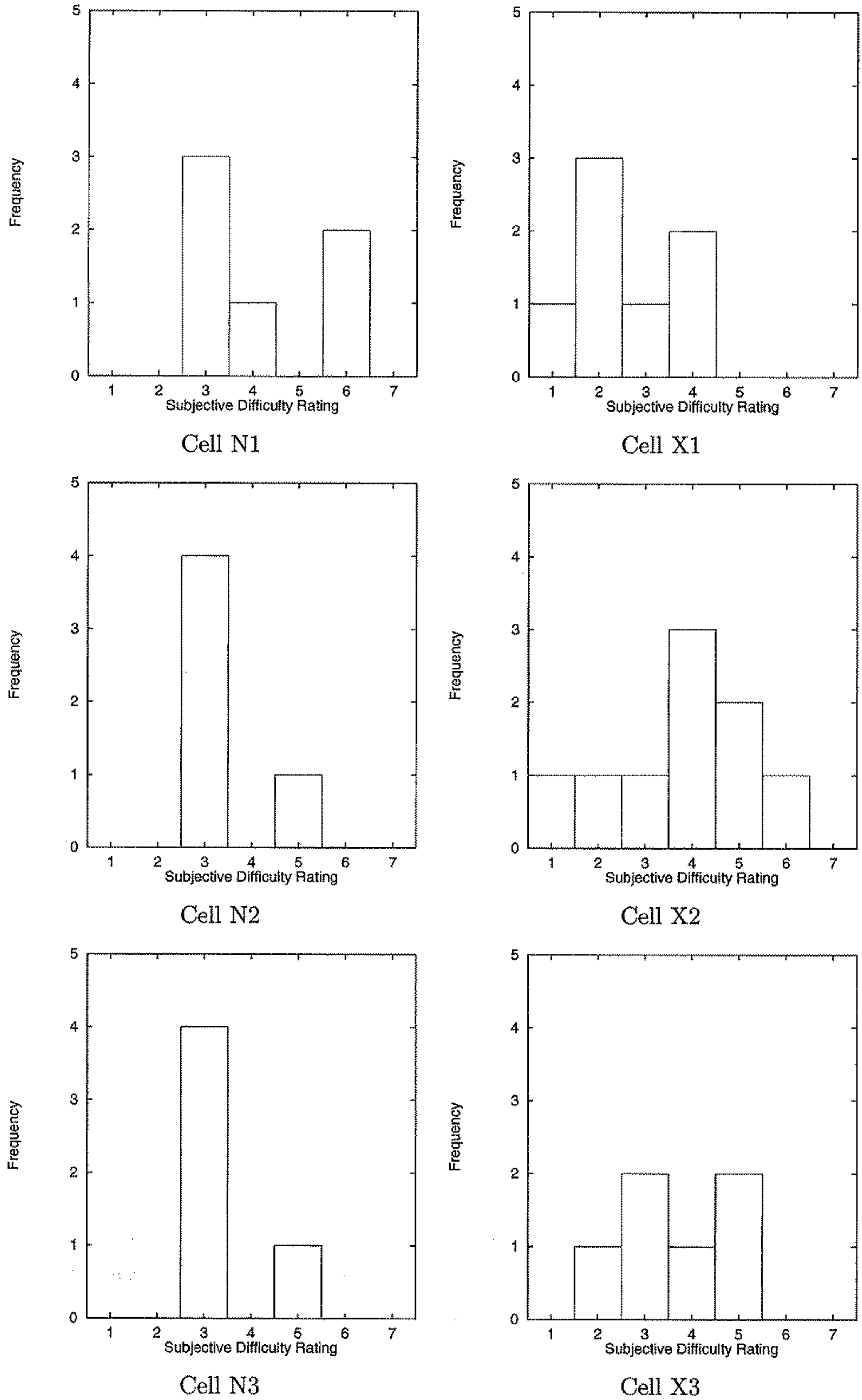


Figure 7.11: Subjective Distance Reading Difficulty

proportion of expert observations (397 observations out of 440 = 90%) than of novice observations (241 observations out of 320 = 75%), which supports this hypothesis.

Time taken by each subject over the task may be an indication of care taken, but this was not measured. This should be accounted for in future experiments.

Influence of Display Type

From the models in Equations 7.1–7.3 and the analysis of azimuth error, the results were much as expected: A 2D plan-view presentation affords the greatest accuracy for judgment of azimuth angle, followed by a stereo-3D display, with a pseudo-3D display being the least accurate. The smaller bias in interpreting the stereo-3D display is consistent with the assertion of Ellis, Tharp, Grunwald and Smith that stereo viewing will reduce errors in exocentric judgments of azimuth angle caused by the viewer's overestimation of slant angle (see §3.2.3, p. 65). Subjects tended to underread azimuth angle (more so for the pseudo-3D display than for the stereo-3D display); however, displays with different viewing parameters may exhibit different biases.

For observations of relative distance, from Equations 7.10–7.11, holding B_x constant, the 2D PPI display gives coefficient of λ_D closest to unity (and hence the greatest accuracy), followed by the stereo-3D and pseudo-3D displays in order of decreasing accuracy. Again, this is as anticipated.

Influence of Target Position on Azimuth Angle Observations

For the 2D display, the vertical placement of the target pairs on the screen (i.e. whether they were displayed towards the top or bottom of the screen) was found to have an unexpected significant influence. It is possible that this may have resulted from the display screen being perpendicular to the viewer's line of sight only at one location (eye level), so subjects had to cope with target pairs on the display not being perpendicular to the line of sight even if presented in plan view. This would be a function of distance of the eyepoint from the display surface and the eye level. (Unfortunately, eyepoint position was not tightly controlled and varied during the experiment as subjects were asked to adjust the chairs on which they were sitting to a comfortable height, and could move the chair slightly further or nearer to the screen as they saw fit, although the face of the display screen was kept vertical and its position on the desk in front of the viewer was constant.) One subject (an expert subject using the PPI display) alluded to such effects on the interpretation of relative distances in his comments after the task (Quote D.30 in Appendix D), and it is speculated that if such effects influenced interpretation of distance, they would also influence the interpretation of azimuth angle.

For the 3D display types, increasing the z-depth of the target pairs tended to impart a negative bias on the reading of azimuth angle, the bias being greater for the stereo-3D display than for the pseudo-3D display.

The left/right placement of the targets on the screen (B_x) was found to have a slight but significant influence on azimuth angle observations. This influence was found to vary between the display types.

Some of these findings were at odds with the results of the pilot experiment. The experiment was not designed to make a thorough exploration of the effects of target position on the observation of azimuth angle, but the results suggest that there may be significant effects which should be investigated in further studies.

Influence of Target Position on Distance Observations

There was no significant influence of ϕ (the angle between the line joining T_1 and T_2 and a radial line from the centre of the display passing through B), which is contrary to anticipated from subjective reports (seven subjects, 2 novices and 5 experts, commented that distances were easier to judge if the targets lay on a radius from the centre of the display, since range rings give best distance guidance along a radius and the least guidance where the line between the targets is tangential to a range ring). It is possible that although some subjects reported reading to be more *difficult* in cases where the targets lay on a tangent to a range ring, they may have taken more care in these cases such that accuracy was not significantly affected.

Unexpectedly, no significant effect of B_y (which is proportional to z-depth in the 3D displays) was evident, but there was a significant influence of B_x, which, broadly speaking, is negative if the targets are shown on the left hand side of the display, and positive if they are on the right hand side. The range of B_x in the stimuli being $-20 < B_x < 20$, this gives the following ranges for $e^{\gamma B_x}$ and for the overall coefficient of λ_D at $B_x = -20$ and $B_x = 20$:

Thus, for the PPI display, there is a slight tendency to overread distance in the

D	$e^{-20\gamma}$	$e^{20\gamma}$	λ_D coeff
1	1.020	0.980	1.01 – 0.97
2	0.923	1.083	0.88 – 1.03
3	0.980	1.020	0.94 – 0.98

Table 7.3: y_D Linear Model Coefficient Ranges

left side of the display and a slight tendency to underread distances at the right side of the display. For the pseudo-3D display, this tendency is reversed and slightly stronger. For the stereo-3D display, there is a tendency to underread distances on both sides of the display, with a greater tendency on the left than on the right.

Six subjects (all ATCOs—perhaps because novices simply accepted the distortion?) commented that because the 3D projection distorted the range ring circles into ellipses the scale differences between the display x and y (east and north) axes were confusing. These should account for some of the error of reading distances in different parts of the 3D display, but there should be influence of B_x as well as B_y on y_D.

Again, the experiment was not designed to explore the effects of target position on the observation of distance, but the results suggest that there may be significant effects which should be investigated in further experiments.

Subjective Difficulty

One expert subject commented *'it's a lot harder to see angles when viewing an elliptical display as opposed to a flat 2D screen'.* This seems reasonable given the unfamiliarity with 3D displays. Three subjects using the 3D display formats further reported that angles were more difficult to read in certain parts of the display (Quotes D.85, D.77 and D.78). It was also expected that the fact that the distances between the range rings are not subjectively uniform in the 3D display (due to their projections being ellipses) would make distance reading with the 3D displays more difficult than with the 2D display.

The azimuth angle reading subjective difficulty results suggested that for the 2D display, experts found the task easier than novices, but both groups found the task equally difficult for the 3D formats. This may be explained by the greater experience of ATCOs at reading angles than novices on the PPI display; however, the 3D displays were unfamiliar to all and so all subjects may have found them equally difficult for reading angles.

7.5 Task 2: Information Extraction

This task was divided into two parts. In the first part, subjects were required to select the highest and lowest aircraft from a number of scenarios. In the second part, subjects were required to select the aircraft with the closest horizontal proximity.

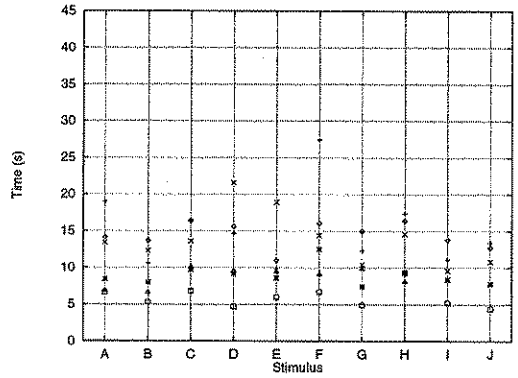
7.5.1 Overall Problems

A number of problems were found which applied to both parts of this task and also to later tasks.

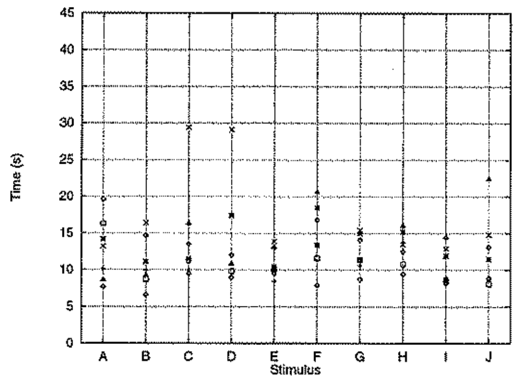
- Selection was using a mouse on the air plots on the 3D display. Despite the training, some subject forgot this and initially tried to select the ground plots instead.
- For the 3D display, the datablocks were more likely to overlap, especially using the stereo display. This made subjects more likely to have to click on datablocks than for the other displays.
- Some subjects reported the fact that the mouse sometimes kept sticking (despite being cleaned prior to each experiment) and this sometimes made it quite difficult to select an aircraft, even if they identified it in good time.
- The relatively small size of the aircraft symbols (and thus the area sensitive to mouse clicks) made selection difficult, especially with a sticking mouse.
- There were occasional problems with people who were familiar with single-button mice sometimes pressing the wrong button on the three-button mouse used in the experiment.

7.5.2 Altitude Extraction

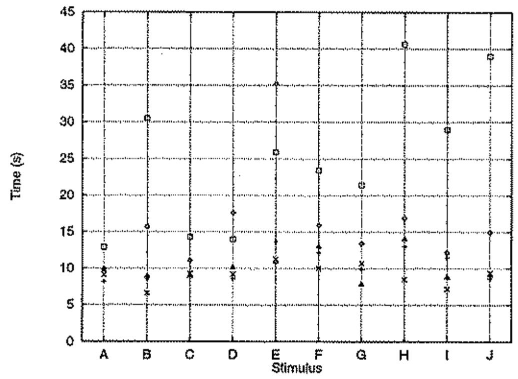
Altitude extraction times y_{AT} for each stimulus are shown in Figure 7.12. It can be seen that there are some data points with much greater times than the rest. These may have been due to the problems mentioned in §7.5.1 above. The technique of leverage plots [How92] showed that the points were below the threshold at which they could be considered as outliers, so analysis proceeded using all data.



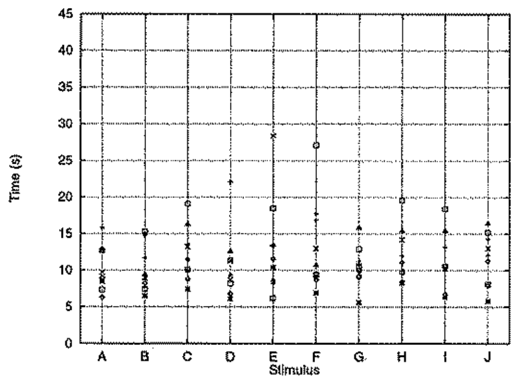
Cell N1



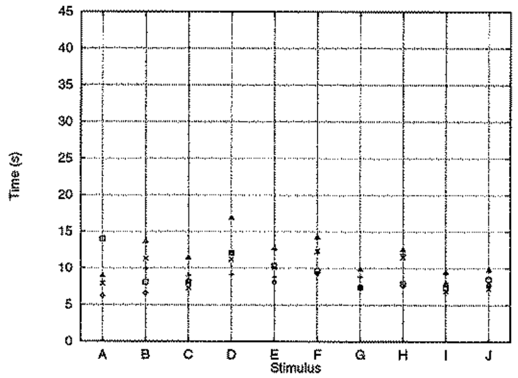
Cell X1



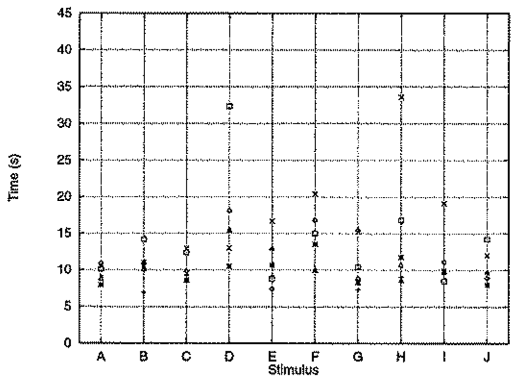
Cell N2



Cell X2



Cell N3



Cell X3

Each subject's data are represented by a different point style.

Figure 7.12: Altitude Extraction Time y_{AT}

Descriptive statistics of y_{AT} are shown in Table C.13. A two-factor ANOVA analysis was performed on the time data (Table C.14), and single-factor ANOVA was also employed to compare cells against each other individually (Table C.15). These analyses suggest the following regarding altitude extraction times:

1. Comparing display types in the novice group:
   (a) There is a significant difference between the PPI and pseudo-3D displays (faster altitude extraction using the PPI than the pseudo-3D format).
   (b) There is a significant difference between the PPI and stereo-3D display, with the stereo-3D display giving the lower mean extraction time.
   (c) There is a significant difference between the pseudo-3D and stereo-3D displays, with the stereo-3D display giving a lower mean extraction time.
2. There are no significant differences between display types in the expert group.
3. There is a borderline significant difference ($p = 0.052$) between novice and expert groups using the 2D PPI display, with a lower mean time for novices than for experts.
4. There is no significant difference between novice and expert groups using the pseudo-3D display.
5. There is a significant difference between novice and expert groups using the stereo-3D display, with novices being significantly faster than experts (for novices, $\overline{X} = 9.6$s, for experts $\overline{X} = 12.0$s).

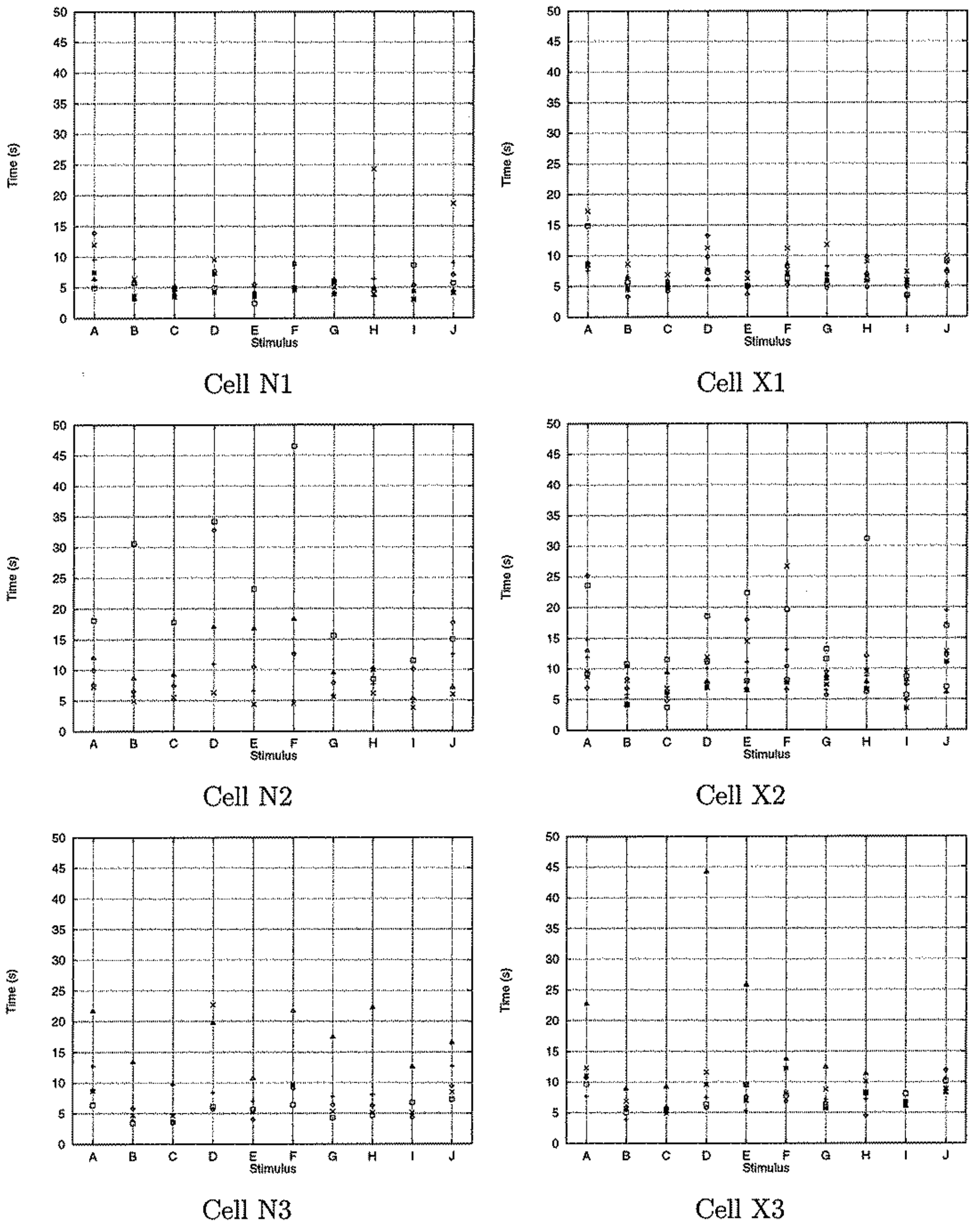
Subjects were asked to rate the difficulty of the altitude extraction task on a discrete scale of 1–7. Descriptive statistics are shown in Table C.16. Analysis by two-factor ANOVA failed to find any significant influence of group or display on subjective altitude extraction difficulty. However, some subjects reported in their comments that the 3D presentation made the task easier than by using datablocks alone.

7.5.3 Horizontal Proximity Extraction

Plots of horizontal proximity extraction times y_{HT} versus stimulus are shown in Figure 7.13. As with the altitude extraction task, it can be seen that there are a few points with much greater time than the rest, possibly due to the problems described in §7.5.1, but leverage plots again revealed that it was not safe to discard such points as outliers. Descriptive statistics of horizontal proximity extraction time are shown in Table C.17. A two-factor ANOVA analysis (Table C.18) suggests that horizontal proximity extraction time is significantly smaller for the 2D PPI than the stereo-3D display, and that the stereo-3D display gives significantly lower extraction time than the pseudo-3D display. Subject group was not found to be significant.

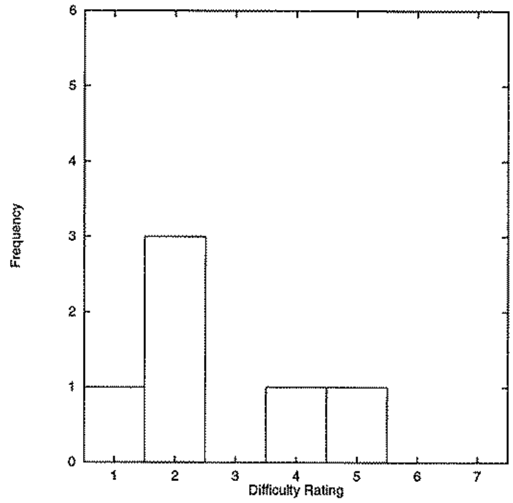
Subjects were asked to rate the difficulty of the proximity extraction task on a discrete scale of 1–7. Histograms of the subjective task difficulty are shown in Figure 7.14. Descriptive statistics are shown in Table C.19. A two-factor ANOVA analysis (Table C.20) suggested strong influences of the display and of the interaction between display and group. These were investigated further using single-factor ANOVA to compare individual cells or groups of cells (Table C.21). Regarding horizontal proximity extraction subjective difficulty, these analyses suggest that:

1. There are no significant differences within the novice group.
2. There are significant differences within the expert group:
   (a) The mean and median estimated difficulties are lower for the 2D PPI display (median value of 1.0) than for either of the 3D display formats (both with medians of 4.0).
   (b) There is no significant difference between the 3D display formats.
3. Using the PPI display, there is a significant difference between novice and expert groups, the experts' estimated difficulty having a lower mean and median than the novice group (expert median difficulty 1.0, novice median difficulty 2.0).
4. Using the pseudo-3D display, there is a significant difference between novice and expert groups, with the novice estimated difficulty having a lower mean and median than the expert group (novice median difficulty 2.0, expert median difficulty 4.0).
5. Using the stereo-3D display, there are no significant differences between the novice and expert groups.

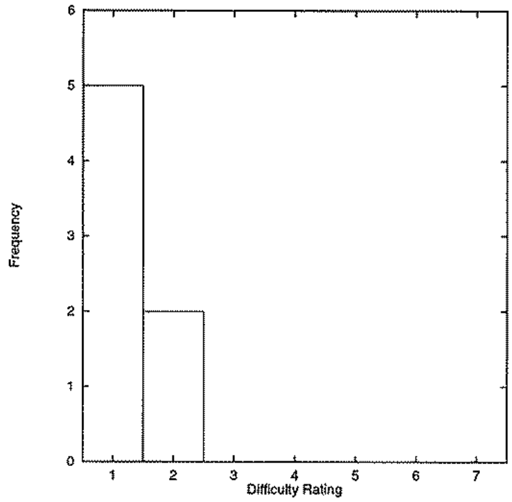


Each subject's data are represented by a different point style.

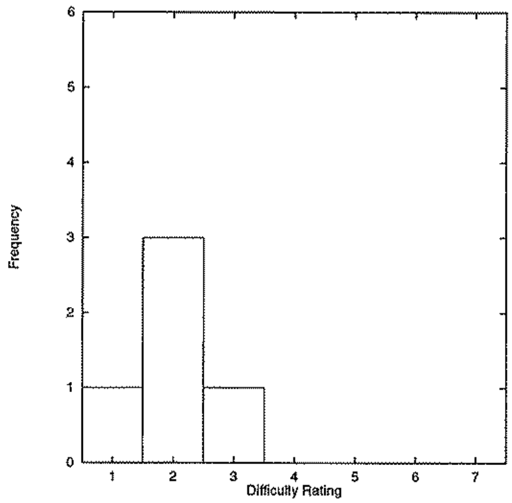
Figure 7.13: Horizontal Proximity Extraction Time y_{HT}



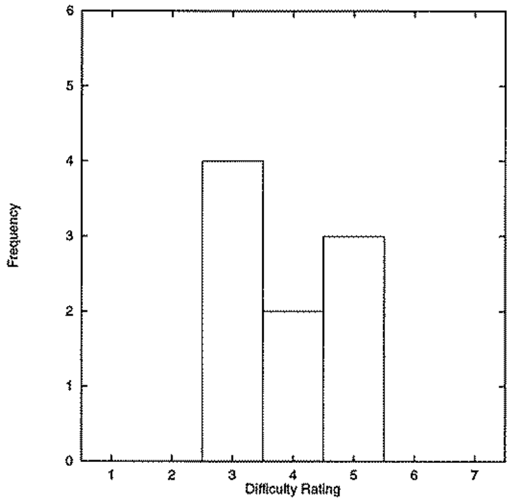
Cell N1



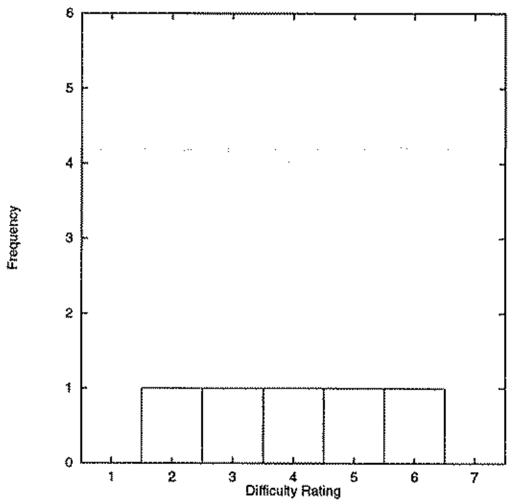
Cell X1



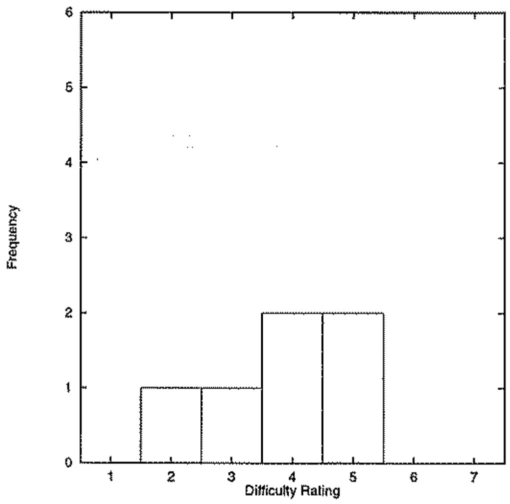
Cell N2



Cell X2



Cell N3



Cell X3

Figure 7.14: Estimated Horizontal Proximity Extraction Difficulty

7.5.4 Discussion

Altitude Extraction

Subject comments (§E.1) were used to gain insight into how subjects were performing the altitude extraction task. These suggest that subjects using the 3D display formats used the length of drop lines to narrow down the search, then used the datablocks for confirmation and accuracy. However, this was not so useful where drop lines were of similar length. Some subjects also reported problems when comparing drop lines of similar length which were displayed far apart (e.g. Quotes E.39 and E.44). A number of subjects reported that the drop lines made the task easier than with a 2D display.

Contrary to expectation, for the expert group, no significant differences in altitude extraction times were found between the different display types. Some expert subjects reported that they tended to use the mode-C readout in the datablocks instead of the length of the drop lines in the 3D presentations; reasons given were force of habit (Quote E.24 in Appendix E) and doubts about accuracy using the drop lines (Quotes E.29 and E.32). This bias toward the use of the datablocks could explain the lack of significant difference of performance between display types.

For the novice group, altitude extraction was found to be fastest for the stereo-3D display ($\overline{X} = 9.6$s); next fastest for the 2D PPI display ($\overline{X} = 11.2$s); and slowest for the pseudo-3D display ($\overline{X} = 14.2$s). This was unexpected, and may be due to the problems with position ambiguity and clutter found in the pilot 3D display (§5.5.3), with inadequate depth cues possibly making the pseudo-3D display difficult to interpret. It is postulated that with adequate depth cues (in this case, with stereopsis), a 3D display of aircraft position affords faster performance than a 2D display for determination of height due to the pictorial visualisation of the vertical dimension. However, if depth cues are not adequate, then the potential benefits of the 3D presentation may be negated such that performance is worse than with a 2D presentation.

It is speculated that differences in performance between novices and experts (novices being significantly faster than experts for the PPI and stereo-3D displays) could be due to either novices being more 'computer literate' and so more experienced in using mice, or due to behavioural differences in using datablocks and drop-lines to extract the information from the display. Subjects were instructed to work as quickly and accurately as possible, but expert subjects may have been more concerned with accuracy (which would tend to place an emphasis on datablock information) rather than speed (which would tend to emphasise the use of the drop-lines in the 3D display formats).

Horizontal Proximity Extraction

The fastest proximity extraction times occurred for the 2D PPI display, next fastest for the stereo-3D display and slowest for the pseudo-3D display. Subject group was not significant.

Histograms of errors (incorrect selections) are shown in Figure 7.15. It can be seen that the majority of incorrect selections were in certain stimuli. Moreover, some of these stimuli gave incorrect selections in nearly all cells (stimuli A and J) whereas some gave incorrect selections only in the 3D display cells (stimuli D and F). Stimuli A, D, F and J all contained more than one pair of aircraft for which horizontal separations were very similar, but for stimuli D and F, one of the pair with similar separations lay along a N–S direction (i.e. coplanar with the line of sight) and the other pair lay along an E–W direction (i.e. orthogonal with the plane of the line of sight). The fact that stimuli D and F gave incorrect selections only in the 3D display formats is therefore probably due to the ambiguity of determining position along the line of sight in a 3D display (Quote E.74). (Recall that in the plan view display, horizontal separations are displayed unambiguously.)

Some subjects commented that they found the elliptical range rings in the 3D projections difficult; this may have made it slower to use. The evidence of difficulty ratings supports the suggestion that experts found the PPI display less difficult to use than the 3D formats. There may have been an additional search time for the 3D displays (cells 2 and 3) since the instructions stipulated that the subject had to select the aircraft *air* plots, but closest horizontal proximity must be determined on the *ground* plots, which must then be followed up the drop line to the air plots.

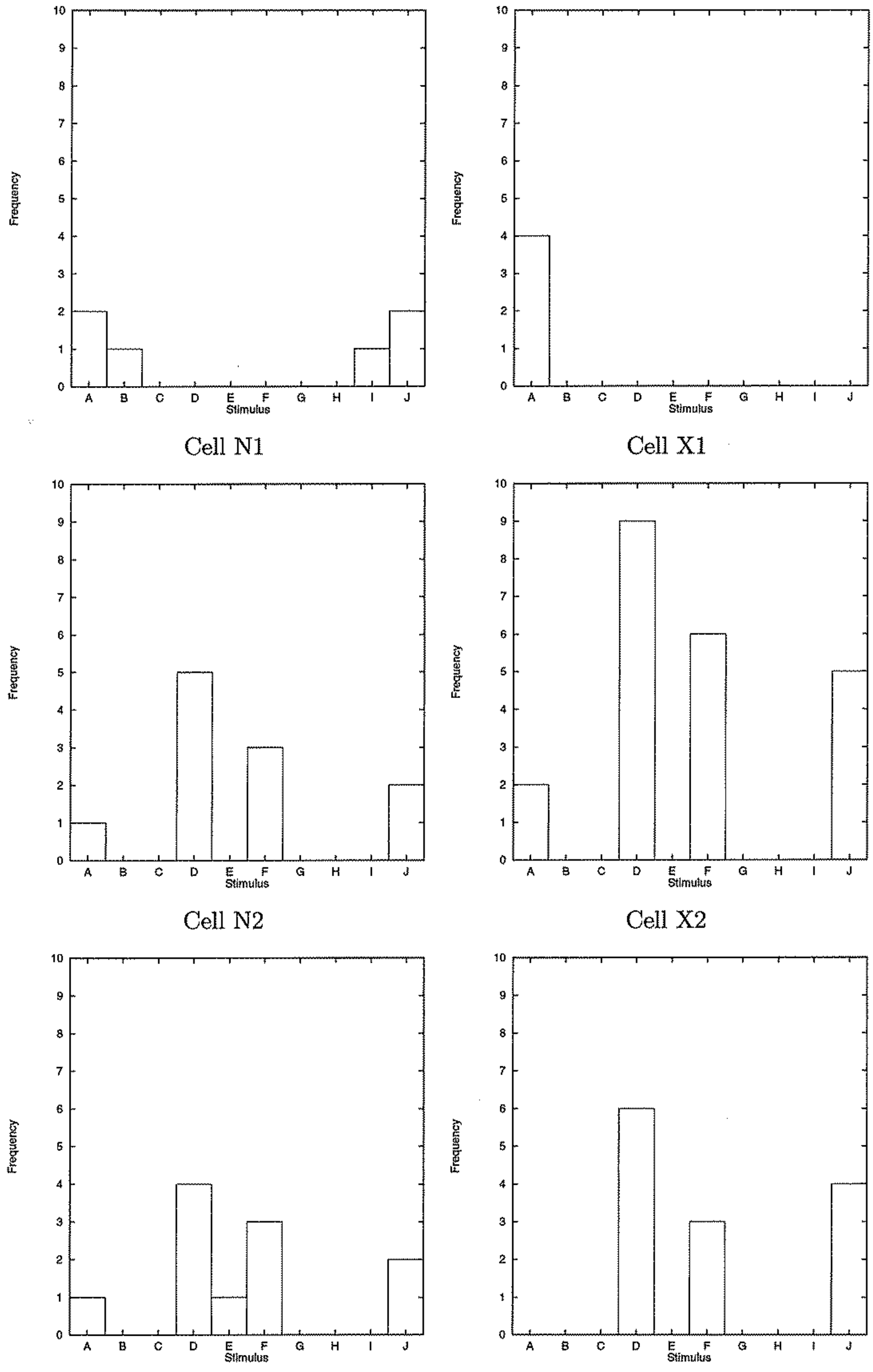


Figure 7.15: Horizontal Proximity Error Histograms

7.6 Task 3: Memory Task

7.6.1 Recall Times

Plots of memory recall times are shown in Figure 7.16. Descriptive statistics for display D crossed with group G and for number of aircraft A crossed with sample type R are shown in Table C.22. Two-factor ANOVA analyses were carried out to test the effects on recall time of display and group (D against G, Table C.23), and to test the effect of number of aircraft and whether or not the scenarios were random (A against R, table C.24).

The analyses suggest that number of aircraft, scenario randomness, display and group all had significant effects on the recall times, but that there was no significant interaction between the factors in either of the analyses.

1. Experts showed significantly faster recall times than novices.

2. Recall time was fastest for the 2D PPI, second fastest for the stereo-3D display and slowest for the pseudo-3D display.

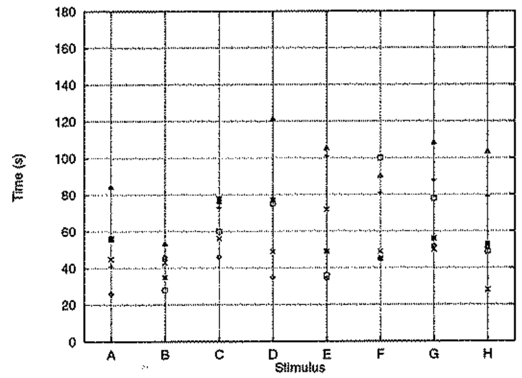
3. Recall time was significantly faster for scenarios taken from real traffic samples than for randomly-generated scenarios.

4. Recall was fastest for 3 aircraft scenarios, then was roughly level for the 5, 7 and 9 aircraft scenarios. This may be due to 'saturation' occurring for these scenarios.

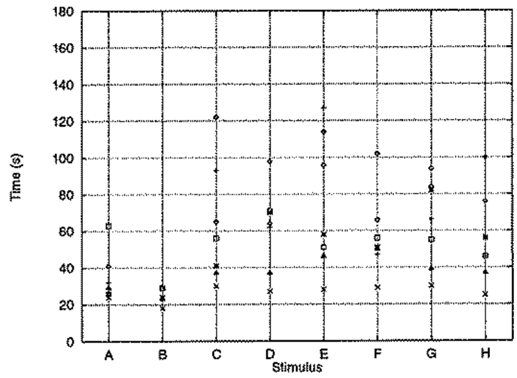
7.6.2 Recall Analysis Method

When it came to analysing the reconstructed scenarios, there was great difficulty in arriving at useful measures of accuracy of recall of position, height and callsign, especially for scenarios in which the number of aircraft was greater than 3.

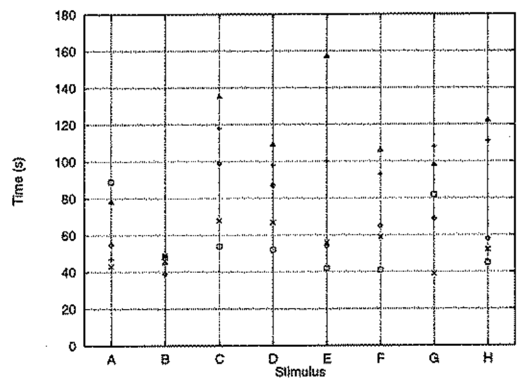
For position what is required is some index of 'goodness of fit' of the recalled positions with the actual positions. The obvious method is to measure the distance of each recalled point from the actual position, but this only works well where there is an easily discernible one-to-one correspondence between recalled points and stimulus points; i.e. where the overall *pattern* the points make correspond reasonably closely between the stimulus and recalled scenarios. However, if the patterns differ



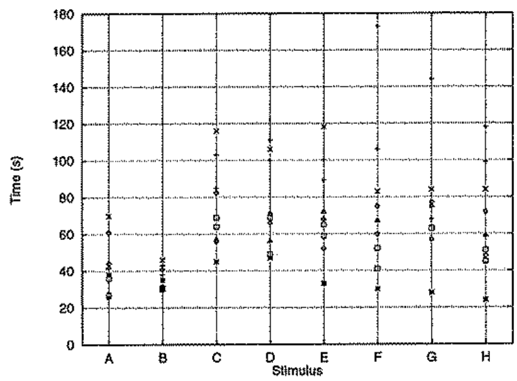
Cell N1



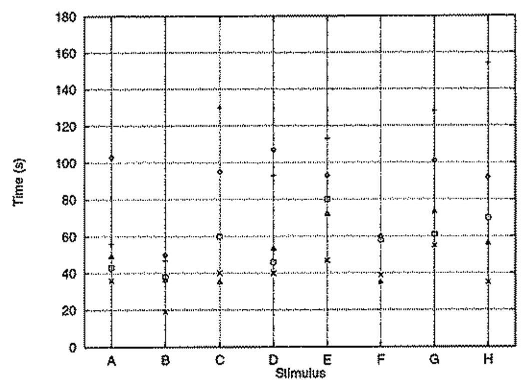
Cell X1



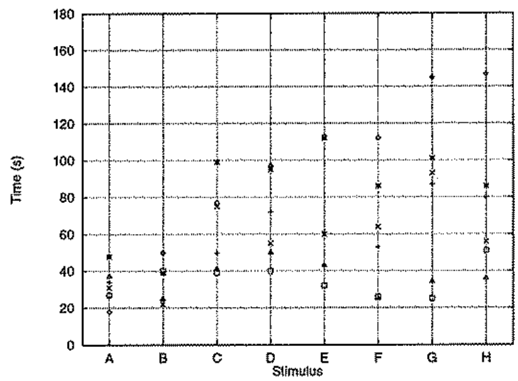
Cell N2



Cell X2



Cell N3



Cell X3

Figure 7.16: Memory Recall Time

significantly, for example if the recalled scenario contains false additional points or points far removed from the stimulus pattern, then this approach becomes more difficult, since it becomes a more subjective judgment as to which points in the stimulus scenario correspond to which points in the recalled scenario. Further, there is evidence that subjects were remembering aircraft in 'clumps' so that the positions of the aircraft within a 'clump' may have been accurate, but the overall placement of the clump may have been inaccurate, and this approach fails to take this into account.

The accuracy of callsigns was difficult to judge. There were several classes of error. For callsign/aircraft correspondence:

- Callsign assigned to incorrect aircraft.
- No callsign assigned to aircraft.

and for callsign correctness:

- Alphabetic callsign correct, but numeric component incorrect:
  - Numeric component missing.
  - Numerals transposed.
  - Some numerals correct, some incorrect.
- Alphabetic callsign incorrect, but numeric component correct:
  - Alphabetic component missing.
  - Alphabetic component totally wrong.
  - 3 letter designator recalled instead of 2 letter (e.g. BAW instead of BA) in the case of expert subjects (see below).
- Callsign totally incorrect.

As the number of aircraft increased and the stimulus pattern and the recalled pattern became less and less alike, determining whether a callsign was assigned to an aircraft correctly became more and more difficult, especially if the callsign itself was incorrect.

The accuracy of heights were most difficult of all to judge. Again several types of error were identified:

- Correct height assigned to incorrect aircraft.
- Incorrect height:
  - Digits transposed or otherwise permuted.
  - Heights rounded in some way.
  - Missing digits (e.g. only the first digit recalled).
  - Relative heights correct, but numbers incorrect or missing (e.g. by drop line length and/or numbers).
  - Relative heights incorrect.
  - No height information.

Most subjects did not try to record drop line length but numeric values or nothing at all. Again, accuracy with these sorts of error was very difficult to define or determine.

These problems ruled out numerical analysis of the results, since suitable measures of accuracy could not be determined. A simple count of errors divided by categories also proved to be impossible; it was not always clear which category an error belonged to, especially in cases of assigning a callsign or height to a particular aircraft.

It was therefore decided to analyse the data by reviewing the scenarios together with the material gathered in the questionnaire as to how subjects remembered and recalled the information, and possibly to relate this to subjective difficulty. Subjective difficulty ratings are examined first.

7.6.3 Subjective Difficulty

Subjects were asked to rate the difficulty of components of the task on a discrete scale of 1–7: Overall, Recall of position, Recall of identification (callsign) and Recall of height. It was intended that these should be relative to each other; although this was not explicitly stated on the instruction sheet, this was clarified to those subjects who asked. Histograms of the estimated task difficulty are shown in Figures 7.17–7.20.

For recall of callsign, two subjects made *two* entries: One for the alphabetic component, one for the numeric component (both rated the alphabetic component

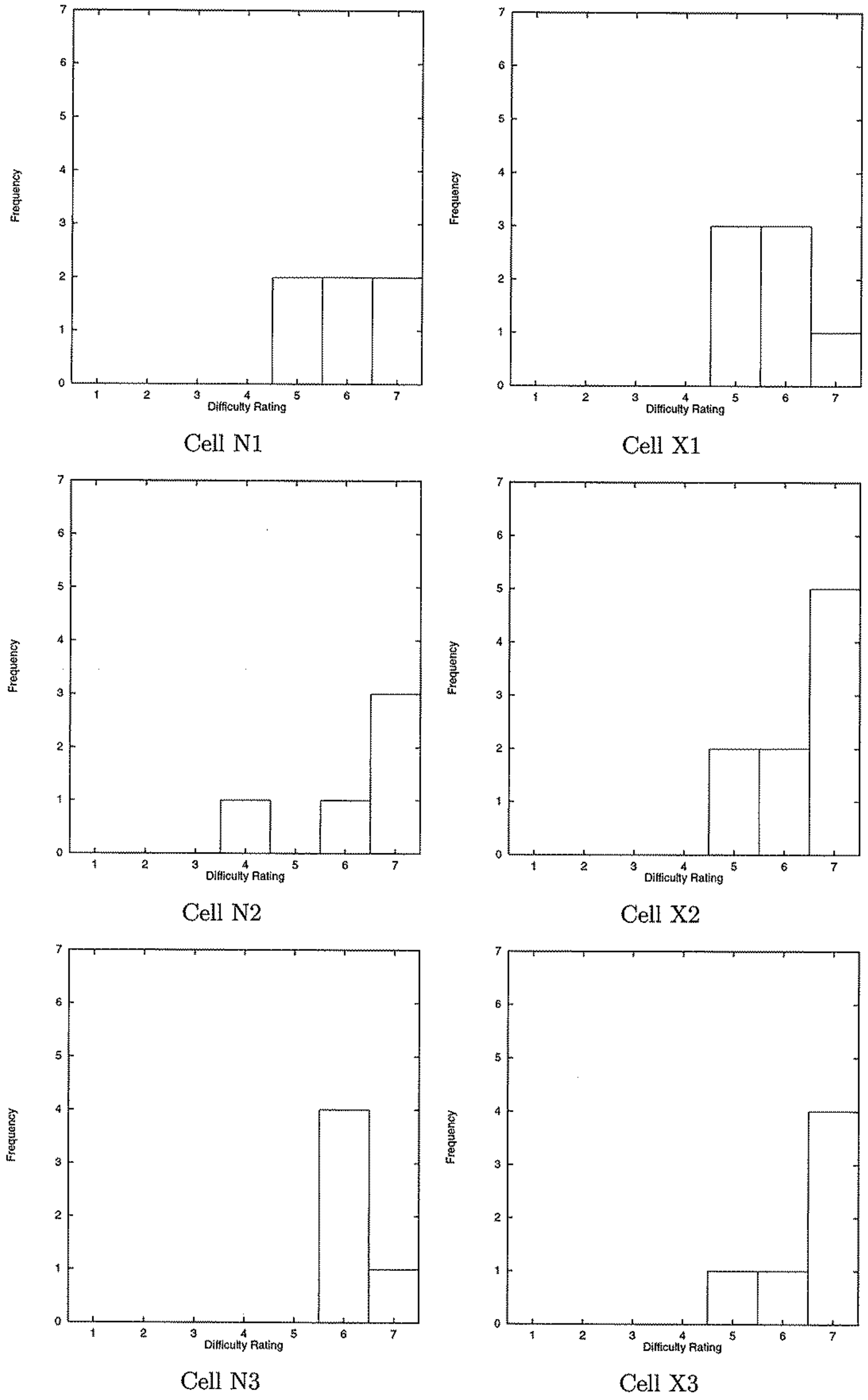


Figure 7.17: Estimated Overall Memory Recall Difficulty

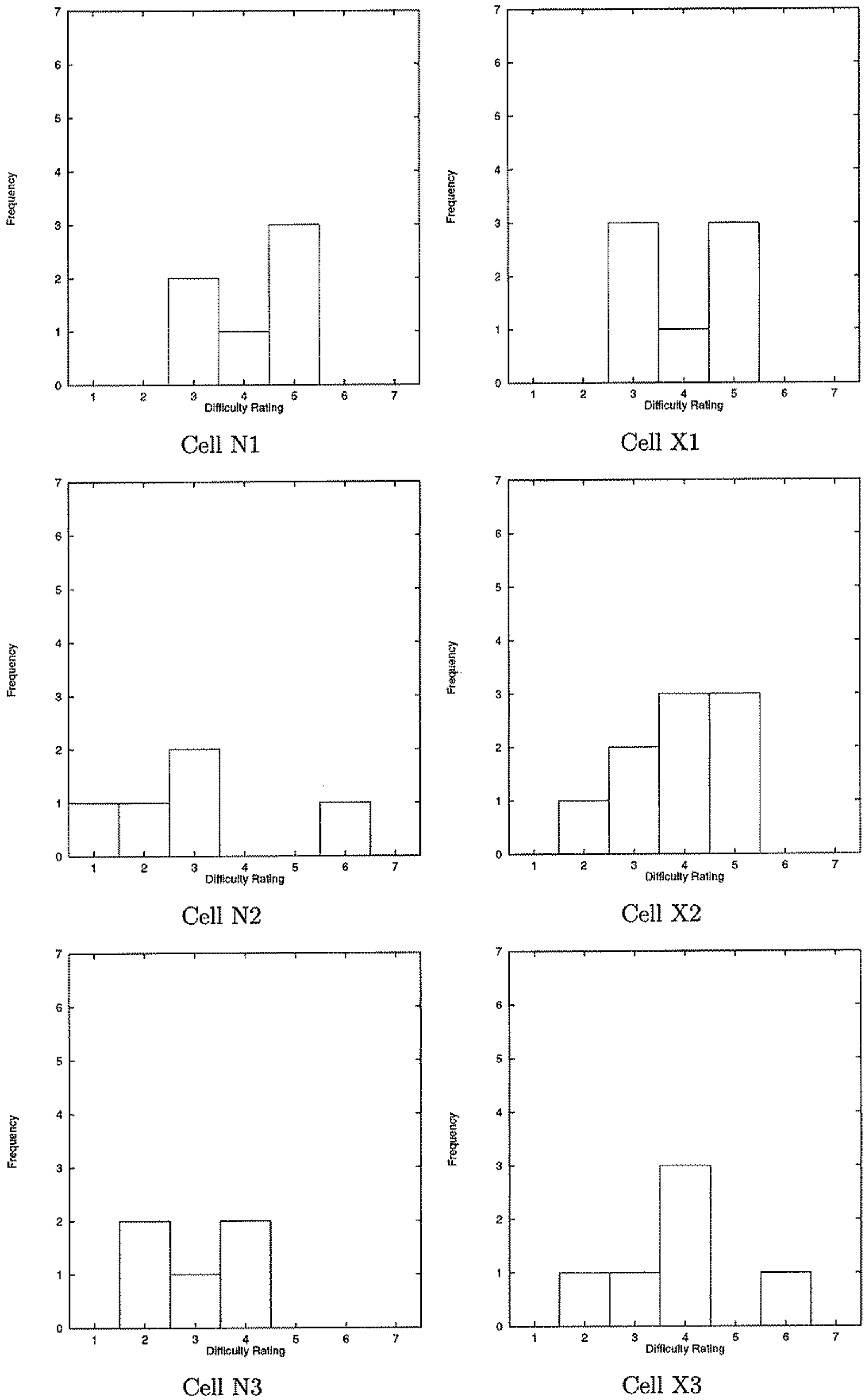
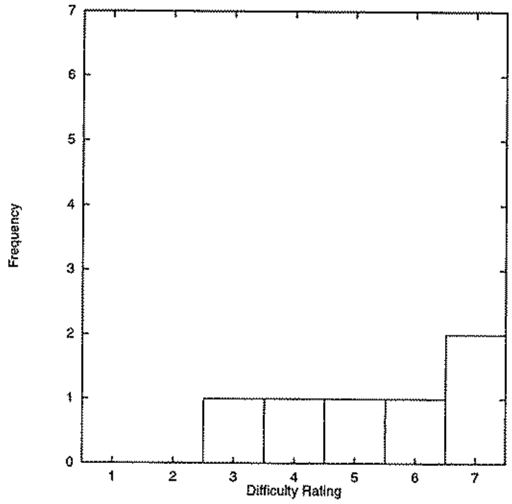
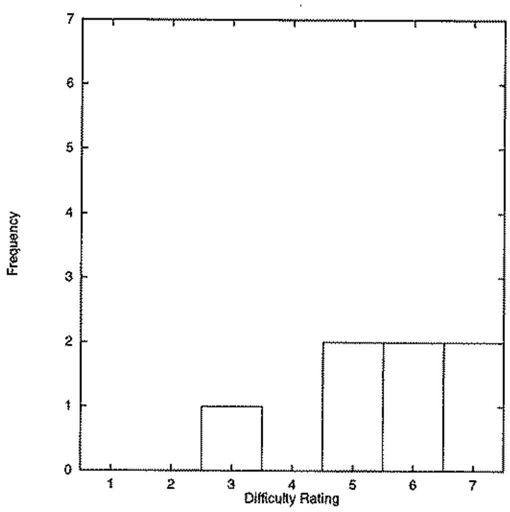


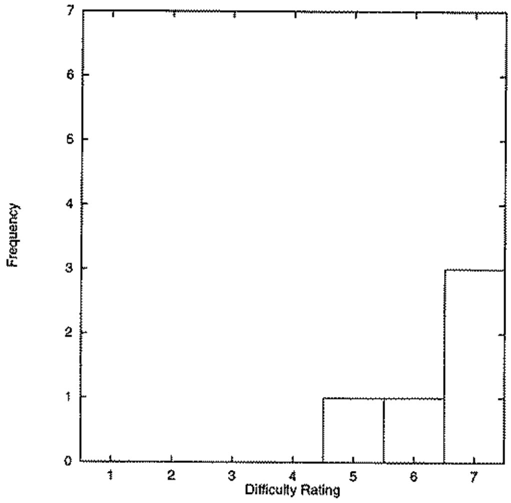
Figure 7.18: Estimated Position Recall Difficulty



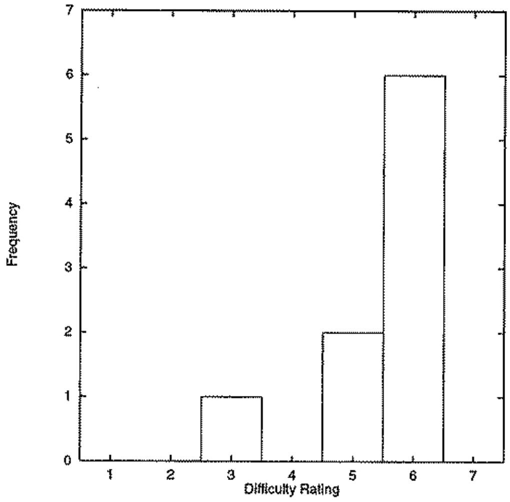
Cell N1



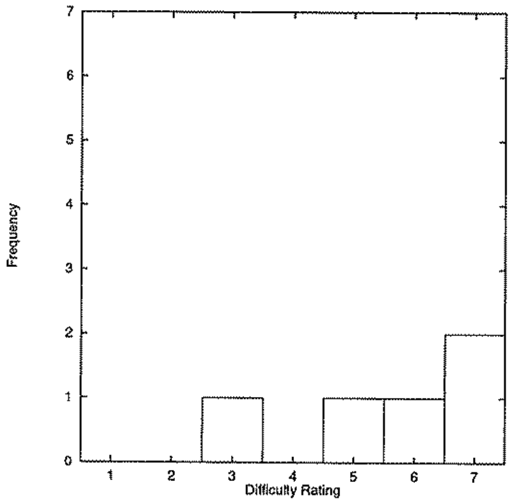
Cell X1



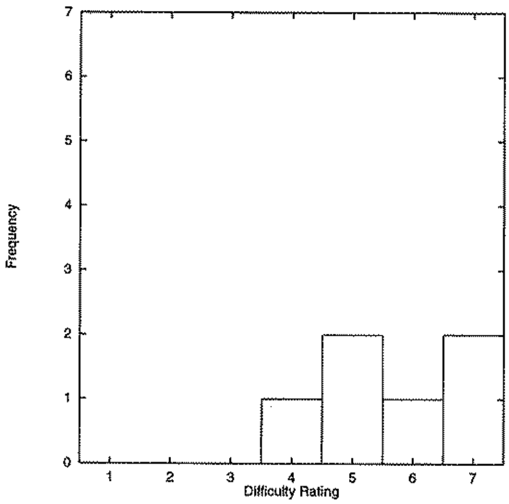
Cell N2



Cell X2



Cell N3



Cell X3

Figure 7.19: Estimated Identification Recall Difficulty

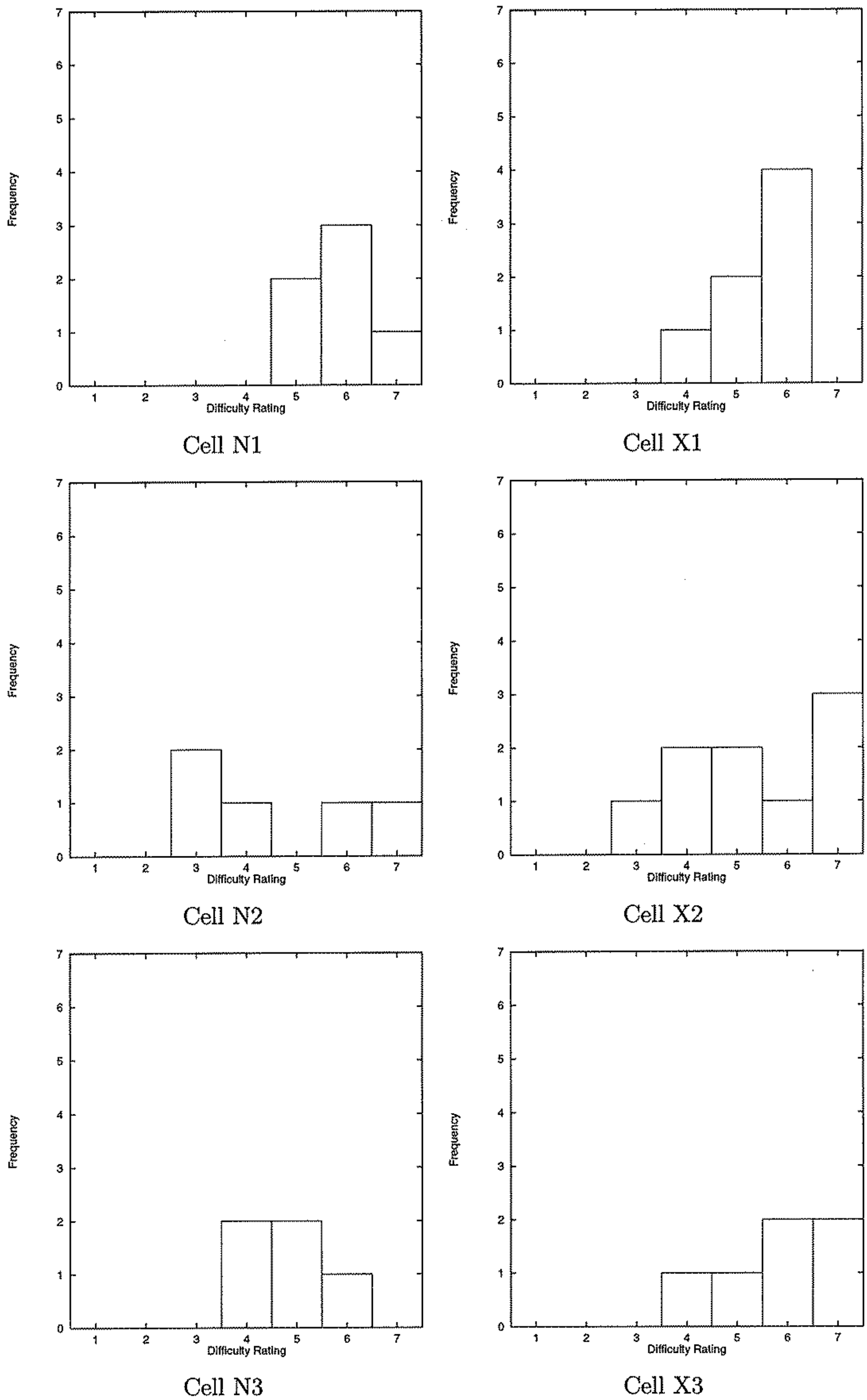


Figure 7.20: Estimated Height Recall Difficulty

difficult as 2, the numeric component difficulty as 6). For these entries, the higher rating (6) was chosen. One subject indicated a difficulty of 4, but wrote that the alphabetic component was easier than the numeric component.

Descriptive statistics for recall difficulty ratings are given in Tables C.25–C.28. ANOVA analyses looking at the effects of display type and group failed to find any influence of these factors on the subjective difficulty ratings, but suggested that subjects found recall of position to be easier than either height ($F = 34.54$, $F_{75,0.05} = 3.97$, $p < 0.01$) or identifier ($F = 46.05$, $F_{75,0.05} = 3.97$, $p < 0.01$) (median difficulties: Position = 4.0, height = 5.3, identifier = 5.6).

7.6.4 Recall Analysis

Position

Despite the difficulties outlined above in measuring position recall accuracy, position error data were obtained for scenarios containing 3 and 5 aircraft. Position accuracy was measured by overlaying transparent acetate sheets on which a stimulus and a 1 mm grid were printed onto a response sheet, deciding which recalled aircraft corresponded to which aircraft on the stimulus (not necessarily the closest, but the one which best matched the 'pattern' of the stimulus) and measuring the x and y offsets (amount east/west and north/south of the target) in millimetres. For scenarios containing 7 and 9 aircraft, it was too difficult to see which of the recalled aircraft corresponded to which aircraft in the stimulus.

The position error x (east/west) and y (north/south) components were not found to be correlated with display type (unlike Burnett & Barfield's memory experiment which showed recalled position to be consistently to the north of target stimulus for the 3D display format).

Absolute position error (distance of aircraft recalled position from true position $= \sqrt{x^2 + y^2}$) was analysed by two-factor ANOVA analyses looking at the influence of display and group (Table C.30) and number of aircraft and scenario randomness (Table C.31). (Descriptive statistics are shown in Table C.29.) Display and the interaction between display and group were found to have a significant influence on position error. Group itself was not a significant factor. The 2D PPI display was found to give smaller position errors than either of two 3D display types for both novices and experts. For novices, the pseudo-3D display gave smaller position errors than the stereo-3D display; this was reversed for experts.

Number of aircraft and scenario randomness were both found to have a significant influence on position error. As expected, the aircraft positions in the 3-aircraft scenarios were recalled more accurately than those in the 5-aircraft scenarios, and non-random scenarios were recalled more accurately than random ones.

In recalling position, some subjects reported doing an initial count of the number of aircraft to serve as a check (e.g. Quote F.52). Strategies for remembering aircraft position were largely centered around looking for patterns of some description. Novice subjects (apart from one) did not recognise traffic patterns for what they were but used various visualisation strategies: Remembering patterns by geometric shapes (e.g. Quotes F.1, F.26), joining the dots to make pictures (Quote F.6), clustering aircraft into 'clumps' and making patterns within those clumps, and referring clumps and individual aircraft to the video map/range rings (e.g. Quotes F.33, F.64). Some expert subjects recognised traffic patterns and tried to memorise in terms of these where possible (e.g. Quotes F.117, F.102, F.110), since these they reported to be easier to remember (e.g. Quotes F.117, F.129). Where they did not recognise a traffic pattern, they used similar strategies to the novices.

One subject reported that 'lone' (Quote F.126) or conflicting aircraft (Quote F.128) were easier to remember (another subject made a similar comment verbally). This may be an example of controllers identifying aircraft which do not conform to the pattern or which 'do not look right' (see footnote on p. 39).

Subjects reported that position was easier to remember than height or identification, which is supported by the evidence of the subjective difficulty ratings.

Height

Strategies for remembering aircraft heights were found to change as the number of aircraft increased, and according to whether the aircraft were in a random pattern or not. For the 3-aircraft scenarios, most subjects were successful at simply remembering the numbers. As the load increased, some subjects resorted to other methods, such as grouping into height bands and looking for patterns, and others decided to forget heights altogether to concentrate on other, easier aspects, or only to remember salient heights. The fact that combinations of strategies were used, and that the strategies changed with number of aircraft and scenario type, makes analysis of subject statements by grouping into categories quite difficult. However, bearing this in mind, it may be instructive to try to categorise statements. Below are given height recall strategies, with the number of subjects reporting using that strategy and the breakdown of those subjects by cell.

Numeric Memory A lot of subjects used memory of numbers in combination with other strategies, but a number reported using this as the main or sole memory strategy, even in 3D display formats. Of these, two reported that salient drop line heights stuck in the memory (Quotes F.25, F.164). Note that all bar one of the eight subjects who reported using numeric memory were using 3D displays.
8 subjects: N1:1, N2:2, N3:1, X1:0, X2:3, X3:1[1]

Height Patterns Some scenarios containing traffic patterns showed approaching and departing traffic. Such aircraft would be in 'clustered' in close geographic proximity and would have similar heights. Aircraft leaving a holding stack, joining the extended runway centreline and then flying down it to the runway would be descending, so the traffic stream would have decreasing height. Similarly, a stream of aircraft departing an aerodrome would have ascending heights. Four novice subjects reported identifying such sequences, which helped to stick in memory. Six ATCOs reported using familiar traffic patterns, and their knowledge of the heights of the aircraft within those patterns or the sequence of heights, to remember scenarios.

Of all the subjects who reported using height patterns or sequences, only one of these (an expert) was using a 2D display format. It is speculated that the 3D format may have helped height sequences rapidly to be identified.
10 subjects: N1:0, N2:2, N3:2, X1:1, X2:2, X3:3

Grouping by Similar Height/Height Rounding
7 subjects: N1:1, N2:3, N3:0, X1:0, X2:2, X3:1

Length of Drop Line Only two subjects (both novices) reported trying to remember the length of the drop line.
2 subjects: N1:0, N2:2, N3:0, X1:0, X2:0, X3:0

[1] i.e. 8 subjects reported using this strategy; one from cell N1, two from cell N2, one from cell N3, three from cell X2 and one from cell X3.

Relative Heights Five subjects reported trying to remember relative heights of aircraft. Relative height may be a reference to height patterns, but it is difficult to be certain.
5 subjects: N1:0, N2:2, N3:1, X1:0, X2:0, X3:2

Visualising Heights Four subjects (all using 2D displays) reported trying to visualise heights.
4 subjects: N1:2, N2:0, N3:0, X1:2, X2:0, X3:0

Other Strategies

Association with Historical Dates One subject studied history as a hobby and tried to associate numbers with dates (Quote F.72).

Association with Airline Code (Quotes F.84 and F.125).

Pattern Violations and Conflicts One ATCO reported looking for "anything unusual in the heights of two aircraft relative to their positions" (Quote F.103); one reported that conflicting heights were remembered first (Quote F.133).

Three subjects reported the drop lines of being little or no use, apart from for picking out extremely high or low aircraft. Apart from these comments, there is insufficient information to be able to link any particular height memorisation behaviour with any particular display format.

Callsign

The callsign was composed of two parts: A two-letter 'airline' code (in some cases real, in some cases bogus) and a three digit 'flight number' code.

Although two-letter callsigns were used, real aircraft callsigns now contain three characters (for ATC purposes), which was not appreciated at the time that the experiment was designed. A British Airways aircraft was known in the past by the callsign prefix 'BA' but the prefix is now 'BAW'. Some expert subjects recalled the three-letter prefix instead of the two-letter prefix, and one afterwards reported verbally that he didn't notice the discrepancy. This may be because although airlines are identified by a callsign prefix, controllers are required to speak to the aircraft using a similar but related name (e.g. BAW2024 would be called 'Speedbird 2024', DLH102 would be called 'Lufthansa 102' etc.) and refer to the aircraft by that name or a similar abbreviation (e.g. 'the Speedbird', 'the Luftie' etc.) when talking to others face-to-face. When looking at a familiar callsign, it is speculated that expert subjects do not remember the code but the name associated with the code, and when recalling they note the code associated with the name.

ATCOs still found some of the callsigns familiar, and reported that the use of familiar callsigns helped. Although the numeric codes were random, some callsigns inadvertently corresponded to real aircraft, and some ATCOs noticed familiar aircraft at the incorrect location (e.g. a London–Paris flight well to the North of Heathrow); one ATCO reported that this particular flight was particularly salient for this reason. Some novices also found some airline codes familiar (e.g. BA, AF) since flights are still known by the two-letter prefix codes for ticketing purposes, and reported that these were retained in memory better.

For unfamiliar codes, both novices and expert used mnemonics or similar strategies (e.g. trying to make words out of the prefixes of two adjacent callsigns).

From the recalled scenarios and the subjects' comments, the numeric component of the callsign was the least likely element of the scenario to be remembered. A numeric height code at least relates to something real so subjects could try to visualise height or relative height between aircraft, or look for sequences, but as a subject observed, the flight numbers were "just a random bunch of numbers" (Quote F.50). Some subjects attempted to make 'patterns' out of the numbers; one reported that some sequences were memorable (e.g. 468, 137) (Quote F.121), and some tried to associate the height and the callsign, but by and large, the existence of any pattern was due to coincidence.

7.6.5 Discussion

Effect of Number of Aircraft

As the number of aircraft in the scenarios increased, subjects tried alternative strategies to remember the various elements, and became selective about what they tried to remember depending on how easy it was to commit those elements to memory. Generally speaking, the elements of the scenarios in order of ease of memory (easiest to remember first) were:

1. Position (especially of relative positions of groups of aircraft)

2. Callsign alphabetic prefix

3. Height

4. Callsign flight number

This is supported by the subjective difficulty ratings. However, within a scenario, some element could be particularly salient (for example, a flight number or sequence of heights could be remembered where other heights or flight numbers were not).

The relative difficulties were much as anticipated, except that subjects did not appear to use the direct visualisation of height in the 3D displays. It was thought that a 3D display might better convey approximate heights, so retention of these would be better than for the 2D display. It was anticipated that as memory load increased, subjects would be more inclined to remember approximate heights and relative, using the patterns formed by the drop lines in the case of 3D displays, rather than the precise heights. However, this did not appear to occur.

One explanation is that the task may have encouraged the subjects to try to record precise height information where possible, instead of more approximate or relative height information. Since heights were presented in a digital format, and could easily be drawn as such on a piece of paper, subjects may have been biased towards using 'number memory' instead of trying to remember relative heights as the memory burden increased. At least one subject did not appreciate that relative height information would have been acceptable (Quote F.17). Further, there was no adequate method indicated of how to record the relative heights of aircraft. The instruction sheet indicated that a 'best guess' was acceptable, and that subjects could draw the drop-line length to indicate approximate height, but there seemed to be a reluctance to do so; there is evidence that some subjects chose to remember some more memorable elements more accurately (such as position) than other aspects poorly, concentrating their memory resources on easier things (e.g. Quote F.141).

Effect of Scenario Randomness

Scenario randomness is perhaps a misnomer. As anticipated, subjects seemed to be better at recalling scenarios derived from traffic samples rather than random scenarios, but this appears to be due to the fact that the aircraft selected from traffic sample scenarios tended to form discernible groups or patterns as regards both their positions and heights (e.g. streams of aircraft landing or taking off), whereas aircraft in random scenarios were scattered. Novices had no prior knowledge of the traffic patterns, but were still able to identify and use patterns to help them to remember the scenario. It is likely that if scattered aircraft were selected for the traffic sample scenarios, and if made-up scenarios were less random but contained groups of aircraft with height sequences within the groups, then the made-up scenarios would be remembered better than the traffic sample scenarios.

Effect of Subject Group

As expected, expert subjects were able to apply their knowledge of traffic patterns in scenarios which contained these, and some reported that familiar patterns were easier to remember than unfamiliar ones (e.g. Quote F.117). Familiar callsigns would have made recollection easier still (Quote F.98).

Effect of Display Type

The position recall accuracy data which were obtained suggested that 2D displays gave better accuracy than either of the 3D display formats. This could be due to the ambiguity of reading distances along the LOS in 3D displays (which could also explain Burnett & Barfield's finding that aircraft positions recalled from a 3D presentation were biased significantly towards the north of the true aircraft position—the display was viewed from the south and so distances in the north-south plane would have been subject to ambiguity). However, the fact that the 3D displays had to be reconstructed in plan caused a problem with at least one subject (Quote F.182), and this may also have had an influence.

It is possible that 3D display formats allowed for more rapid identification of height patterns and approximate heights, but this task yielded no evidence to support or refute this.

Critique

Some useful information was gained from this task regarding how subjects behaved when trying to memorise a great deal of information presented as different codes. However, the task failed to address adequately the chief concern of this thesis, namely, the effect of display type on the memory of spatial information. Due to the extremely heavy load on subjects' memory, recalled information tended to be poor and was such that most quantitative data could not be statistically analysed. Analysis therefore had to focus primarily on subject comments, using the recalled scenarios as supporting evidence, which is a less desirable state of affairs than the opposite case.

The questionnaire used to gain subject comments was limited in that it was an open questionnaire rather than one which asked specific questions. It was therefore open to the dangers of subjects forgetting to give information, or not giving it as it might be considered 'obvious' or non-pertinent, or simply not thinking of it.

Overall, it was felt that the task tried to gain evidence to support too many hypotheses, and ended up gaining only circumstantial evidence pertaining to some of the hypotheses. It would have been better if it had been targeted more specifically, resulting in a more sensitive experiment design. Future experiments should therefore be carried out to test each of the hypotheses in isolation, if possible, especially on the key questions of how display format influences memory and recall of spatial information.

7.7 Task 4: Conflict Detection Task

7.7.1 Conflict Detection Responses

For each of six scenarios (which contained conflicts or not according to Table 6.4) subjects either indicated that a conflict existed or did not. If a scenario contained a conflict and the subject selected that conflict, or if the scenario contained no conflict and the subject indicated as such, this was classified as a *correct* response. A response was *incorrect* where subjects either failed to identify the correct conflict in scenarios containing a conflict, or indicated a conflict in scenarios which did not contain a conflict.

The number of correct and incorrect responses for each scenario and cell are tabulated in Table C.32. The percentage of incorrect responses for each cell are shown in Figure 7.21. These were analysed by logistic regression. It was found that subject response depended significantly on the display type ($\chi^2 = 6.73$, $\chi^2_2(0.05) = 5.991$) but not on group or the interaction between display and group. The proportion of incorrect selections per display type were: 2D PPI 12.8% (10 of 78), pseudo-3D 31.0% (26 of 84), stereo-3D 24.2% (16 of 66). Thus, significantly more incorrect selections occur for subjects using the 3D display formats than for subjects using the 2D PPI display.

Histograms of the percentage of incorrect responses in each cell for each scenario are shown in Figure 7.22. It can be seen that certain scenarios give a much greater proportion of incorrect selections than others. Possible reasons for this are discussed in §7.7.6 below.

7.7.2 Conflict Detection Times

Conflict detection times are plotted in Figure 7.23. Descriptive statistics are given in Table C.34. A two-factor ANOVA analysis (Table C.35) found that subject group had a significant influence, but display and the interaction term did not (expert group (41.2 ± 20.6 s, $n = 74$) was significantly faster than the novice group (49.1 ± 18.5 s, $n = 43$)).

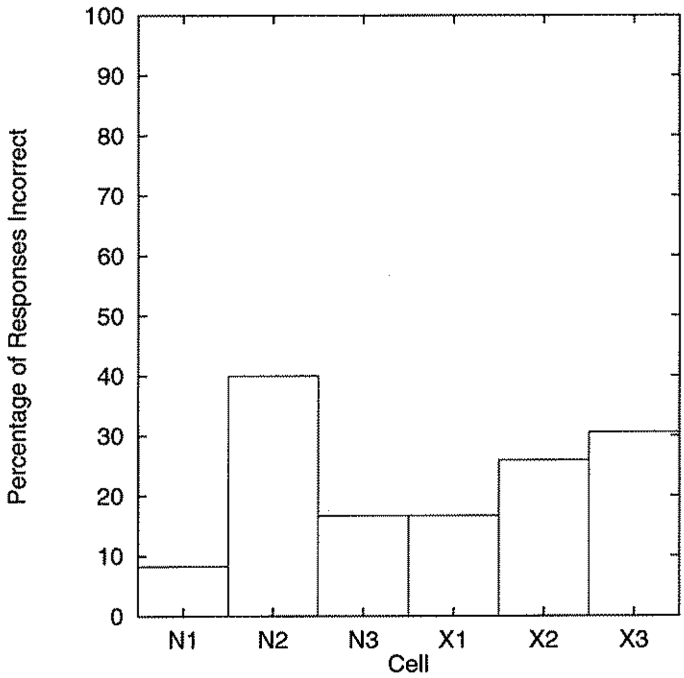


Figure 7.21: Incorrectly Detected Conflicts by Cell

7.7.3 Conflict Detection Techniques

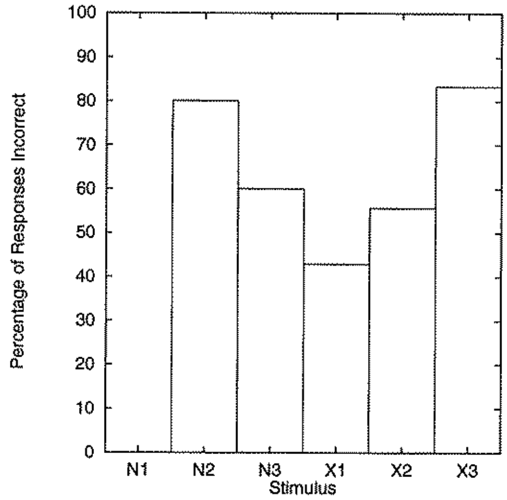
The task questionnaire asked subjects to note down how they carried out the conflict detection task. The responses are given in §G.1. These reveal the techniques used by novice and expert subjects to be remarkably similar (except that novices tended to disregard speed). The overall technique seems to be as follows:

1. First, scan for immediate conflicts by looking at

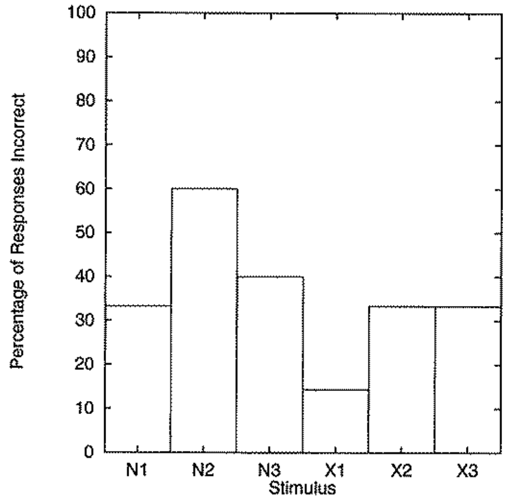
    - ground position (for lateral separation)
    - height (for vertical separation)

    (some subjects reported looking at position first, some at height first, some did not indicate which they considered first, if anything).

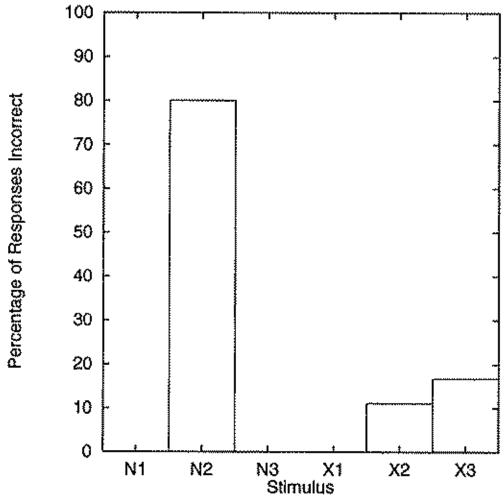
2. 'Background' aircraft unlikely to be in immediate confliction (e.g. aircraft clear of anything else). For the other aircraft, look at manœuvres:



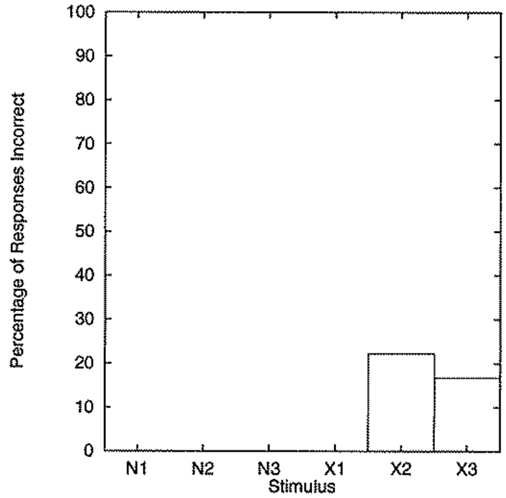
Scenario A



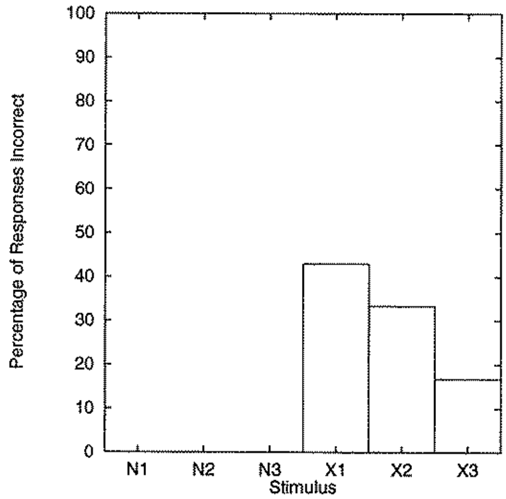
Scenario B



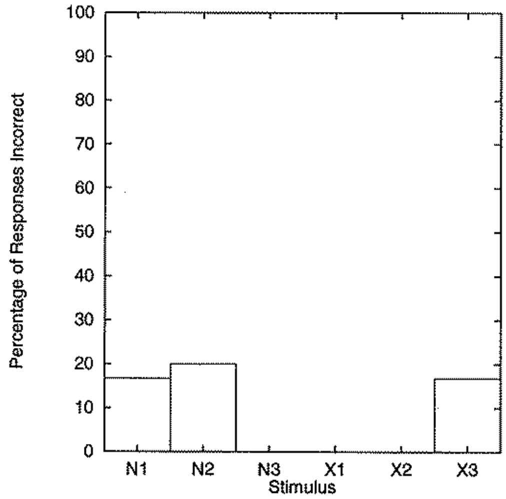
Scenario C



Scenario D

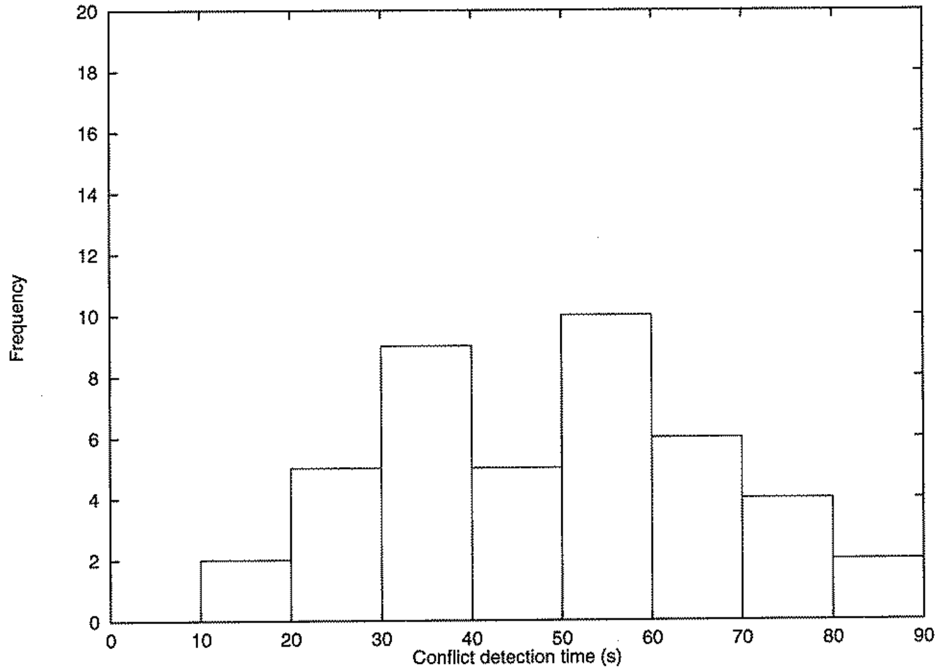


Scenario E

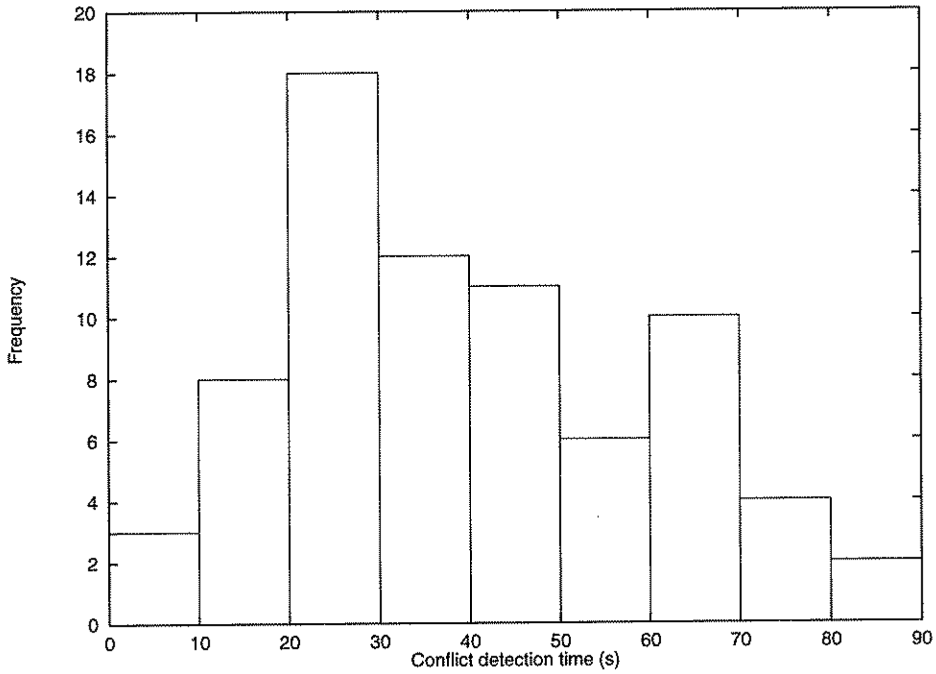


Scenario F

Figure 7.22: Incorrectly Detected Conflicts by Scenario



Novice Group



Expert Group

Figure 7.23: Conflict Detection Time Histograms

- tracks (as the trailing histories appear)[2]
- vertical manœuvres

to look for impending conflictions.

3. Continue scanning, with higher priority to some aircraft than others.

For the 3D display formats, where subjects reported the type of position information they used, ground position was mentioned far more than air position. (The distinction does not exist in the 2D display, of course!)

7.7.4 Conflict Resolution

Conflict resolution instructions were divided into four categories: Horizontal manœuvre, Vertical manœuvre, Speed and Combined (combined horizontal *and* vertical manœuvre and possibly speed). No pure speed resolution instructions were given by subjects. Histograms of conflict resolution instruction type versus cell is shown in Figure 7.24, and a histogram of conflict resolution instruction type versus scenario is shown in Figure 7.25. From Figure 7.24 it is apparent that the class of conflict resolution instruction given is not dependent on display or subject group.

In a few cases, the type of resolution instruction given was observed to depend on when the conflict was spotted. For example, in scenario C, one aircraft was climbing through the level of the other and passed within 0.3 n.m. horizontally if the scenario was left to run uninterrupted. Subjects often entirely missed the conflict, or spotted it sufficiently in advance to level the climbing aircraft off at a lower altitude until the aircraft had crossed. However, the experimenter observed that a few subjects spotted the conflict only when the vertical separation was less than 1 000 ft, so a vertical instruction would have been ineffective and a horizontal conflict avoidance was the only alternative[3].

[2]When the scenarios started, aircraft did not have trailing histories—these built up as the scenario unfolded, so track information could not be immediately assessed.

[3]However, even a horizontal avoidance instruction would have probably been ineffective in this scenario with $< 1\,000$ ft vertical separation. A similar situation probably existed in an anecdote in which a US controller is quoted: "Lear7P, if you still read me, traffic is no longer a factor."!

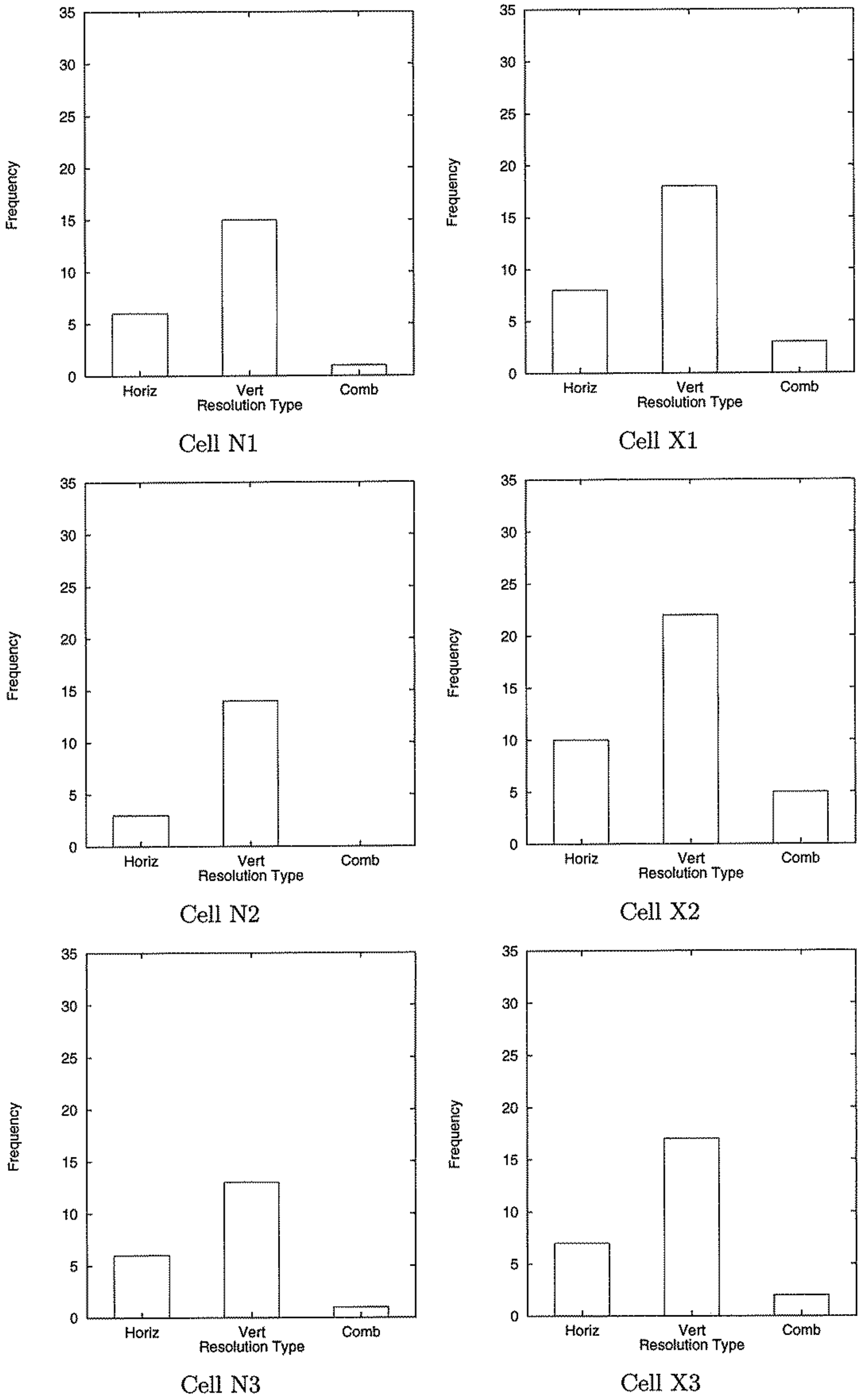


Figure 7.24: Conflict Resolution by Cell

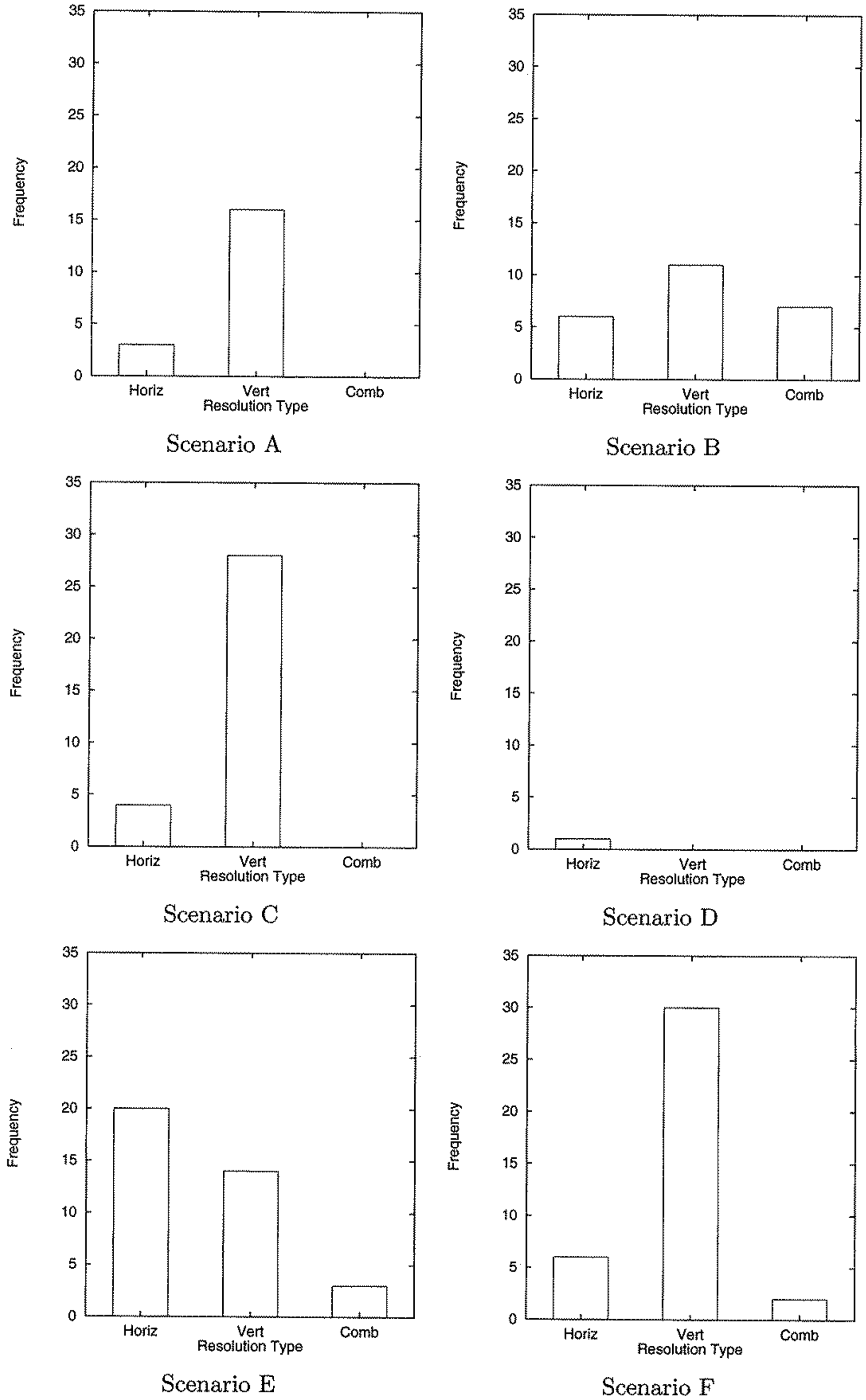


Figure 7.25: Conflict Resolution by Scenario

7.7.5 Subjective Difficulty

Subjects were asked to rate the difficulty of the conflict detection task on a discrete scale of 1–7. Histograms of the estimated task difficulty are shown in Figure 7.26. Descriptive statistics are shown in Table C.33. Analysis by two-factor ANOVA revealed no significant influence of display, group or the interaction of the two.

7.7.6 Discussion

Conflict Detection Responses

Conflict detection is a difficult task to assess because it is partly subjective. Despite the inclusion of route codes, some expert subjects were unsure of where aircraft were going, which they would have known had they been actually controlling the aircraft:

> *As some aircraft were turning, not knowing whether they were going to roll out on a safe heading or not doesn't help. In the "real world" as you would have more idea of what the aircraft* should *be doing the task would have been slightly easier."* (Quote G.48)

If subjects had been controlling the aircraft, they would have a plan whereby some of the potential conflicts would have been resolved. However, some controllers reported verbally that not knowing the overall plan made the task more difficult. They may have thus perceived possible conflicts and preferred to take action on a 'better safe than sorry' basis. As Burnett and Barfield describe:

7.1 ... *because each controller develops a different traffic management style, one controller may perceive a conflict exists for any given situation and a different controller may perceive that a conflict is non-existent for the same situation if certain procedures are immediately executed.* [BB91]

It is therefore a bit of a misnomer to classify some responses as 'incorrect' in cases where individuals perceived a conflict to exist where there wasn't one. The incorrect responses of scenarios A and D are certainly in this category, since there were no occurrences of impending or actual loss of separation in either of these scenarios.

Scenario A is particularly interesting. Figure 7.22 shows that more 'incorrect' conflicts were detected for the 3D displays than for the 2D display (logistic regression analysis supports this, finding display to have a significant effect ($\chi^2 = 7.388$, $\chi^2_2 =$

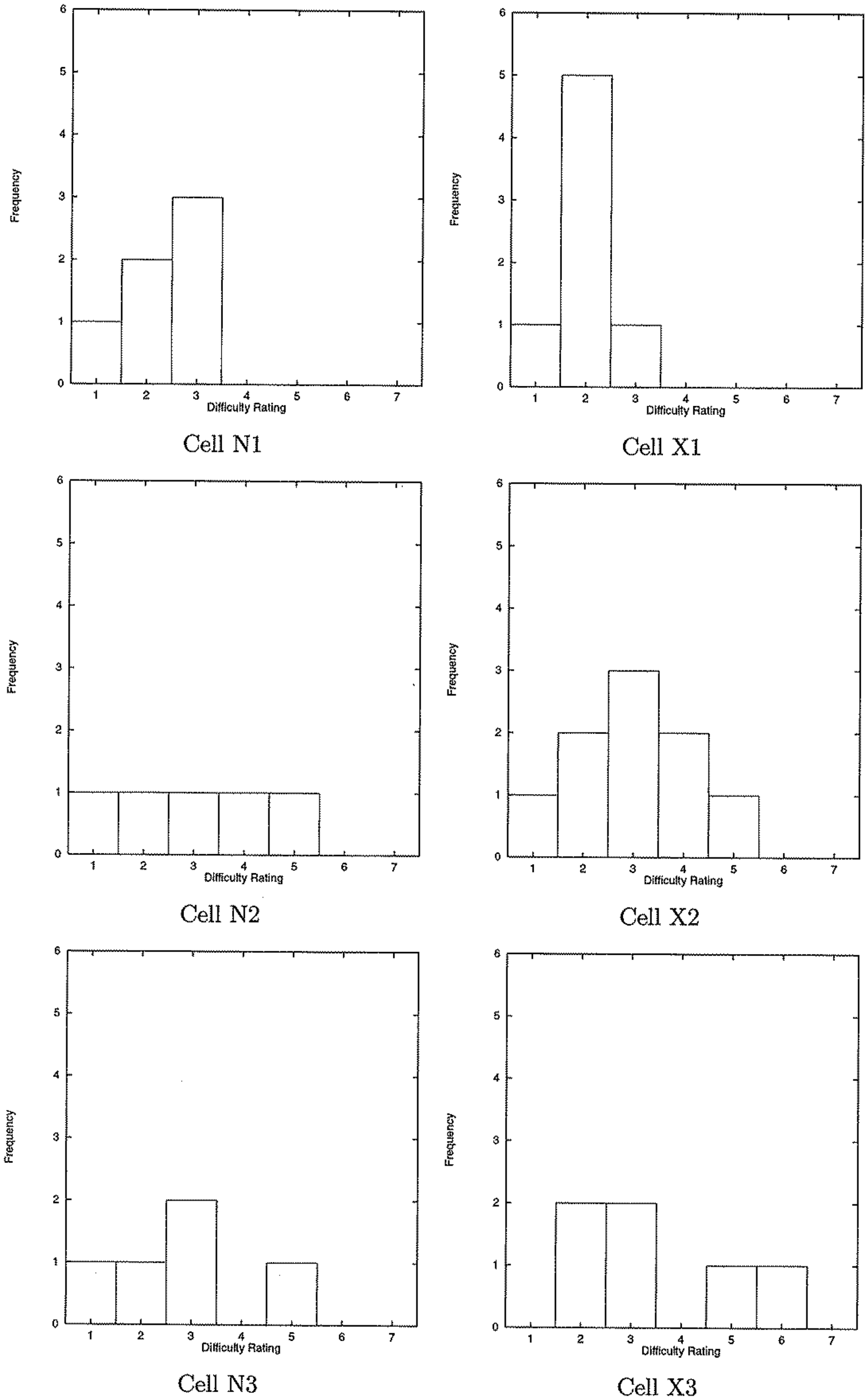
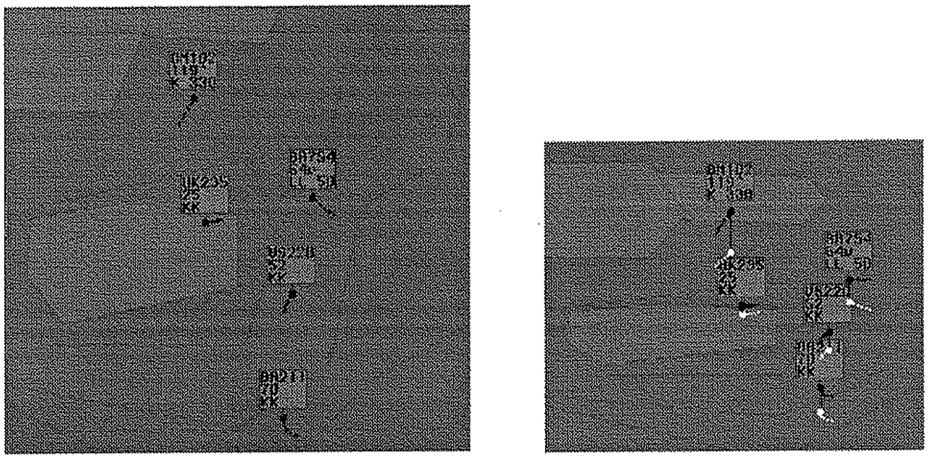


Figure 7.26: Conflict Detection Subjective Difficulty

5.991)). In all cases, the same two aircraft were falsely perceived to be in conflict (VS220 and BA754). Figures 7.27 (a) and (b) show the appropriate portions of the display in the 2D and pseudo-3D display formats respectively. VS220 (the aircraft



(a) 2D PPI Display (b) Pseudo-3D Display

Figure 7.27: Conflict Scenario 'A' Close-Up

further south) is level at 5 200 ft and bound for Gatwick, whereas BA754 (further north) is at 6 400 ft descending for 5 000 ft and bound for Heathrow. The vertical separation of the two aircraft will therefore be less than 1 000 ft at some point in the near future. This is still safe provided that the aircraft remain more than 3 n.m. apart laterally. At the time of Figure 7.27 (about 1 minute into the scenario) VS220 has just rolled out of a right turn; earlier in the scenario, it was heading due north, and its projected track would have intersected with that of BA754. However, its *speed* is lower than that of BA754, which means that VS220 would never have lost separation.

The author speculates that the situation is less clear with the 3D display (and hence may lead to more subjects incorrectly reporting a conflict between the two aircraft with the 3D format than with the 2D format) because both aircraft are heading approximately north, which is roughly along the LOS. This may have caused the distance between the two aircraft to have appeared to be smaller than it actually was, and because the trailing histories appear to be closer together, this may have partially masked the differences in the speed between the two aircraft.

True 'incorrect' responses were where subjects failed to detect an extant conflict. Scenario 'B' contains many responses in this category (17 out of 19 responses were cases where the subject failed to spot the conflict). This scenario involved an 'altitude bust' where an aircraft failed to level off at its cleared altitude and came into conflict with aircraft in a holding stack. Only constant monitoring of all traffic, ensuring that aircraft were doing what they were meant to be doing, would have revealed this conflict, and this suggests a tendency either to check some traffic at a higher priority than others (referred to in §2.3 as having traffic in the 'background' and 'foreground'—see for example Quotes G.13 and G.28) or to trust the cleared level (Quote G.40). Logistic regression analysis failed to find any significant effect of subject group on the number of incorrect responses in this case.

A common problem encountered was excessive clutter due to datablocks where a number of aircraft were in close proximity (e.g. Quotes G.63 and F.1). This particularly appeared to be a problem with scenario C, where there were several aircraft around the location of the conflict and where all the incorrect responses were cases of the conflict not being identified. The problem appeared to be worse in the 3D display cases than in the 2D displays (compare Figures 7.27 (a) and (b)) because of the greater amount of symbology in the 3D displays which could overlap with data blocks.

Conflict Resolution

§2.6.3 described the study of Ellis, McGreevy and Hitchcock in which pilots were found to carry out avoidance manœuvres 'somewhat earlier' and more frequently in the vertical dimension when using a perspective traffic display (Figure 2.4) than a plan view display. It was hypothesised that similar behaviour might occur when using a 3D display for conflict resolution. This was found not to be the case—the display format was not found to have any significant effect on conflict resolution instruction.

Conflict resolution is a much different task than pilot evasive manœuvres. Pilot evasive manœuvres are usually carried out to avert impending collision, when aircraft are already in close proximity. In contrast, potential conflicts are usually identified by controllers well before separation is lost. In resolving a conflict, the controller will often want to create the minimum amount of disruption to either aircraft, so if an aircraft is established on a course, the vertical dimension gives the controller the means to separate aircraft without causing them to deviate from their routes. For a climbing or descending aircraft crossing another level aircraft, it is therefore preferable to stop the climb or descent until the aircraft have crossed if further climb or descent will bring the aircraft into conflict. This may explain the dominance of vertical resolutions instructions in Figure 7.25. (The exception to this is scenario E which had two aircraft converging while flying in the same airway. Here the 'correct' resolution instruction was a horizontal one to parallel the tracks, since one aircraft's course would have eventually taken it out of the airway.)

Summary

The fundamental questions with which this study is concerned are how display and subject group influence this task.

Surprisingly, the behaviour of the expert and novice groups was found to be very similar. The only significant difference in quantitative results was that experts tended to identify conflicts sooner than novices. The methods they used, the classes of resolution instruction given and the number of 'correct' responses they gave were all similar. (However, it should be stressed that this was only a simple part-task concentrating purely on the detection of conflicts, and the real air traffic control task is a lot more complex.)

Contrary to the findings of Burnett and Barfield's experiment, there was no evidence that the 3D display aided conflict detection (and at least one subject made a note as such (Quote G.68)). There were significantly more 'incorrect' responses than 'correct' responses for the 3D display formats than for the 2D display, and no significant influences of display type on conflict detection time or conflict resolution instruction were found. Reasons for the latter have already been discussed above.

Possible reasons for the failure of 3D displays to show any significant performance benefits over 2D displays are:

1. The ambiguity of position along the display LOS inherent in a 3D display may lead to problems with detecting conflicts between two aircraft where one or both are travelling approximately along the LOS, as discussed above.

2. Excessive clutter. Subjects reported aircraft-related symbology being 'garbled' against the datablocks (e.g. Quotes G.56 and G.66), and that excessive dat-

ablock overlap increased workload (e.g. Quote G.49). The 3D display formats contained twice as much aircraft-related symbology (ground and air plots and trailing histories) as the 2D display format and so were more prone to this problem than the 2D displays. The display scale chosen was also very wide and the displayed area covered several 'sectors' (see Quote G.60), making the display particularly crowded in areas of activity.

3. There is evidence that subjects considered vertical and horizontal information separately when assessing conflicts, and when referring to horizontal position, they referred to ground position. The air position symbols would not therefore have been of much use.

The contradictory results of this and Burnett and Barfield's experiment show that the investigation of this area is not simple. A future study should perhaps repeat this experiment, but with an improved design of display less prone to the problems of label clutter.

7.8 Task 5: Chaser Task

7.8.1 Selection

The first stage of the task was to select the aircraft closest to the point on the ground at the centre of the displayed area. Dependent variables were correctness of selection and selection time.

Histograms of incorrect selections are shown in Figure 7.28. As can be seen, incorrect selections are only in stimuli C, D and E. Logistic regression analysis did not find any significant influence of display, group or the interaction term on the number of incorrect selections.

Selection times (for all selections, both correct and incorrect) are plotted in Figure 7.29. These show a gamma, rather than a normal distribution, and the analyses were carried out accordingly. Table C.36 shows descriptive statistics. Analysis by two-factor ANOVA (Table C.37) suggests that:

1. The 2D display gives faster selection than the 3D formats for the expert group, but not for the novice group.

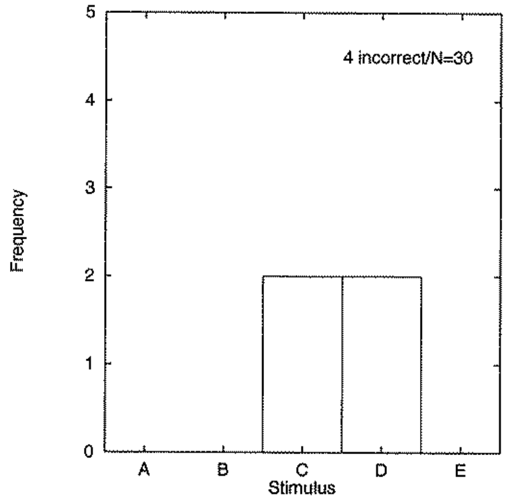
2. Experts are faster than novices for the 2D and pseudo-3D displays, but novices are faster than experts using the 3D stereo display.

7.8.2 Interception

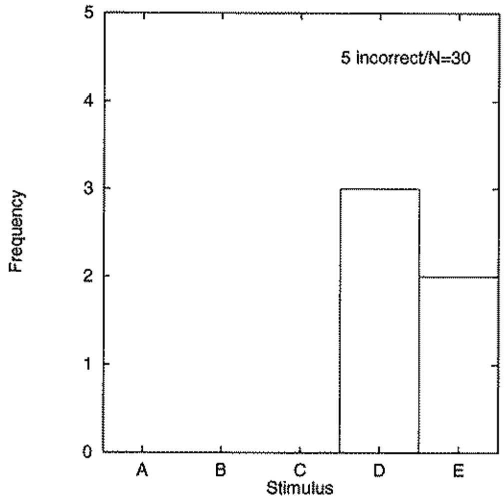
Time to intercept was recorded and divided by the initial distance from target to normalise the results in the case of an incorrect selection. Only three 'no-intercepts' were recorded, so it was expected that removing these would not significantly alter the results. Normalised interception times are plotted in Figure 7.30. Descriptive statistics are shown in Table C.38. Analysis by two-factor ANOVA (Table C.39) showed that the 2D PPI display gave the fastest interception times, with the stereo-3D display the second fastest and the pseudo-3D display the slowest (the means and standard deviations for observations for each display type are shown in Table 7.4).

$D = 1$	$D = 2$	$D = 3$
$2.1 \pm 1.7, 65$	$3.2 \pm 2.5, 68$	$2.8 \pm 1.9, 54$

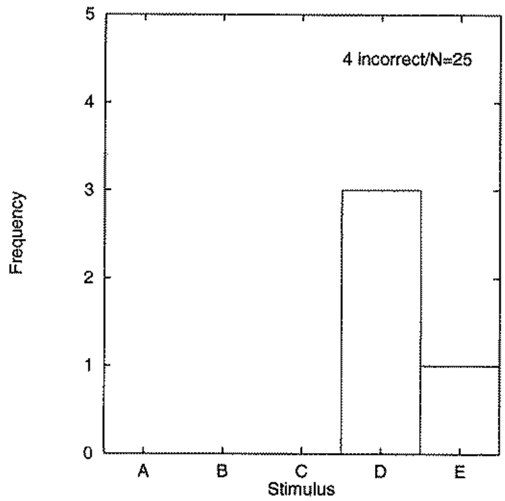
Table 7.4: Task 5 Normalised Interception Time: Summary by Display



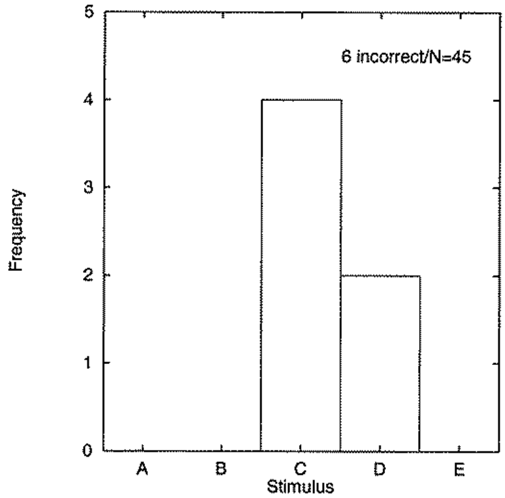
Cell N1



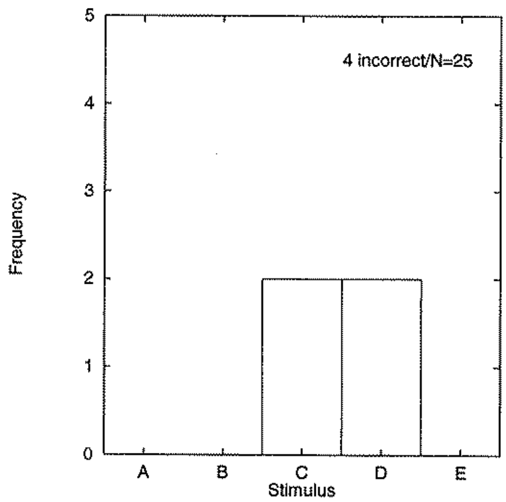
Cell X1



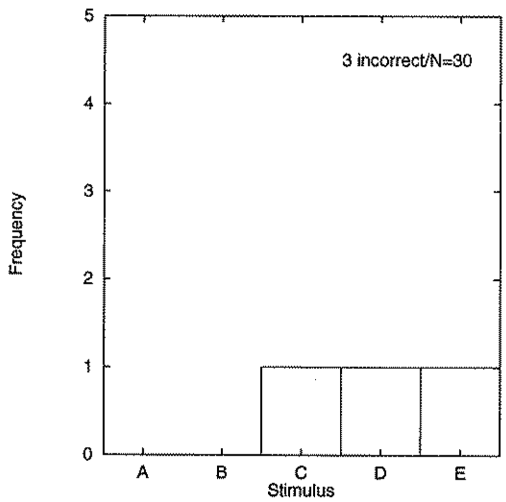
Cell N2



Cell X2



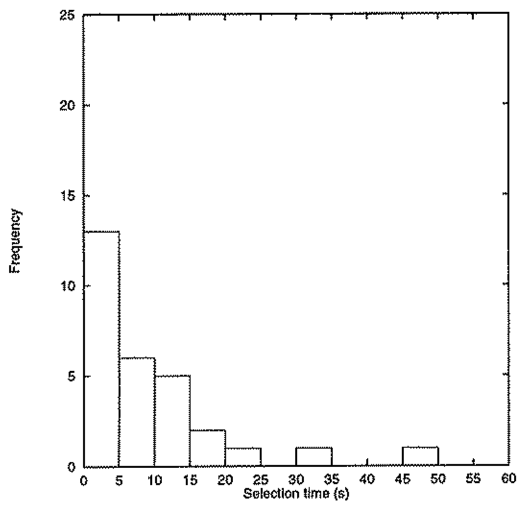
Cell N3



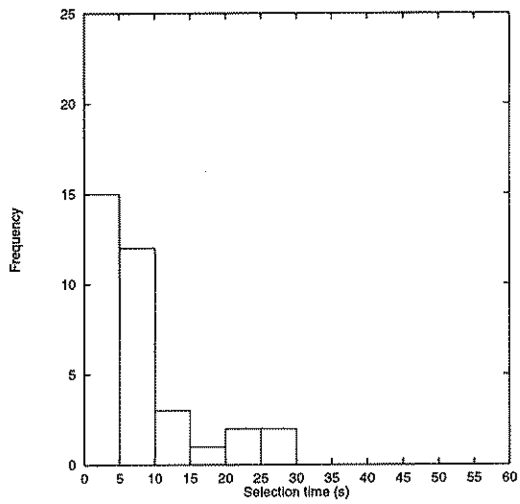
Cell X3

Number of incorrect selections out of total number of selections made are indicated on each histogram.

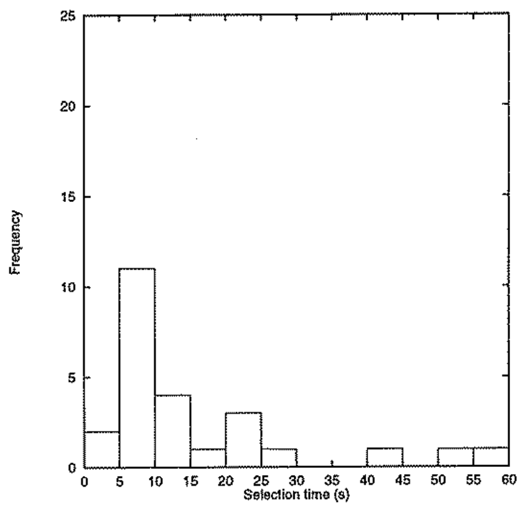
Figure 7.28: Task 5: Incorrect Target Selection Histograms



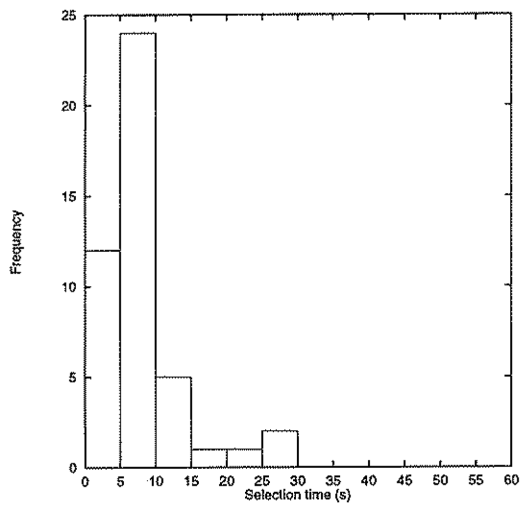
Cell N1



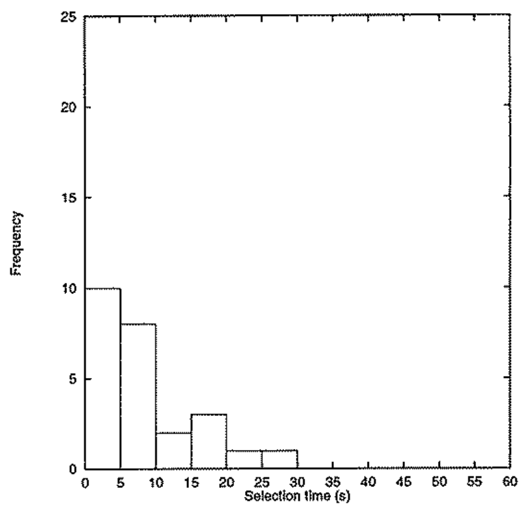
Cell X1



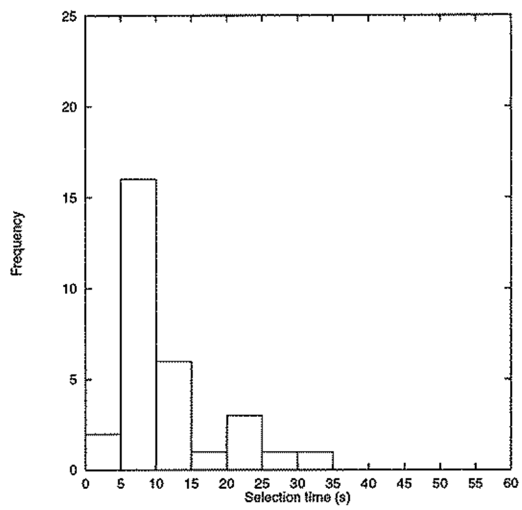
Cell N2



Cell X2

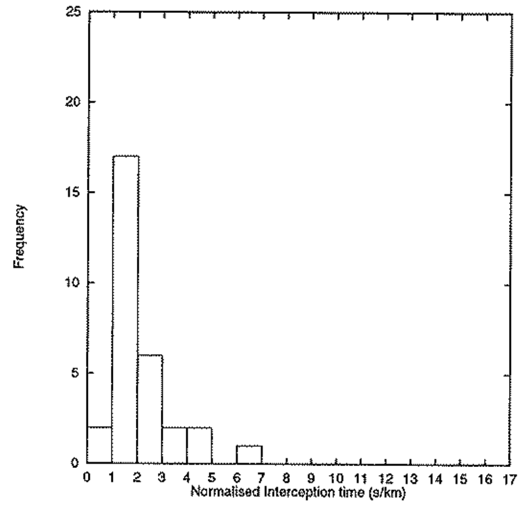


Cell N3

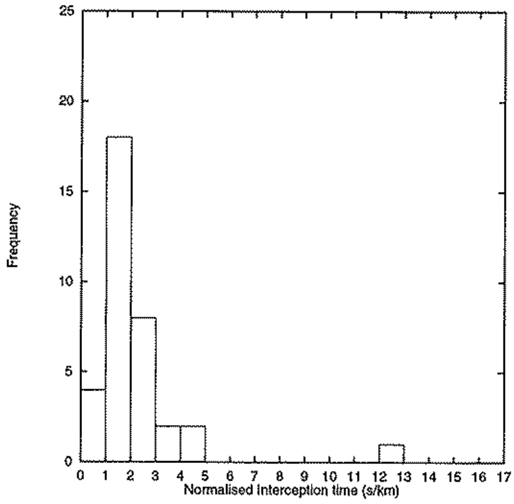


Cell X3

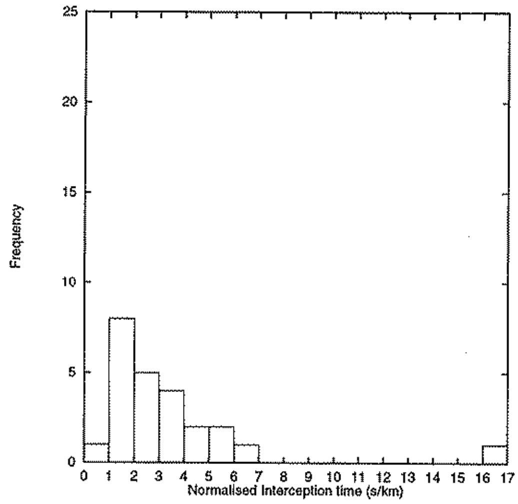
Figure 7.29: Task 5: Selection Time Histograms



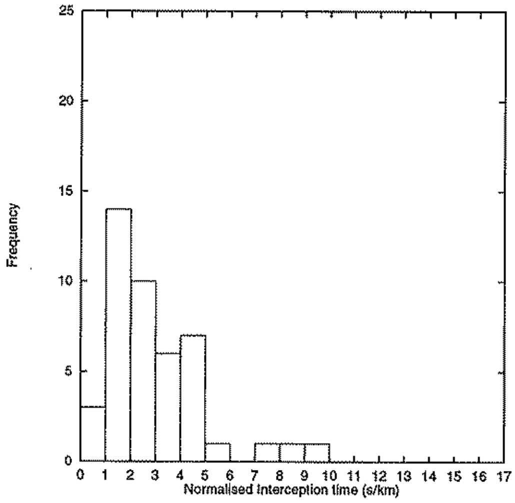
Cell N1



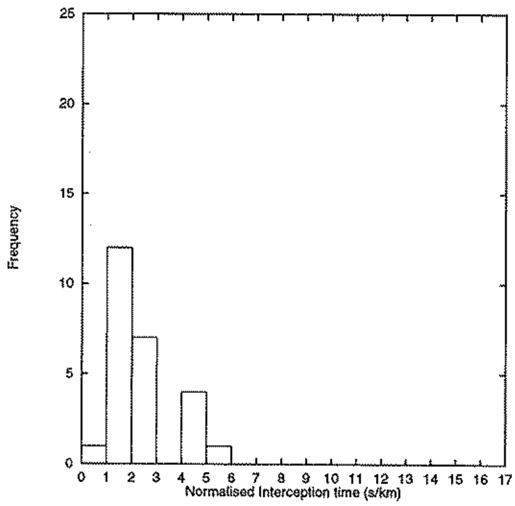
Cell X1



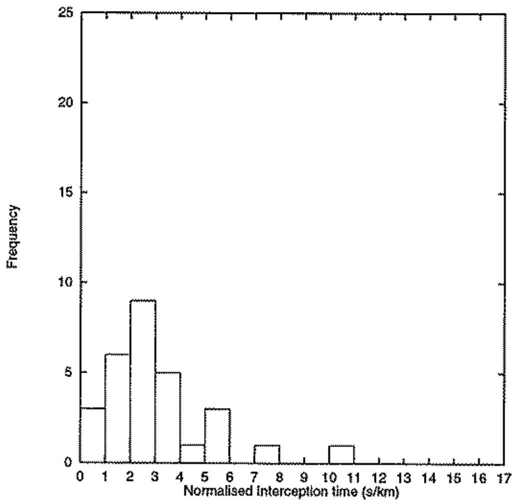
Cell N2



Cell X2



Cell N3



Cell X3

Figure 7.30: Task 5: Normalised Interception Time Histograms

Three sources of difficulty were observed in performing the interception task:

1. For both 2D and 3D, precise determination of height required the subject to use the mode-C readout in the datablock. However, datablock overlap caused a problem when the target and chaser were in close proximity; subjects had to click on a datablock in order to see obscured information, and this required them to move the cursor away from the control sliders, often at times where control was required (since the aircraft were in close proximity and adjustment of the chaser's trajectory was often required for the final phase of the intercept) (e.g. Quote H.30).

2. Two separate controls were provided for control in the vertical and horizontal planes, rather than an integrated control mechanism. This appeared to cause difficulty for some subjects, since chaser path control required co-ordinated use of both sliders.

3. Pitch control tended to be difficult:

   - The pitch slider was quite sensitive, controlling a large range of pitch angle (±90°) within a small travel (e.g. Quotes H.28 and H.34).
   - There was no 'dead-zone' around 0° pitch, which made levelling off the chaser quite difficult (e.g. Quote H.16).
   - There was a time lag built into the response of the chaser—it did not instantaneously assume a new command pitch angle but pitch change rate was limited (a similar rate restriction applied to heading as well) (e.g. Quote H.9).

   The upshot of this was that height control tended to require almost constant attention.

7.8.3 Subjective Difficulty

Subjects were asked to rate the difficulty of the conflict detection task on a discrete scale of 1–7. Histograms of the estimated task difficulty are shown in Figure 7.31. Descriptive statistics are shown in Table C.40. A two-factor ANOVA analysis examining the influence of display type and subject group found no significant influence from either of these factors or from the interaction between them.

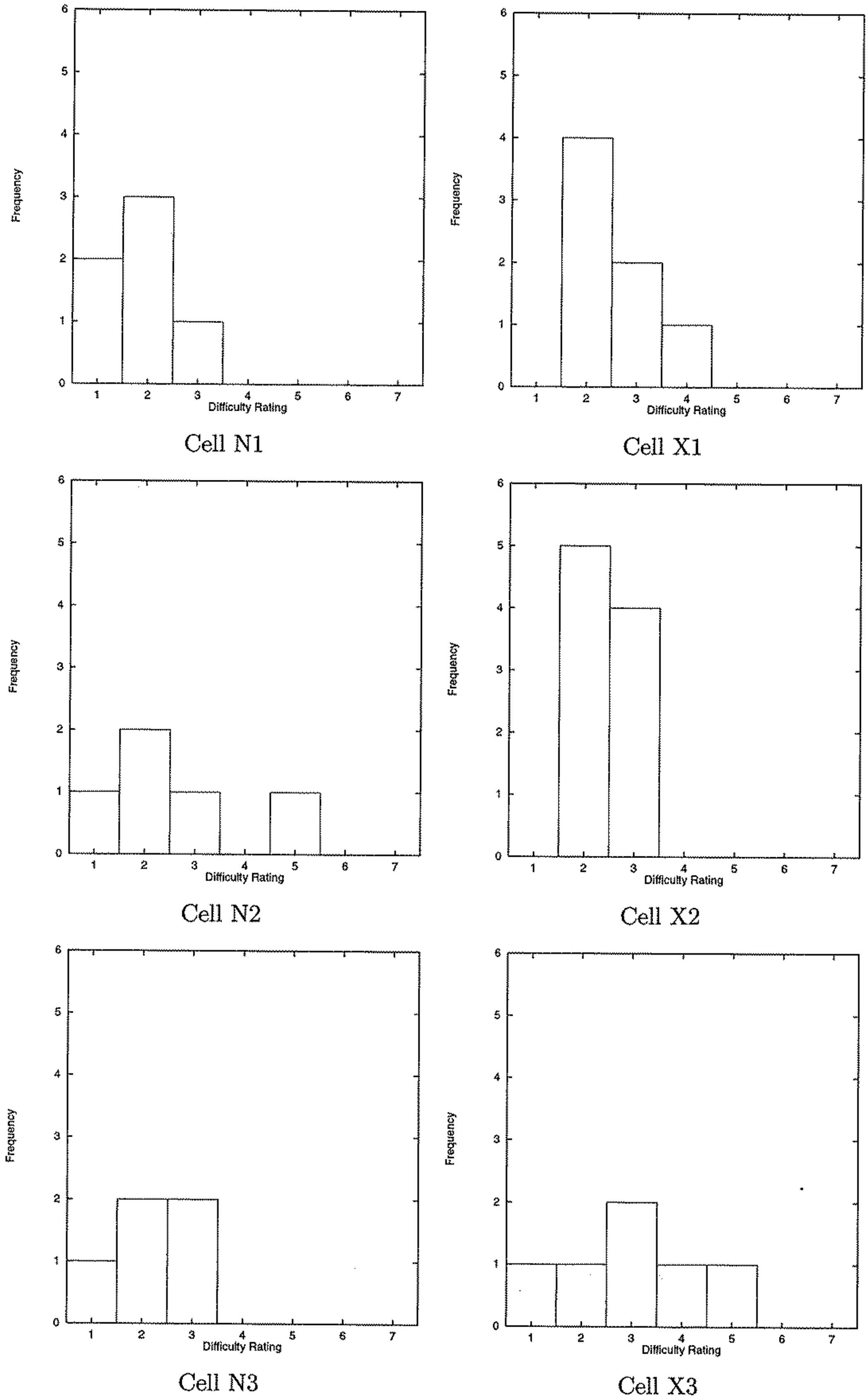


Figure 7.31: Chaser Task Subjective Difficulty

7.8.4 Discussion

Selection Times

Overall, it appears that the 2D display formats give faster selection times than the 3D display formats. However, the reason for the interaction between display and group (for example, why cells X3 and N2 had greater selection times than the other cells) is not known. Further, the number of selection errors was not significantly different between the experiment cells.

These results were contrary to expectation, and contrary to the results in Bemis, Leeds and Weiner's study (§2.6.3, p. 51). The task required aircraft to be selected by absolute, slant distances rather than by purely horizontal or vertical distances, and it was expected that since these distances are visualised directly on a 3D display, selection would be quicker than on a 2D display. There are possible reasons as to why this expected behaviour did not occur:

1. Some subjects had problems interpreting the instruction sheet; it was not clear to some that selection was to be in terms of slant, rather than horizontal, distance. This had to be explained verbally by the supervisor in several cases (sometimes by drawing diagrams), and it is possible that some subjects did not understand the instructions fully.

2. The scenarios did not sufficiently exercise judgment of slant distance, or the difference between the slant distance and horizontal distance was not great.

3. Aircraft selection difficulty due to equipment problems described in §7.5.1.

The discrepancy between the findings here and other earlier results remain to be investigated for future work.

Interception Times

It was expected that the 3D display formats would give faster interception times than the 2D format; however, the 2D format gave faster performance than the 3D formats. Unfortunately, the questionnaire did not explicitly ask how subjects performed the task. However, the separation of vertical and horizontal controls may have also caused subjects to consider the horizontal and vertical dimensions separately (especially since pitch control was a lot more difficult that heading control), partly negating any benefit of a 3D presentation. Some expert subjects still preferred using the datablocks for height extraction, even where the datablock obscuration was causing difficulty reading height and drop line information was available (e.g. Quote H.21 and H.30), although at least one expert subject reported using the drop lines (Quote H.36) and one belatedly realised that he could have used them, albeit with an accuracy penalty (Quote H.22).

7.9 Immersive Display Demonstration

7.9.1 Display Problems

As originally designed, the radar visualisation VR display showed a view of the same area as the 2D and 3D displays, using a texture map to implement the video map. Aircraft were represented by tetrahedra at the air position, standing on a 'post' which connected the air position with its plan position on the ground. Trailing histories were shown both in the air and on the ground. Datablock text was selected by the mechanism described in §6.2.3—the viewer was required to look at an aircraft to bring up its associated datablock, which was scaled and orientated every 0.1s such that it would always face the viewer and would appear at a constant size.

This was implemented and working before the main experiment. The only real problem was that 'pixellation' of the video map was obvious, and consequently the range rings and some airway boundaries were illegible. This could have been alleviated by mapping the video map image onto a larger number of smaller polygons, with a performance penalty. The datablock selection method was a little slow but worked effectively.

However, shortly before the main experiment was due to commence, the machine was upgraded to use a PixelPlanes II image generator. When the machine was returned, it was found that the application was 'broken' in subtle and mysterious ways. Running the same application code, the video map was now no longer displayed (the texture mapping mechanism had presumably changed), some of the 'posts' were afflicted with transparency (i.e. they were no longer solid-shaded but transparent) and the datablock selection performance was now totally inadequate—the time delay between looking at an aircraft and the datablock being selected, and the time to display re-orientated and re-scaled text objects, were unacceptably long. Further, the backlighting on one 'eye' on the HMD had failed, so that the image presented to one eye was dimmer than that presented to the other eye (this problem was not fixed until the last two subjects).

There was insufficient time to fix these problems before the main experiment, and some of the problems could not be fixed despite help from the manufacturer's technical support. The system therefore had to be demonstrated with these problems.

7.9.2 Subject Opinions

The VR display was demonstrated to air traffic controllers. As described in §6.4, this involved three parts: A demonstration of a kitchen visualisation, an air traffic visualisation and verbal description of a possible application. Subjects were asked to note their comments, and these are shown in Appendix I.

General Comments

Subjects commented on the technical problems with the display: Weight, poor resolution, update rate, and difficulties with 'navigating' in the virtual environment (see comments in §I.1). For operational application in ATC, these comments indicate that current technology would have to be a lot more mature than at present.

Radar Visualisation

Subject comments are given in §I.2. Immersive displays were universally rejected for radar control work. There was evidence that subjects found the sensation of depth greater than in either of the 3D TTW displays in that vertical information seemed to be much more apparent (Quotes I.17, I.18 and I.20). However, reasons for not using an immersive display for radar control work included:

- An exocentric viewpoint is required, where the controller is outside the scene and able to see everything, otherwise there is the danger of missing something outside the FOV (Quote I.13).
- One controller cited the need for a fixed image against which to assess location and movement for conflict detection (Quote I.24). Another cited the changing viewpoint as potentially disadvantageous (Quote I.25).
- Potential for disorientation (Quotes I.16 and I.27).
- No obvious benefits over present equipment (e.g. Quotes I.14, I.15, I.22 &c.).

Other considerations include:

- Physical and physiological effects of prolonged use of immersive equipment are not known (Quote I.23).

- Air traffic control is a highly cooperative system, where controllers have to operate as a team, and another point raised was the need to cooperate with other controllers, rather than to be encased in one's own 'world' (Quote I.11).

- The controller would still require access the various tabular displays (e.g. flight strips, maps, weather reports) and controls (e.g. R/T channel selection), and these would have to be provided in an immersive environment.

The reaction to the immersive radar visualisation was therefore negative.

Kitchen Visualisation

Subject comments are given in §I.3. Responses to the kitchen visualisation (to demonstrate potential VCR applications!) were a lot more positive than for the radar visualisation, although with reservation. Subjects mostly saw applications in VCR simulation for limited training and potentially for the design of new VCRs, especially assessing lines of sight.

The author has visited two companies specialising in VCR training equipment. Companies such as these can already provide full VCR simulators which are akin to flight simulators, using real furniture within the simulated VCR and television projection to reproduce the scene out of the window, for high fidelity but at a very high cost. More limited computer-based training aids, displaying a more limited view on one or two monitors, are also available. Limitations of VR compared to such simulators and training aids include the lack of ability to write on flight strips (quote I.42), poor resolution and cumbersome technology, but VR certainly has the potential to be more cost-effective and flexible than a full VCR simulator in some part-task training applications, if not for full-task VCR simulation.

Subjects generally liked the idea of using VR for applications such as modelling the new Heathrow VCR. Some had already seen a computer animation video of proposed layouts, but these are not interactive in any sense. Controllers liked the possibility of using a VR simulation to assess lines of sight (which was raised many times as an area of concern), layout of furniture, access etc. It is postulated that such things would be easier to assess in an immersive visualisation than in a TTW visualisation, and this would allow for effective prototyping and development.

Proposed Application

A possible future operational application of VR, described in §6.4, was verbally presented to subjects. Their comments are given in §I.4. To recap, the proposal was for a lightweight see-through HMD which would allow aircraft and other symbology to be superimposed on the real world, allowing VCR controllers to stay 'head-up' in low visibility conditions, such as at night or in fog. It was suggested as a possible future application since the current technology is not sufficiently mature.

Subjects generally saw this as a potentially useful application, but only if the technology was sufficiently mature.

7.10 Summary of Results

The main aim of these experiments was to examine the effects of display type (2D, pseudo-3D and stereo-3D) and subject group (ATCO or non-ATCO) on subject performance over a number of tasks. The results of these experiments are summarised below and discussed further in the next chapter.

The first task examined the readings of azimuth angle and horizontal distance between two objects. It had been anticipated that since ATCOs are trained on a 2D display, they would read these parameters to a greater accuracy than untrained individuals. It was further anticipated that perceptual biases inherent in 3D displays would render them less accurate compared to a 2D display. The results bore out the latter, but not the former hypothesis: Subject group did not significantly affect accuracy of reading either parameter. This may have been due to behavioural differences between the two subject groups; there was evidence that ATCOs attempted to read the display only to a limited accuracy, while novices seemed to take more care over their readings and attempted to read the parameters to as great an accuracy as possible, but only performed as well as the ATCOs. (However, the subjects' self-estimated accuracy was found not to reflect true performance for either subject group.)

The second task examined the speed of selecting the highest and lowest aircraft, and the aircraft in closest horizontal proximity. Regarding the altitude extraction task, it was anticipated that subjects would select the highest and lowest aircraft in less time using the 3D displays than the 2D format. While this turned out to be the case for novice subjects, display type was not found to influence significantly the response time of expert subjects. Further, it was anticipated that expert subjects would be faster than novices, but it was found that novices performed no slower than experts on average, and indeed were significantly faster in some cases. The observed differences between the two groups may again be due to behavioural differences; there is evidence that the ATCOs were predominantly using the datablocks rather than exploiting the 3D visualisation, either out of accuracy concerns or force of habit, and ATCOs may have emphasised accuracy rather than speed. Regarding the horizontal proximity extraction task, the results indicated that the 2D display format gave faster response times than the 3D formats, but that there were no differences between the performance of the two subject groups.

The third task examined how subjects recalled a static air traffic scenario. All subjects found this task to be very difficult, but the results suggest that familiarity with the airspace and traffic patterns made the task easier where these could be exploited. The results also suggest recall of ground position may be more accurate when the stimulus is 2D rather than 3D, but the fact that the scenario had to be reconstructed as a plan view regardless of the format of the stimulus image may have biased the results.

The fourth task looked at the detection and resolution of conflicts. Display type was not found to have an effect on how quickly conflicts were detected, but subjects using the 2D display were found to make significantly more correct responses than those using the 3D display formats. The way in which conflicts were resolved was not found to depend on display type or subject group.

The final task was an interception task. Performance measures were: Time to select a 'threat' aircraft based on a horizontal proximity criterion, and time to intercept. In both performance measures, the 2D display was found to give smaller times than the 3D formats; however, there may have been factors biasing performance in favour of the 2D display and further investigation is necessary.

Chapter 8
Conclusions

8.1 Overview

This chapter starts by summarising the work of this thesis: What it aimed to address and how it was addressed; problems encountered, particularly sources of error; the conclusions; and its contributions. Future work is then discussed.

8.2 Introduction

This thesis described an empirical exploration comparing 2D and 3D displays of object position in three-dimensional space for tasks involving control over the positions and paths of those objects, particularly for air traffic control.

Currently, ATC uses a 2D plan-view display of air traffic with height displayed numerically. Such a non-integrated display of spatial information imposes a cognitive workload in requiring the viewer to integrate the information mentally to appreciate the three-dimensional disposition of the aircraft in space. A 3D display may present a single, integrated pictorial representation of all spatial dimensions which potentially requires less cognitive effort to interpret and may present spatial relationships more clearly. Tasks such as ATC are safety-critical, real-time tasks; potentially critical decisions may have to be made 'on the spot', and a high workload may degrade the quality of those decisions. Three-dimensional displays can potentially offer a reduced workload thereby allowing the controller to spend more time on other aspects of the task or to handle more traffic.

However, 3D displays have associated penalties. In choosing between 2D and 3D displays, Wickens and Todd [WT90] cite two main research domains which must be considered:

1. 3D display research: Two key factors are the costs of position ambiguity along the display line of sight and the inherent distortion of distance judgments along axes which are not parallel with the viewing plane. If tasks such as

ATC require such judgments to be made with precision, then a 3D rendering may not be suitable.

2. The proximity compatibility principle: "Tasks of a more integrative nature involving (for example) the comparison between data points will benefit from more 'object-like' displays, whereas tasks requiring the focus of attention on a single dimension or single object will be better served by more separated bargraph or digital displays".

This thesis explored some of these issues by conducting experiments which compared subject performance using two-dimensional and two types of three-dimensional display (pseudo-3D and stereo-3D) across a number of tasks which were aimed at important elements of using displays for spatial command/control tasks. Two groups of subjects were used, air traffic controllers and students, as it was felt that ATCOs would be biased towards the 2D display. The tasks and display formats used in this thesis built upon previous research by a number of others, notably Strutt, Burnett & Barfield, Ellis *et al.* and Wickens *et al.* It was expected that results would generally support the findings of this previous research, as well as extend it to areas not previously covered (principally, a comparison of 2D and 3D display formats for ATC involving a stereoscopic 3D display and an investigation of differences between non-air traffic controllers and ATCOs). In addition, an immersive (virtual reality) display was demonstrated to gain opinions regarding possible future application of this technology to ATC.

8.3 Summary of Main Findings

8.3.1 The Effect of Display Format

The illusion of depth in a 3D display is modelled as a weighted additive function of the depth cues employed. Research suggests that in relatively static displays, stereopsis is a particularly salient depth cue. Significant differences were therefore expected to be found between subject performance using the stereoscopic and pseudo-3D display formats, with the stereoscopic display expected to give 'better' performance than the pseudo-3D display (where the definition of 'better' varies with the task).

Distance, Angle and Height

Because 3D displays suffer from perceptual biases in reading of azimuth angle, and distort distance judgments along axes not perpendicular to the LOS, it was expected that for tasks requiring azimuth angle and horizontal distance to be interpreted, the 2D display would be more accurate than a 3D display. However, for tasks requiring rapid interpretation of vertical information it was expected that subjects viewing a 3D presentation would respond quicker than subjects using a 2D display.

Linear perspective is an important depth cue, but since the displayed visual angle subtended by an object of constant objective size varies depending on its depth, this raised the question of whether or not a display incorporating linear perspective would be suitable for ATC, where operators are required to judge horizontal distances, azimuth angles and relative heights between aircraft. It was found that subjects made more accurate observations of azimuth angle when these were displayed using a 3D parallel projection than with a perspective projection but surprisingly, the accuracy of observation of relative horizontal distances was *not* significantly affected by the type of projection. As a result of this finding, a parallel projection was adopted for the 3D display formats for subsequent experiments.

Observations of azimuth angle and relative horizontal distance between two targets were then compared between the 2D, pseudo-3D and stereo-3D displays. For both horizontal distance and azimuth angles, it was found that the greatest accuracy resulted from the 2D display, with the stereo-3D display the next most accurate and the pseudo-3D display the least accurate. This was consistent with other research, notably on the effects of 3D perspective viewing on the exocentric judgment of azimuth angle by Ellis *et al.* Interactions from other variables (such as the position of the targets in the display) were also found.

The speed of extraction of altitude and horizontal proximity information was also examined. For horizontal proximity extraction, the results were as expected with the 2D display giving the shortest extraction times and the stereo-3D display being faster than the pseudo-3D display. It was expected that altitude extraction speed using the 3D formats would be either the same as or significantly faster than for the 2D format (as found by Burnett & Barfield). However, no significant differences were found between the display types for the expert group. It was speculated that expert subjects may have been predominantly using the numeric height information, even in the 3D scenarios, either due to their training or out of concerns for accuracy. The results for the novice group were equally unexpected: Fastest was the stereo-3D display, next fastest was the 2D display and slowest was the pseudo-3D display. One explanation of these results is that inadequate depth cues and/or poor design in the pseudo-3D display might make it prone to misinterpretation (see §5.5.3), but the addition of stereopsis might alleviate this: It is therefore possible that inadequate depth cues caused the pseudo-3D display to exhibit worse performance than the 2D display, whereas a stereo-3D display had adequate depth cues and so performed better than the 2D display as expected.

Memory Recall

A memory recall task suggested that the accuracy of recall of aircraft ground position was higher when subjects were required to recall a scenario presented in 2D than when subjects were required to recall the same scenario presented in 3D. Unlike a similar experiment by Burnett & Barfield, no consistent bias in recalled position was found for the 3D display formats. However, the memory task tried to test the memory of a large number of different entities (position, height and callsign) simultaneously and so overloaded the subjects and subsequently compromised recall of these entities. This made it very difficult to determine the effect of display type on the recall any of these entities in isolation, and tasks exploring these issues in future should be better targeted (e.g. designed to concentrate on the effects of display type on each of these entities in isolation).

Conflict Detection and Interception

All the above tasks required subjects to read parameters from a static scene. Two tasks were presented which involved interpretation of dynamic rather than static scenarios: A conflict detection task and an interception (chaser) task. The latter also required subjects actively to control an interceptor, where the former merely required passive response.

In the conflict detection task, subjects using the 2D display correctly identified significantly more conflicts than those using either of the 3D displays (with the stereo-3D display giving more correct identifications than the pseudo-3D display). Two reasons were postulated for this: Position ambiguities in the 3D display resulting in aircraft appearing to be in conflict where they were not (or potentially *vice versa*), and excessive clutter. However, the latter may be largely a feature of the display design, rather than something inherent in the three-dimensional nature of the display, and the experiment should be repeated with a display design less prone to clutter.

The times in which the conflicts were detected were not found to be significantly influenced by display type, contrary to the findings of an experiment by Burnett & Barfield. This may be partly due to the 3D displays being more prone to clutter, but also due to the way in which subjects detected conflicts. There is evidence to suggest that subjects tend to examine the horizontal and vertical dimensions separately when looking for conflicts, in a 3D display looking at ground position and height and largely ignoring the air positions. This would tend to negate the advantages of a 3D presentation. The type of instruction given by subjects to resolve the conflict was also not found to be influenced by display type.

The chaser task required subjects to select the closest target to a point, using judgment of absolute distance rather than simply considering horizontal and vertical components, and then to guide a 'chaser' to intercept the chosen target. The target selection had been expected to support the findings of a similar experiment by Bemis, Leeds and Wiener, which showed greater accuracy and speed in selecting a 'threat' using a 3D display than a plan view display. However, in this experiment, the results of speed and accuracy of selection of the interceptor were inconclusive for no clear reason (although a number of reasons related to problems with the experiment are postulated in §7.8.4).

It was expected that interception times would be faster for the 3D displays than the 2D displays, but it transpired that the 2D display gave faster interception times than either of the 3D display formats. This may again have been for reasons to do with the experiment itself rather than the display formats: Control of the 'chaser' in the vertical was rather more sensitive than control in the horizontal, and the horizontal and vertical controls were separate, which may have encouraged subjects to consider the two dimensions separately, partly negating the benefit of the 3D format. The stereo-3D display gave quicker interception times than the pseudo-3D display, as anticipated.

8.3.2 The Effect of Subject Group

Because ATCOs are highly trained in the use of the 2D plan-position indicator display for air traffic control, it was expected that their performance at tasks using a 2D display would be better than for novices using the same display format. However, it was not known whether the same differences would be observed between the novice and expert groups when presented with the same tasks using 3D displays. If expert performance was significantly better than novices using 3D displays, this would suggest that their training and experience using 2D displays would also partly carry over to the 3D display.

Surprisingly, analysis of observations of relative horizontal distance and azimuth angle revealed no significant differences in accuracy between the performance of ATCOs and non-ATCOs. However, evidence suggested that whilst novices were attempting to read the parameters as accurately as possible, ATCOs were only attempting to read the display to an accuracy which they either thought was realistically achievable or was sufficient for their routine tasks. Subject group was not found significantly to influence performance at speed of horizontal proximity extraction.

Based on subject comments as well as on quantitative results, differences between the two groups did appear in some tasks where familiarity with air traffic control in the displayed area was advantageous. ATCOs were found to detect conflicts significantly more quickly than novices (although there were no differences between the two groups regarding the correct identification of conflicts) and reported using familiarity with traffic patterns in their memory of scenarios. Differences also appeared in tasks requiring extraction of vertical information, where evidence suggests that ATCOs tend to use the numeric height information in the datablock, even in a 3D display where a visual indication of height is given. Reasons cited included force of habit and concerns over accuracy.

8.3.3 Discussion

Returning to the research domains discussed in §8.2 above, three of the experiment tasks were based on and aimed at verifying the findings of 3D display research (namely, observation of azimuth angle and horizontal distance, speed of extraction of horizontal and vertical information and memory of a scenario), whereas two were aimed more at exploring the issues raised by the proximity compatibility principle (namely, conflict detection and interception).

The tasks exploring human cognitive aspects when interpreting 3D displays had findings much as predicted: 3D displays are less accurate than a plan view display when it comes to precise reading of parameters which lie on axes not coplanar with the viewing plane. In other words 3D displays convey horizontal information (horizontal distance and azimuth angle) less accurately than 2D displays; they may convey approximate height information more quickly than a 2D plan-view display (since height is represented graphically rather than numerically), but that height information cannot be determined accurately without resorting to numerical readouts.

The key question from the point of view of 3D display research is therefore, as Wickens and Todd state, whether or not tasks such as ATC require such judgments to be made with precision greater than that afforded by a 3D display. If the answer is 'no', then 3D displays may be rejected for operational radar control use *within the current operational framework* (since, of course, alternative procedures and control methods could be found where the drawbacks of 3D displays present less of a problem). If the answer is 'yes, with penalties', then one must weigh the penalties carefully—improved precision in reading angles and distances could be afforded for example through training, symbolic enhancements, measuring 'tools' (e.g. an electronic measuring tape giving the distance between two highlighted aircraft) etc. but these have the penalties of expense, display clutter, workload etc. Even on a 2D display, some ATCOs use a 'calibrated pen' or 'calibrated thumb' to read off approximate distances, for example on tracks involving several turns, but it is doubtful that a 3D computerised measuring stick would be so convenient or rapid in use (especially since electronic equivalents exist in some 2D equipment but ATCOs still seem to prefer a 'calibrated pen' instead).

The problem with tasks exploring purely cognitive issues is that they can only provide useful information about certain aspects of tasks taken in *isolation*. However, they do not adequately address whether or not the characteristics of a given display make it suitable for a given task *as a whole*. The aim of the conflict detection and interception tasks was to have subjects perform a range of cognitive actions together in carrying out a task rather than individually. The conflict detection task also asked how subjects performed the task to see whether the task was of an integrative nature or required the focus of attention on parameters singly. The results suggested that although the conflict detection task is integrative in some ways, subjects solving the problem tended to consider the horizontal and vertical separately. This was not just observed in ATCOs, who may solve problems in this manner because of current training, procedures etc.; it was also observed in non-ATCOs who were performing the task for the first time. This raises an important question about the usefulness of 3D displays for radar control: If the task requires the operator to consider the dimensions separately, then this may partly negate the benefits of a 3D display. It also raises still further issues: Did subjects perform the task in this way because of inadequate representation of depth in the 3D presentation, because of the way the problem was defined, or because of some preconceptions? If subjects were trained to perform the task in a more integrated fashion exploiting the characteristics of a 3D display, would they then perform better using a 3D display than a skilled operator would using a 2D display?

8.4 Potential Sources of Error

8.4.1 Introduction

The experimental methodology and results are here critically reviewed with regards to the validity of results and lessons which could be learned for future experiments. This section discusses possible sources of error. These may account for some of the differences between the findings of this study and other studies, show caveats in interpreting the results of this study and highlight areas that should be accounted for in future experimental work.

8.4.2 Subject-Related Sources

The following sources of error are variables that were not controlled, largely due to the method by which the subjects were obtained (i.e. on a volunteer basis).

Gender Differences

As shown in Table 7.2, both novice and expert groups of subjects were predominantly male. Evidence has been found suggesting differences in spatial ability between individuals due to gender [MJ89], but this could not be taken into account in this study due to the small number of female volunteers.

Subject Computer Skills

The experiment necessitated the use of a computer, and required subjects to use a mouse. One possible source of error is the subjects' levels of confidence with computers and expertise in using a mouse.

An attempt was made to take this into account in the entry questionnaire with a question "Have you used a computer mouse or trackerball before?" (§B.2, question 13). It was assumed that use of a mouse is a skill at which it is easy to acquire proficiency, with the corollaries that (a) all responding 'yes' to the question would be proficient, and (b) that it would be a simple matter to bring non-proficient subjects up to a comparable level of proficiency through the training sessions for the tasks. The question also assumed that proficiency with mice implied proficiency with trackerballs (a commonly-used input device in ATC) and *vice versa.*

Informal observation by the experimenter revealed that these assumptions were incorrect, and that subjects varied in the confidence and speed with which they used a mouse. Unfortunately the question "Have you used a computer mouse of trackerball before?" did not measure the level of proficiency of the subject.

The experimenter noted that novice volunteers were largely individuals who were of a 'technical bent'. All except two were from the faculties of science or engineering; one of the exceptions was a lecturer from the medical faculty, the other was a teacher of English as a Foreign Language. The air traffic controllers seemed to the experimenter to contain a greater mix of individuals: Some mentioned that they had attended particularly because they were 'interested in computers' and the applications of computer technology, others said that they were not 'computer literate'.

This suggests that the novices as a group may have had a higher level 'computer literacy' with general-purpose computers and greater proficiency at using mice[1], and this may have influenced the relative performance between the two groups.

Occasional difficulties were also observed with some subjects possibly mis-operating the buttons of the three-button mice used in the experiment, either by pressing the wrong button or by pressing more than one button at once. This may have been

[1] Although the ATC system is highly computerised, the author has observed that being an ATCO does not imply 'computer literacy', possibly since the computers are special-purpose rather than general-purpose. One ATCO told the author that he knew "nothing about computers" even though it was pointed out that he operated a computerised system every day.

due to the individuals being either unfamiliar with mice, or used to single-button or two-button mice. Again, the training sessions prior to each task were aimed at reducing or eliminating such effects, but it seems that learned behaviour may be more difficult to modify than was anticipated.

Air Traffic Control Background

The expert subjects had differing levels of operational experience, with some having extensive radar control experience, others having primarily ground control experience with limited work on radar. An attempt was made to measure operational experience in the entry questionnaire (§B.2, questions 8 and 9), but it was felt that correlating operational experience and performance would not be meaningful for two reasons.

First, radar control is not a standard task but varies according to the demands of the particular task—for example, en-route control is different from approach control, and even within the area of approach control there are several different radar controllers with different tasks handling traffic arriving at Heathrow.

Second, there is also a question of how recently subjects had been radar operators. Some pilot subjects who had been radar controllers during their operational careers stated that they were 'rusty', which implies that the recency of radar experience as well as the total length of experience influence expertise. The main experiment subjects had a variety of backgrounds: All were from the control tower at Heathrow airport, and so all were operating in a variety of non-radar control positions in the visual control room (ground movement planning, ground movement control, arrivals, departures), but a few were also current radar operators, additionally operating in the London Special VFR and Thames Radar radar control positions. One controller was on his first assignment and his only radar experience had been in training.

These factors could introduce variations in the performance of expert subjects.

Subject Vision Differences

Subjects using the stereoscopic display did not have their stereoacuity measured. Variation in stereoacuity may give different levels of stereopsis (depth perception due to retinal disparity) between the subjects using the stereoscopic display. It is recommended that in future experiments, stereoacuity is measured so that it can be accounted for.

All expert subjects must have normal colour vision due to medical requirements of their profession, but no such restriction was placed on the novice subjects. One novice subject reported being colour-blind.

8.4.3 Display-Related Sources

The following are sources of error related to the display; either the implementation of the display itself, or the viewing of the display.

Stereoscopic Display

One of the reasons that a stereoscopic display has not been used in a comparative study in ATC before may be the complications peculiar to such displays. In this experiment, a time-multiplexed system based on LCD shutter glasses was used to present the disparate left/right eye images to the subjects. This particular implementation had certain limitations, and introduced differences between the stereo-3D and pseudo-3D displays which were not just confined to the absence or presence of stereopsis.

1. Due to the way in which stereo was supported by the Silicon Graphics workstation, the vertical resolution of the stereo-3D display was half that of the pseudo-3D display, although the horizontal resolution remained the same. In particular, this changed the aspect ratio of datablocks and their text, and increased the level of clutter over that of the pseudo-3D display.

2. Since the LCD glasses are blanked (made opaque) periodically, they significantly block the transmission of light, with the result that subjective display brightness was lower for the stereo display than for the pseudo-3D display.

3. The level of 'ghosting' or 'cross-talk' between the left and right eyes (due to the LCDs not being fully opaque when blanked) was not measured.

4. There is a slight flicker viewing the stereo display due to the LCD glasses being switched at 60 Hz.

Control of Eyepoint

Subject eyepoint was controlled only to the extent that the display monitor was placed at a fixed position on a table, and subjects were seated at the table at the same place. However, subjects were free to move the chair forward or backwards and to adjust its height for personal comfort. This may be partly responsible for an observed variation in azimuth angle estimate with target position, as has already been discussed in §7.4.3.

This brings into question how tightly viewer eyepoint should be controlled, both for experiments and in practical implementations of 3D displays. Significant variations in eyepoint can be expected to affect perception of the display, but tight control of eyepoint (for example, head clamps) may result in discomfort on the part of the subject and would be completely impractical in an operational environment.

In the stereoscopic display, no attempt was made to account for different subject inter-pupillary distances. This should be accounted for in future experiments.

8.4.4 Experiment Implementation-Related Sources

Some implementation-related sources of error identified are listed below.

1. Selection was using a mouse on the air plots on the 3D display. Despite the training, some subjects forgot this and initially tried to select the ground plots instead. Also in a 2D display, if a subject has correctly identified an aircraft by its ground target, he/she then had to click on the corresponding air plot. These would tend to increase the time required for selection of aircraft in a 3D display format compared to a 2D display format.

2. Some subjects reported the fact that the mouse sometimes kept sticking (despite being cleaned prior to each experiment) and this sometimes made it quite difficult to select an aircraft, even if they identified it in good time.

3. The relatively small size of the aircraft symbols (and thus the area sensitive to mouse clicks) made selection difficult, especially with a sticking mouse.

4. In the memory recall experiment, subjects had to reconstruct the radar screens on a 2D piece of paper, even if the stimulus image had been in 3D. This could have influenced recall times.

8.4.5 Implications

This research has shown some of the difficulties and pitfalls in implementing three-dimensional displays in general and stereoscopic displays in particular. Implementing stereopsis gives a range of additional considerations which must be taken into account: The effects of cross-talk between the images, the variation of stereoacuity and inter-pupillary distances between subjects, and other issues related to implementation (such as flicker, reduced subjective display brightness, etc.). In experiments comparing 2D and 3D displays, the design must minimise any differences which are not directly related to the display format (for example, in these experiments, subjects should have been allowed to click on either the ground plot or the air plot of a particular aircraft to select it).

Regarding experiments which measured time to select an aircraft, the author infers that the overall effect of the sources of error mentioned above would be to slow the performance of subjects using the 3D displays, particularly the stereo-3D display. Specific sources of error which are expected to contribute to this are:

- The higher level of clutter on the 3D displays, due to the larger amount of symbology.
- The larger datablocks on the stereo-3D display (due to halving of the vertical resolution) lead to greater label overlap and clutter than the pseudo-3D or 2D displays.
- In the 3D displays, the subject was only allowed to select the aircraft on the air plot, rather than either the air or ground plots.

This suggests that results showing a shorter times for the 3D display formats (particularly the stereoscopic display) than the 2D format have a higher level of confidence associated with them than results which show the 2D format to give faster times.

8.5 Immersive Displays

Radar control was envisaged as the main task which might benefit from a three-dimensional visualisation, and so an immersive display of an air traffic scenario was demonstrated to subjects who were air traffic controllers. As a possible application of VR technology, this was universally rejected by subjects. Technology concerns aside, the main question seems to be the suitability of such displays for radar control purposes.

The author speculates that radar control is an *exocentric* task; the operator assumes a 'God's Eye' view of the airspace for which he or she is responsible and direct events within it from the perspective of an outside observer. The position of the aircraft relative to the observer is irrelevant and so 'presence' is not required. An immersive display gives an *egocentric* perspective, which, according to some subjects, may be suitable from a pilot's point of view (since the disposition of traffic relative to the 'ownship' is of interest), but not that of a controller. An immersive display could be envisaged whereby the airspace would be contained in an area separate from the operator, like observing fish in a tank from outside; however, this gains nothing from presence and so partly defeats the purpose of an immersive display.

The proposed VCR 'head-up display' application was, however, much better received. VCR tasks seem to be egocentric tasks, and ground movement controllers in particular are continually controlling aircraft 'heads-up'; aircraft have locations relative to the controller rather than being just points on a screen and the controller must turn his head and/or body to look for the aircraft. The proposed application might therefore have been better received because it would enable controllers to remain 'visual' and to retain the egocentric perspective in conditions where they would currently be forced to use a plan-view display. It might also eliminate the need to go 'head down' to consult auxiliary displays (such as runway approach monitors) since this information could also be shown 'head up'. However, technology would have to advance considerably for this to be feasible, and the likely high implementation cost would have to be weighed carefully against any potential benefits.

The most immediate application of current generation (or impending) immersive display technology seems to be for training or similar applications. The possible use of VR for modelling the new control tower was very well received, and architectural visualisation is certainly one of the areas for which VR appears to be suitable.

8.6 Conclusion

8.6.1 Contributions

To recap, the main reason for using 3D displays for ATC is the potential for reducing workload by eliminating the mental integration required in interpreting a 2D radar display with numeric height information. This seems to be based on the assumption that an integrated spatial mental image forms the key element in the 'picture'—overall awareness of the state of the traffic and decision-making.

The author contends that three-dimensional displays may be suitable for overall monitoring or conveying rapidly the disposition of traffic, but may not be so useful for making control decisions. There is some evidence to suggest that at least some air traffic controllers do attempt to translate the image with which they are presented into a three-dimensional visualisation. However, there are other, non-spatial elements in the air traffic controller's 'picture', and the non-spatial elements as well as the spatial elements are involved in decision-making. This research has found that although some controllers report trying to 'visualise' the traffic state in three dimensions (perhaps for *general awareness* of the traffic situation), when it comes to performance of some tasks which require more than just general awareness for overall monitoring but concentration on and manipulation of aircraft trajectories (for example, actively searching for conflicts) there is strong evidence of a tendency to consider the vertical and horizontal dimensions separately. If the vertical and horizontal dimensions are considered separately for active control, then this favours the use of the current 2D display rather than an integrated display based on the proximity compatibility principle. This is the key finding and major contribution of this thesis. The use of both 2D and 3D displays in parallel (the 3D for an overall view, the 2D for precision in control), such as proposed by Strutt, might therefore offer the best of both worlds, but it is likely that controllers would need training in order to take full advantage of both sources of information.

Other contributions are:

- An investigation of the performance of stereoscopic 3D displays relative to both pseudo-3D and 2D displays. As expected, stereoscopic 3D displays seem to give better task performance generally than pseudo-3D displays.
- An investigation into the differences between novice and expert subject groups

in the performance of tasks with 2D and 3D displays.

- A demonstration of virtual reality technology to air traffic controllers, with the aim of gathering opinions about future applications of VR to ATC.
- A simple aircraft model to simulate aircraft for the purposes of an ATC display. This was used in a subsequent publication as a simple model for controlling flying objects in a VR environment.

8.6.2 Other Applications

The emphasis of this thesis has been on considering the practicality of using three-dimensional displays for operational purposes. While the use of 3D displays for radar control is still open to debate, it might be practical to use 3D displays for non-operational purposes.

Because ATCOs are highly trained with the 2D displays, mental filtering and sorting of traffic to gain the 3D picture is practically second nature and a rapid process. However, it is a skill which takes time to develop and a 3D display might be useful for the novice.

Three-dimensional displays might therefore be useful in training or other visualisations such as reviewing incidents, airspace planning, evaluating tools such automatic conflict alert aids etc., which may benefit from integration into a 3D display to give an clear idea of the spatial relationships between aircraft rather than from a non-integrated view such as a 2D PPI which may be more suitable for control.

Introducing 3D displays in non-operational rôles might be a good path to follow for developing them further for similar applications, possibly eventual operational use. Since the applications are not critical, this may be a good way of gaining more experience and information about the practical benefits and drawbacks of 3D displays, and ways to overcome shortcomings, for real applications, rather than just measuring subject performance in a research laboratory.

8.6.3 Future Work

On the question of whether or not 3D displays may be useful for command/control of objects in real-time, safety critical areas, this thesis has raised more questions than it has answered, and has highlighted areas for further research. A display must be designed to support the operator's task, but while the objectives of the task may be clearly defined, the way in which the operator thinks about and carries out the task may not be so clear. Current research into the air traffic controller's 'picture', in particular on the rôle of a spatial mental picture (if such exists) in monitoring and control, remains incomplete. This question should be addressed in order to that benefits that 3D displays might bring to ATC can be more fully understood.

Further work is required to address the possible problems with perceptual ambiguities and biases in reading 3D displays. While it is clear that these exist, what is not clear is the magnitude of error that can be tolerated in air traffic control. If the errors are within tolerable limits (or can be made to be so by training or symbolic enhancement, for example) and if the operators are aware of the types of error intrinsic in such displays, then this may remove one obstacle to the acceptance of 3D displays. However, more research is required to determine precisely what the acceptable limits of errors, and to control them if they are too great. Further research is also required on the optimisation of display parameters and depth cues, since these can greatly affect the legibility of the display.

One possibly significant factor not investigated in this thesis is head-slaving the 3D TTW display, giving a movable viewpoint and motion parallax. This might be a solution to problems such as ambiguity of position and cluttering of objects along the same LOS, since with a head-slaved display the viewer can easily view the scene from a slightly different angle. Head-slaving might also enhance the sense of depth afforded by the display.

While evidence from 3D displays research should be considered, it cannot alone predict how operators will perform using the displays for a complex task. An evaluation of 3D displays for ATC should therefore use a full operational simulation, with conditions as close to real ones as possible. Only with such an evaluation can all the problems and issues be identified.

Appendix A
Pilot Experiment Instruction Sheets

Air Traffic Control Displays
Pilot Study

Thank you for taking the time to participate in this study. Your help is greatly appreciated. The purpose of my research is to evaluate different display technologies for air traffic control. To this end, you will be given three tasks concerned with various perceptual aspects of the job of air traffic control.

- Task 1 will be to read the relative heading and distance between pairs of aircraft on a three-dimensional display.
- Task 2 will be to memorise a static traffic scenario, and then try to recall it.
- Task 3 will be a conflict detection task. A couple of short scenarios will be presented on a display. You will be required simply to identify any potential conflicts (loss of separation) that exist.

After the tasks, you will then be shown some of the different display types and invited to comment.

Please do not hesitate to task the supervisor at any time if there is anything that you do not understand, or which is not clear.
Since this pilot study is running throughout this week, in the interests of not prejudicing the research, please do not discuss this with others until after Friday 30 September.

Task 1
Heading and Distance Estimation

In this task, you will be asked to estimate the heading and distance between 20 different pairs of aircraft on a three-dimensional display.

The display will show a 100 × 100 nautical mile region around the Heathrow radar, looking to the north. A video map is displayed on the "ground" which shows the Heathrow, London City and Gatwick CTAs in a light tan colour and the various airspace boundaries and airways. Range rings are also shown at 10nm intervals centered on the radar head at Heathrow.

Aircraft are represented as black "dots" in the air, with a 'drop' line joining their positions on the ground, which is represented by a white "dot".

Each aircraft has a label (datablock) in a box attached to it. The top line of the datablock contains the mode-A identification number of the aircraft. The bottom line shows its mode-C altitude (00s of feet of flight level). *The altitude will be ignored in this task.*

In the experiment, pairs of aircraft will be shown on the screen one after the other. For each pair, please write on the response sheet provided the estimated heading from aircraft 1 to aircraft 2 and the distance between them in nautical miles. Please take as much time as you like, and try to be as accurate as you can. When you have finished with one pair of aircraft, press the SPACE BAR on the keyboard to move on to the next pair of aircraft. When you have completed all 20 pairs, please tell the supervisor.

Please tell the supervisor when you are ready to proceed.

Task 2
Scenario Recall Task

For this task, you will be shown two static displays showing an air traffic scenario. For each display, you will be asked to memorise the air traffic pattern, and then to draw it on a piece of paper. The first display is just for familiarisation. The second will be recorded.
Each display shows a static air traffic scenario over the same area as the previous task. This will be shown for one minute; after this, it will be removed from the screen and you will then be given a piece of paper showing the radar map of the area. You will then be asked to mark on this piece of paper the positions of the targets and their altitudes from memory. Please try to make a best effort (for example, fill in a best guess for the altitude if you don't remember it precisely), working as quickly and as accurately as possible. Please tell the supervisor as soon as you have finished.
In these displays, some additional information will be given. "History" trails will be shown attached to each aircraft, (if you are shown a 3D display, these will be black for trails in the air, and their "shadows" on the ground will be white) giving its previous positions over the last 8 radar sweeps. A datablock (shown as text characters in a box) attached to each aircraft by a thin leader line shows the mode-A transponder code and mode-C height (00s of feet or flight level), possibly with a small character after it: an up arrow indicating that the aircraft is ascending, a down arrow indicating that the aircraft is descending, or no character to indicate that the aircraft is in level flight.
For this task, I am interested purely in position and height, and the previous positions, identification and vertical trend (climb/descend) can be ignored.

Task 3
Conflict Detection Task

The purpose of this task is for you to detect conflicts in an animated traffic scenario.

Familiarisation Display

Prior to the task, you will be shown a short animation of radar data to familiarise you with the display. The radar data are taken from Heathrow radar, at around 09:00 on a morning in April. The aircraft positions are updated every 6 seconds. Don't worry that the display shows a lot of aircraft; you will be shown far fewer in the task!

The display contains a rudimentary datablock overlap avoidance algorithm — that is, the computer will try to move the datablocks around to avoid them overlapping. They may therefore "jump" as the aircraft move, but this is nothing to worry about. The algorithm used is not perfect, however, and so sometimes the datablocks may overlap for short periods.

If you are shown a 2D display, the datablocks will show full information (mode-A code, mode-C height information and climb/descent, as for the previous task). If you are shown a 3D display, since the height is represented graphically, only the mode-A codes will be shown unless the SPACE BAR on the keyboard is pressed — pressing and holding this key for more than 0.5s will cause the full datablock to be shown. Releasing the key will cause only the mode-A code to be shown again.

Conflict Detection Task

In the task itself, you will be shown two short (less than 4 minute) animated traffic scenarios containing several aircraft. You will be asked simply to watch the scenario unfolding and to tell the supervisor when you think that any aircraft may lose separation.

Here, separation is defined as 3nm laterally, or 1 000 feet vertically (i.e. aircraft must be 3nm or greater apart horizontally if they are within 1 000ft of each other vertically, or they must be 1 000ft or greater apart vertically if within 3nm of each other horizontally). When you think two or more aircraft are in danger of coming into conflict with each other, either immediately or at some time in the next few minutes, please tell the supervisor immediately, with the following information:

Conflict Detection Task contd.

- The mode-A transponder codes of the aircraft involved.
- A conflict resolution manœuvre (e.g. turn aircraft 1052 left 20 degrees; descend aircraft 2047 immediately to FL330).

Please note that there may be more than one conflict in the scenario.

The datablocks also contain additional information — an extra line in the datablock shows the cleared altitude (if any) and route code. For example, the datablock:

```
1023
220v
210 LL
```

refers to aircraft mode-A code 1023, at FL220 and descending to its cleared level of FL210, its destination London Heathrow (ICAO code EGLL).

Virtual Reality Display
Instructions for Use

The virtual reality (VR) machine uses a head-mounted display (HMD) to show a three-dimensional scene filling your field of view which will change as you move your head. The scene is similar to the 3D air traffic displays on the computer, but this time you will be "inside" it as opposed to looking at it "through the window" of a computer screen. Before the task, you will be given a short familiarisation session to get used to it.

The HMD comprises a helmet containing two small television screens, one for each eye, with some wide-angle optics. Spectacles may be worn with this display. The helmet is connected to the VR machine by a cable at the back. Next to the cable is a power switch, and a nut which tightens or loosens the headband. Before you put the helmet on, please ensure that the nut is unscrewed. Then place the helmet on your head and tighten the nut so that the helmet is confortable but will not fall off if you lean over.

Take a little time to look around you when you first become "immersed", to get used to the scene changing when you move your head. Try squatting down and tilting your head to the side, and notice the effect. Also, try turning on the spot, and again notice how the scene changes.

You will also be given a hand-held device on which there are several buttons: three on the top (left, centre and right) and two at the front (top and bottom). You will be able to see the position of the device as an arrow if you look at the hand holding it in the virtual world. Try moving your hand about and notice how then arrow changes direction with it.

Moving in the virtual "world" can be accomplished in two ways — you can either step in any direction or you can "fly". Wlaking anywhere is rather restrictive because of the cable. Flying is therefore the preferred method of moving.

Flying is accomplished with the buttons on the top of the hand-held device. Pressing the left-hand button on the top of the device moves you forward in the direction in which the arrow (i.e. your hand) is pointing. Pressing the right-hand button moves you backwards in this direction. Notice that you can look sideways whilst travelling—just turn your head in any direction whilst keeping the arrow pointing in the desired direction of travel. As an exercise, without taking a step, try to fly to the north of the displayed virtual area.

When you feel that you are familiar with the virtual environment, please tell the supervisor that you are ready to proceed.

Appendix B
Main Experiment Instructions & Questionnaires

B.1 Introduction Sheet

Displays for Air Traffic Control

Thank you for taking the time to participate in this study. Your help is greatly appreciated.

The purpose of this research is to evaluate different display types for tasks such as Air Traffic Control and the like. To this end, you will be given 5 tasks concerned with various perceptual aspects of using displays, and at the end of the experiment you will be asked to do a spatial reasoning exercise. No prior expertise of air traffic control or radar displays will be required.

The tasks you will be asked to perform will be:

1. Reading the angle and distance between pairs of aircraft.
2. Choosing aircraft according to some criterion.
3. Memorising a scene.
4. Detection of impending collisions.
5. A "chaser" task.

Each task will include a practice session for familiarisation.

After each task, you will be invited to fill in a questionnaire for comments. Please feel free to answer as honestly as you like.

Please do not hesitate to ask the supervisor if at any time there is anything which you do not understand.

Since this study is continuing over a period of weeks, in the interests of not prejudicing this research I would be grateful if you did not discuss this with others who have not yet taken part in the study until after August 1995.

→ *Please now complete the questionnaire given to you, and tell the supervisor when you have finished it.*

B.2 Entry Questionnaire

CONFIDENTIAL

Questionnaire 1 Subject no.

1. Name: ______________________________

2. Occupation: ______________________________

3. Age. Please tick the box which applies:

≤ 20	
21–30	
31–40	
41–50	
51–60	
> 60	

4. Gender:

M		F	

5. Handedness:

Left		Right	

6. Eyesight: Do you wear contact lenses/glasses to correct your vision, or do you have any other vision defect (e.g. colour blindness)?

Yes		No	

If answer is **Yes**, please indicate your sight deficiency below by placing a tick in any boxes which apply:

	LEFT	RIGHT
Short-sighted		
Long-sighted		
Astigmatic		
Colour-blind		
Other (specify below)		

Other:

7. Education: Please indicate the highest level of education which you have either achieved, *either complete or incomplete* (tick ONE box).

CSE/O-level/GCSE	
A-level/IB	
City & Guilds	
ONC/OND	
HNC/HND	
BTEC	
Bachelor's degree	
Master's degree	
M.Phil/Ph.D.	
Other (specify below)	

Other:

8. Do you have any operational experience in Air Traffic Control?

Yes		No	

If answer is **No**, please go to **question 10**.

If answer is **Yes**, please summarise your operational experience briefly below. Try to give area with approximate time (e.g. en-route at LATCC 2 years, GMC at LGW 4 years, etc.):

9. Please state what types of ATC display you have used in the past operationally (tick any which apply):

None (procedural control only)	
Monochrome tube	
Colour synthetic display	
Other (please specify below)	

Other:

10. Have you played any Air Traffic Control games for leisure (e.g. *Tracon*)?

Yes		No	

If answer is **Yes**, please state game(s) below, and try to indicate how experienced you are (e.g. how long or how often you have played) at ATC games.

Games:

Novice: 1	
2	
3	
4	
5	
6	
Expert: 7	

11. Do you have any piloting experience (excluding games)?

Yes		No	

If answer is **Yes**, please state what sort of licence you hold/have held and which types and approximate number of hours (e.g. ATPL; glider, 22 hours; single-engine 100 hours; multi-engine 2500 hours):

12. Have you played any flight simulator games for leisure?

Yes		No	

If answer is **Yes**, please state which games, and try to indicate how experienced you are (e.g. how long or how often you have played) at flight simulator games.

Games:

Novice: 1	
2	
3	
4	
5	
6	
Expert: 7	

13. Have you used a computer mouse or trackerball before?

Yes		No	

END

B.3 Task 1 Instruction Sheet

B.3.1 Two-Dimensional Display

Task 1: Angle and Distance Reading

In this task, you will be asked to read the *angle* and *distance* between pairs of "blips" (representing aircraft) on a 2D plan-view radar-type display.

The 2D Display

The display will look like the diagram in figure 1 below. This shows a "video

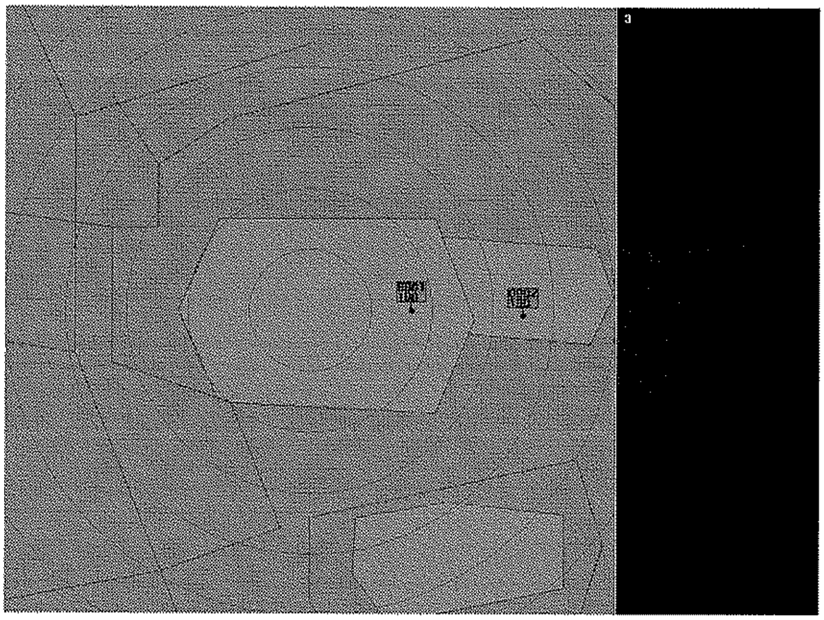
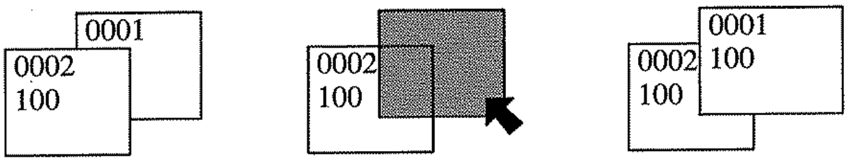


Figure 1:

map" (which is included for completeness, and will not be used here) and concentric *range rings* at 5 mile intervals. You will be using these as an aid to judging distances. North is towards the top of the screen.

Two aircraft are shown on the display, represented by the black filled circles. Attached to these by a thin black line are *datablocks*: these contain text relating to the aircraft. The first line of text shows the identification of the aircraft; the two aircraft are identified as 0001 and 0002. The second line shows the *altitude* of the aircraft, with 0 being at ground level and the number increasing with greater altitude. Height information is not relevant for this task, just the identification of the aircraft and their positions.

Occasionally, the aircraft may be close enough for their datablocks to overlap. In this case, it is possible to see the rear datablock by pressing the LEFT mouse button with the 'cursor' (a red arrow which moves when you move the mouse) over the rear datablock. This is shown in figure 2. In figure 2(a), the rear datablock is obscured. If we 'click' the LEFT mouse



(a) (b) (c)

Figure 2:

button in the shaded area of fig. 2(b), the rear datablock will be brought to the front, as shown in fig. 2(c).

This feature will be demonstrated to you before the start of the task.

The Task

In this task, you will be presented with twenty pairs of aircraft, one at at time, judging the angle and distance between each pair and marking your answers on an answer sheet.

When presented with each pair, please first of all write down in the second column headed `Distance` the distance which you think aircraft 1 is from aircraft 2. Please judge distances by eye only, rather than trying to make measurements on the display with a pen or similar. You can use the range rings to help you judging the distances, since these are at 5 mile intervals.

Next, write the *bearing angle* of aircraft 2 from aircraft 1 (measured *clockwise* from the North) in the `Angle` column. For example, in figure 3(a) the bearing of aircraft 2 from aircraft 1 is about 150°, whereas in figure 3(b), the bearing is about 290°. Please try to be as accurate as possible.

When you have finished the pair, please press the MIDDLE mouse button to move on to the next pair. Please feel free to take as much time as you

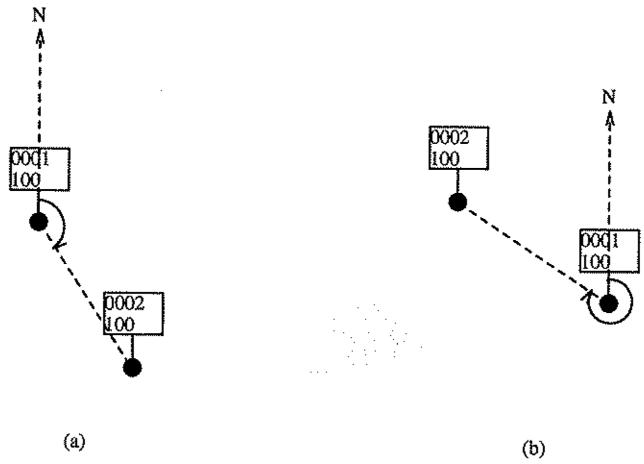


Figure 3:

like—time is not important in this task.

→ *Please tell the supervisor when you are ready to begin.*

B.3.2 Three-Dimensional Display

Task 1: Angle and Distance Reading

In this task, you will be asked to read the *angle* and *distance* between pairs of "blips" (representing aircraft) on a three-dimensional display of aircraft position.

The 3D Display

The display will be similar to that shown in figure 1 below. This shows a

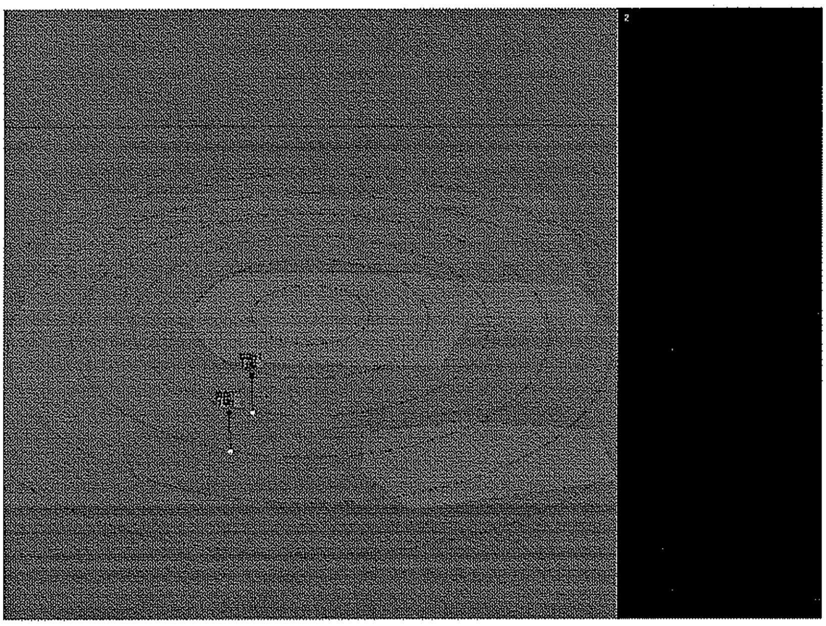


Figure 1:

"video map" (which is included for completeness, and will not be used here) and concentric *range rings* at 5 mile intervals. You will be using these as an aid to judging distances. North is towards the top of the screen.

Two aircraft are shown on the display. Their positions in the air are represented by the black filled circles. The points on the ground which they are directly above are indicated by white filled circles. The air and ground plots are joined by a *drop line* the length of which is proportional to the altitude of the aicraft. The height scale is exaggerated by a factor of 2.

Attached to each aircraft air position symbol by a thin black *leader line* is a *datablock*: this contains text information relating to the aircraft. The first line of text shows the identification of the aircraft; the two aircraft are identified as 1 (0001) and 2 (0002). The second line shows the *altitude* of the aircraft, with 0 being at ground level and the number increasing with greater altitude. Height information is not relevant for this task, just the identification of the aircraft and their positions.

Occasionally, the aircraft may be close enough for their datablocks to overlap. In this case, it is possible to see the rear datablock by pressing the LEFT mouse button with the 'cursor' (a red arrow which moves when you move the mouse) over the rear datablock. This is shown in figure 2. In figure 2(a), the rear datablock is obscured. If we 'click' the LEFT mouse button in the shaded area of fig. 2(b), the front datablock will be made transparent apart from its frame, allowing the datablock behind to be seen, as shown in fig. 2(c).

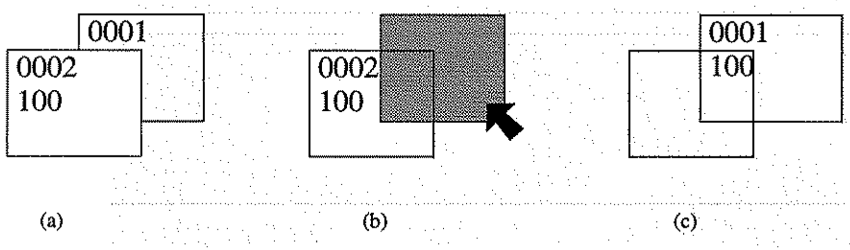


Figure 2:

This feature will be demonstrated to you before the start of the task.

The Task

In this task, you will be presented with twenty pairs of aircraft, one at a time, judging the angle and distance between each pair and marking your answers on an answer sheet.

When presented with each pair, please first of all write down in the second column headed `Distance` the distance which you think aircraft 1 is from aircraft 2. Please judge distances by eye only, rather than trying to make measurements on the display with a pen or similar. You can use the range rings to help you judging the distances, since these are at 5 mile intervals.

Next, write the *bearing angle* of aircraft 2 from aircraft 1 (measured *clockwise* from the North) in the `Angle` column. For example, in figure 3(a) the bearing of aircraft 2 from aircraft 1 is about 150°, whereas in figure 3(b), the bearing is about 290°. Please try to be as accurate as possible.

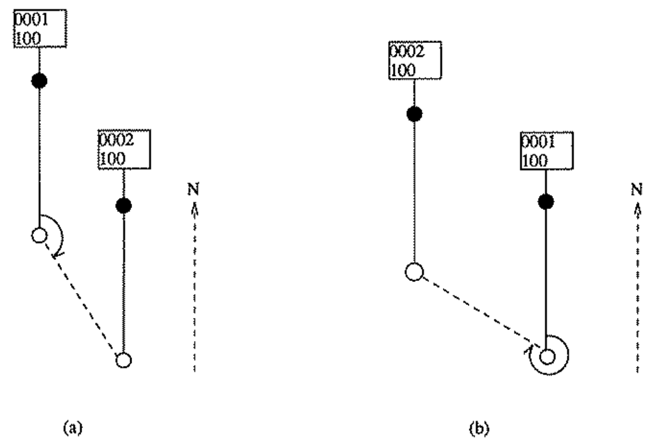


Figure 3:

When you have finished the pair, please press the MIDDLE mouse button to move on to the next pair. Please feel free to take as much time as you like—time is not important in this task.

→ *Please tell the supervisor when you are ready to begin.*

B.4 Task 1 Response Sheet

Task 1: Angle and Distance Reading

Subject: 88 Cell: 1

	Distance	Angle
TR1		
TR2		
15		
13		
7		
5		
14		
9		
1		
18		
17		
16		
12		
20		
2		
10		
19		
8		
11		
6		
4		
3		

B.5 Task 1 Questionnaire

Questionnaire 2	Task 1	Subject no.

1. How did you find reading *distances* from the display?

Very Easy: 1	
2	
3	
4	
5	
6	
Very Hard: 7	

2. How accurately did you think you read *distances* from the display? Please put a mark on the line below indicating your estimated margin of error (for example, if you think you estimated distances to about ±1.5 miles, please put a mark half way between the '1' and the '2' positions on the line).

0 1 2 3 4 5 6 7 8 9 10

Margin of Error

3. Please write any comments which you have about the ease of the display for reading distances. (Please continue your answer on the back of the sheet or request additional paper if necessary.)

4. How did you find reading *angles* from the display?

Very Easy: 1	
2	
3	
4	
5	
6	
Very Hard: 7	

5. How accurately did you think you read *angles* from the display? Please put a mark on the line below indicating your estimated margin of error (for example, if you think you estimated angles to $\pm 5°$, please put a mark on the '5' position on the line).

0 5 10 15 20 25 30 35 40 45

Margin of Error (Degrees)

6. Please write any comments which you have about the ease of the display for reading angles. (Please continue your answer on the back of the sheet or request additional paper if necessary.)

7. Please feel free to write any further comments which you have about the task, the display or anything else which you think worthy of note. Please request additional sheets of paper if necessary.

B.6 Task 2 Instruction Sheet

B.6.1 Two-Dimensional Display

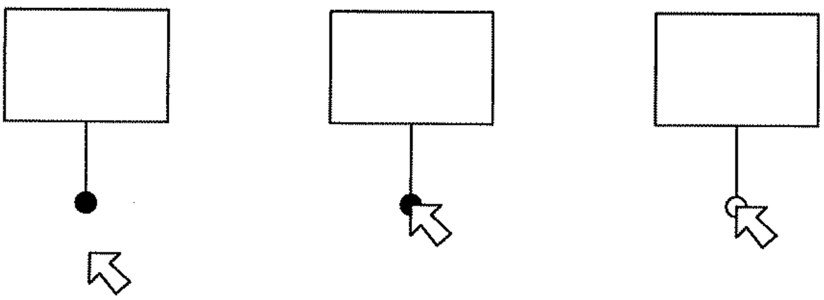
Task 2: Data Extraction

This task is divided into two separate parts. For each part, you will be presented with 10 scenes in turn containing a number of aircraft. For each scene, you will be required simply to select *two* of the aircraft based on a given criterion.

Selecting Aircraft

Each task involves selecting two aircraft. Each aircraft may be selected as shown in figure 1.

You can select an aircraft by moving the cursor point over its symbol (the black circle) (fig. 1(a)) and pressing the LEFT mouse button (fig. 1(b)). The symbol will then turn red to indicate that the aircraft has been selected (fig. 1(c)). If you select an aircraft by mistake, you can deselect it by clicking



(a) (b) (c)

Figure 1:

on its symbol again with the LEFT mouse button.

As described in the previous task, if two datablocks overlap, you can see the obscured one by clicking on it. Clicking on a datablock does *not* select an aircraft (unless you click on an aircraft which is visible through a datablock).

After selecting both desired aircraft, move to the next scene by pressing the MIDDLE mouse button. The system will only allow you to do this when you have selected two aircraft.

You will be given an opportunity to practice this before the start of the task.

1

The Task

This task is divided into two parts.

Part 1 Select the *highest* and the *lowest* aircraft. You may select the aircraft in any order (lowest first or highest first).

Part 2 Select the aircraft which are *closest horizontally* to each other—i.e. do not take height into account when choosing which aircraft are closest together, but select which are the closest aircraft *as if they were all at the same height.*

Your performance will be assessed on speed and accuracy, so please work as *quickly* and as *accurately* as you are able.

You will be given an opportunity to practice before the start of each part.

→ *Please tell the supervisor when you are ready to proceed.*

B.6.2 Three-Dimensional Display

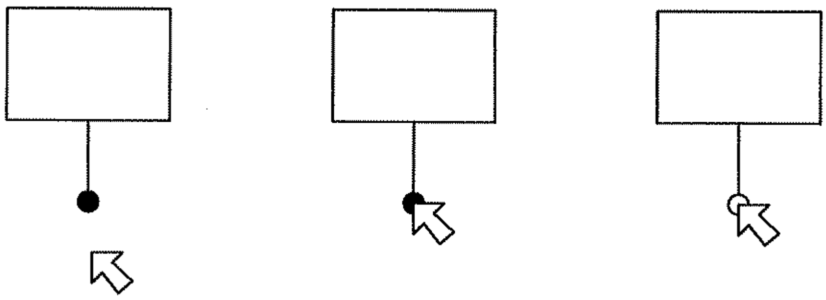
Task 2: Data Extraction

This task is divided into two separate parts. For each part, you will be presented with 10 scenes in turn containing a number of aircraft. For each scene, you will be required simply to select *two* of the aircraft based on a given criterion.

Selecting Aircraft

Each task involves selecting two aircraft. Each aircraft may be selected as shown in figure 1.

You can select an aircraft by moving the cursor point over its air plot symbol (the black circle) (fig. 1(a)) and pressing the LEFT mouse button (fig. 1(b)). The symbol will then turn red to indicate that the aircraft has been selected (fig. 1(c)). If you select an aircraft by mistake, you can deselect



(a) (b) (c)

Figure 1:

it by clicking on its symbol again with the LEFT mouse button.

As described in the previous task, if two datablocks overlap, you can see the obscured one by clicking on it. Clicking on a datablock does *not* select an aircraft (unless you click on an aircraft which is visible through a datablock).

After selecting both desired aircraft, move to the next scene by pressing the MIDDLE mouse button. The system will only allow you to do this when you have selected two aircraft.

You will be given an opportunity to practice this before the start of the task.

The Task

This task is divided into two parts.

Part 1 Select the *highest* and the *lowest* aircraft. You may select the aircraft in any order (lowest first or highest first).

Part 2 Select the aircraft which are *closest horizontally* to each other—i.e. do not take height into account when choosing which aircraft are closest together, but select which are the closest aircraft *as if they were all at the same height.* This means choosing the white ground plot dots which are closest together.

Your performance will be assessed on speed and accuracy, so please work as *quickly* and as *accurately* as you are able.

You will be given an opportunity to practice before the start of each part.

→ *Please tell the supervisor when you are ready to proceed.*

B.7 Task 2 Questionnaire

Questionnaire 3 Task 2 Subject no.

1. How did you find reading the highest and lowest aircraft?

Very Easy:	1	
	2	
	3	
	4	
	5	
	6	
Very Hard:	7	

2. Please write any comments which you have about the ease of the display for picking out the highest and the lowest aircraft. (Please continue your answer on the back of the sheet or request additional paper if necessary.)

3. How did you find reading the aircraft with the closest horizontal proximity?

Very Easy: 1	
2	
3	
4	
5	
6	
Very Hard: 7	

4. Please write any comments which you have about the ease of the display for picking out the horizontally closest aircraft. (Please continue your answer on the back of the sheet or request additional paper if necessary.)

5. Please feel free to write any further comments which you have about the task, the display or anything else which you think worthy of note. Please request additional sheets of paper if necessary.

B.8 Task 3 Instruction Sheet

B.8.1 Two-Dimensional Display

Task 3: Memory Recall

In this task, you will be shown eight scenes in turn. For each scene, you are required spend a period of time memorising it, then recall what you remembered.

Before beginning each scene, you will be given an answer sheet on which to record what you remembered. This will initially be face-down. Please *do not turn the paper over yet.*

The display starts off with a blank screen. When you are ready to start, press the MIDDLE mouse button. The scene will then appear and will remain on the screen for 90 seconds; the screen will then go blank.

Whilst the scene is being displayed, please try to memorise the following information:

- Aircraft position
- Aircraft identification number
- Aircraft altitude

When the screen goes blank, please turn the piece of paper over and try to draw the information on it. This paper is marked with the "video map" and range rings at *10* mile intervals to aid orientation. Please mark on the positions of each aircraft (with a cross, for example) and write the identification and height next to the positions. When you have finished, please hand the paper to the supervisor.

The task will be timed, from when you turn the paper over to when you hand it to the supervisor, so please work as quickly and as accurately as you are able. If you cannot remember the information precisely, *please try to do the best you can.* If you cannot remember the positions or heights exactly, please try to make a "best guess".

→ *Please tell the supervisor when you are ready to begin.*

B.8.2 Three-Dimensional Display

Task 3: Memory Recall

In this task, you will be shown eight scenes in turn. For each scene, you are required spend a period of time memorising it, then recall what you remembered.

Before beginning each scene, you will be given an answer sheet on which to record what you remembered. This will initially be face-down. Please *do not turn the paper over yet.*

The display starts off with a blank screen. When you are ready to start, press the MIDDLE mouse button. The first scene will then appear and will remain on the screen for 90 seconds; the screen will then go blank.

Whilst the scene is being displayed, please try to memorise the following information:

- Aircraft *ground* position
- Aircraft identification number
- Aircraft altitude

When the screen goes blank, please turn the piece of paper over and try to draw the information on it. This paper is marked with the "video map" and range rings at *10* mile intervals to aid orientation. Please mark on the positions of each aircraft (with a cross, for example) and write the identification and height next to the positions. You may find it helpful to draw the approximate length of the drop line. When you have finished, please hand the paper to the supervisor.

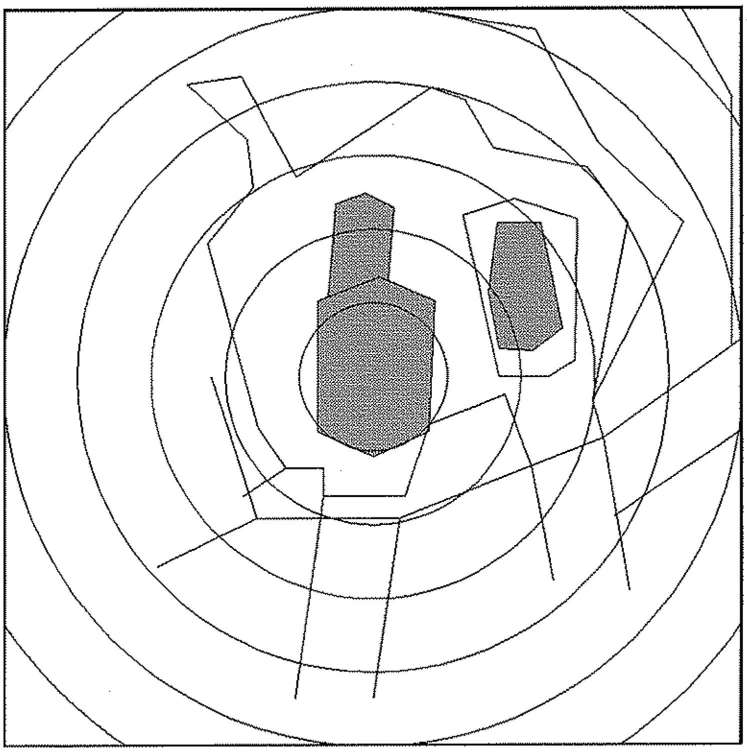
The task will be timed, from when you turn the paper over to when you hand it to the supervisor, so please work as quickly and as accurately as you are able. If you cannot remember the information precisely, *please try to do the best you can.* If you cannot remember the positions or heights exactly, please try to make a "best guess".

→ *Please tell the supervisor when you are ready to begin.*

B.9 Task 3 Response Sheet

Task 3: Practice Answer Sheet

Cell:	-	Trial:	-



Time:________________

B.10 Task 3 Questionnaire

Questionnaire 4 Task 3 Subject no.

1. How did you find the memory task overall?

Very Easy:	1	
	2	
	3	
	4	
	5	
	6	
Very Hard:	7	

2. Which were the easiest and most difficult aspects of the scenes to memorise?

Position:

Very Easy:	1	
	2	
	3	
	4	
	5	
	6	
Very Hard:	7	

Identification:

Very Easy:	1	
	2	
	3	
	4	
	5	
	6	
Very Hard:	7	

Height:

Very Easy:	1	
	2	
	3	
	4	
	5	
	6	
Very Hard:	7	

1

3. Please try to describe how you went about memorising the scenes. For example, did you remember position by the patterns the groups formed, or because they resembled traffic patterns; did you remember the heights by visualising them in your head, or did you rely on memory of the numbers. (Please feel free to use the other side of the paper, or request additional paper if necessary.)

4. Please write any comments which you have about the task, the display or anything else which you think worthy of note. Please request additional sheets of paper if necessary.

B.11 Task 4 Instruction Sheet

B.11.1 Two-Dimensional Display

Task 4: Conflict Detection

The purpose of this task is to examine the detection of potential collisions, or *conflicts*. In this task, you will be shown a number of short scenes (each of 90s duration) containing a number of aircraft which are moving, and be asked to identify any conflicts.

Conflicts

In air traffic control, two aircraft must be separated by certain minimum distances, a horizontal and a vertical. In this task, aircraft are not allowed to approach each other within three miles horizontally and 10 units vertically — that is, if the difference in altitude of two aircraft is less than 10, they must be more than 3 miles apart horizontally, and if they are closer than 3 miles apart horizontally, there must be a height difference of at least 10 between them.

If these criteria are met, two aircraft are said to be *separated*. A *conflict* is defined as a situation where two aircraft, if left continuing their present manœuvres, will lose separation.

The Display

The display is updated at 6 second intervals — i.e. every 6 seconds, the aircraft symbols will 'jump' to their new positions. Range rings are set at *10* mile intervals, not 5 miles as in the first task.

As aircraft positions change, they leave a trail of dots behind them marking where they have been. These enable the course and speed of the aircraft to be estimated from the direction of the trail and the spacing between the dots (see figure 1).

In the datablock, the following information is given (see fig. 1). The first line of text is the identification of the aircraft; the second line gives the current altitude, and after that a character (either 'v' or '^') which indicates whether the aircraft is ascending or descending (this is blank if the aircraft is level). The third line of the datablock contains a two letter destination code (not used here, but included for completeness) and the altitude the aircraft has been given permission to change to. This is blank if the aircraft has not been instructed to change its altitude.

Figure 1 therefore shows an aircraft whose identification is BA102, flying at an altitude of 100 and descending, which has been authorised to descend to an altitude of 98, but no lower.

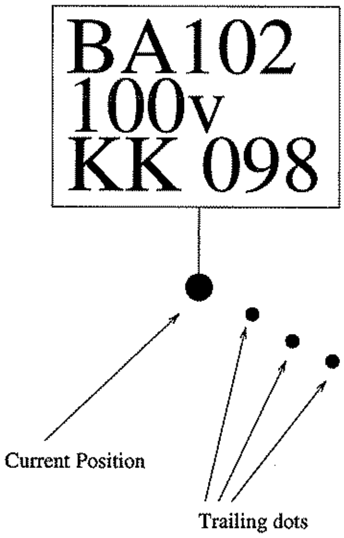


Figure 1:

For examples of a conflict and a non-conflict situation, please look at figure 2. This shows two aircraft whose paths are crossing. Let us suppose that when they cross, the horizontal separation will be less than 3 miles. The situation shown in fig. 2a is *not* a conflict, because although JA124 is descending and crossing the path of BA1284, it has been cleared to level off just above it at an altitude of 110, which is 10 clear of BA1284 and so is safe. The situation in figure 2b, on the other hand, *is* a conflict, because JA124 has been cleared to descend *through* the altitude of BA1284 and will pass within 10 vertically whilst it is less than 3 miles away laterally.

Note also that if JA124 in figure 2a *failed* to level off at its cleared altitude but carried on descending, this would also be a conflict.

The Task

At the start of each scenario, a blank screen will be shown. When you are ready to start, please press the MIDDLE mouse button. The scenario will then start. Initially, there will be no trailing dots, but these will appear as time progresses. Your task will be simply to observe the scenario and to see whether or not a conflict will occur. *Note that scenarios do not necessarily contain a conflict.*

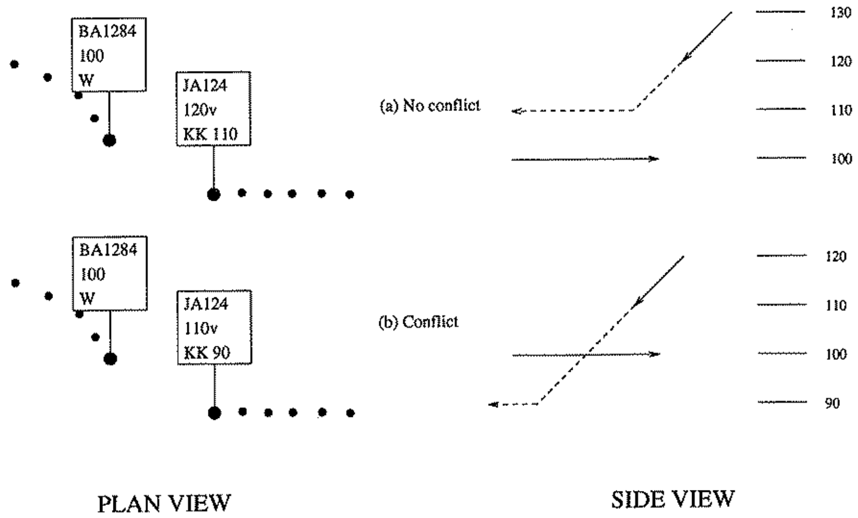


Figure 2:

If you spot a conflict, i.e. a situation where loss of separation either has occurred *or will occur in the near future*, press the MIDDLE mouse button. As soon as you do this, the scenario will stop. Then click on the symbols of the two aircraft which you think are conflicting with the LEFT mouse button (as in task 2). Finally, tell the supervisor what you would have the aircraft do to prevent the conflict from occurring: for example, you could instruct one of the aircraft to level off (i.e. to stop changing altitude), or to turn. When you have selected both desired aircraft, press the MIDDLE mouse button to move to the next scenario.

Please note that once you have pressed the MIDDLE mouse button when you have spotted a conflict, the scenario will stop and you must make a selection by clicking on the two aircraft. If you detected a conflict my mistake (i.e. one which did not actually exist), click on the two aircraft which you thought might have conflicted anyway. You will *not* have the opportunity to reverse your decision.

You can still reveal obscured datablocks by clicking on them with the LEFT mouse button as before.

You will be given a demonstration and practice before the task itself.

→ *Please tell the supervisor when you are ready to begin.*

B.11.2 Three-Dimensional Display

Task 4: Conflict Detection

The purpose of this task is to examine the detection of potential collisions, or *conflicts*. In this task, you will be shown a number of short scenes (each of 90s duration) containing a number of aircraft which are moving, and be asked to identify any conflicts.

Conflicts

In air traffic control, two aircraft must be separated by certain minimum distances, a horizontal and a vertical. In this task, aircraft are not allowed to approach each other within three miles horizontally and 10 units vertically — that is, if the difference in altitude of two aircraft is less than 10, they must be more than 3 miles apart horizontally, and if they are closer than 3 miles apart horizontally, there must be a height difference of at least 10 between them.

If these criteria are met, two aircraft are said to be *separated*. A *conflict* is defined as a situation where two aircraft, if left continuing their present manœuvres, will lose separation.

The Display

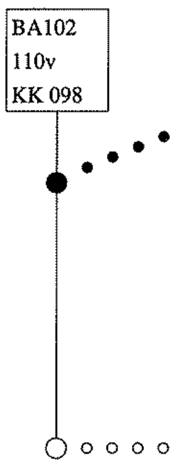


Figure 1:

The display is updated at 6 second intervals—i.e. every 6 seconds, the aircraft symbols will 'jump' to their new positions. Range rings are set at *10* mile intervals, not 5 miles as in the first task.

As aircraft positions change, they leave a trail of dots behind them marking where they have been. These enable the course and speed of the aircraft to be estimated from the direction of the trail and the spacing between the dots (see figure 1).

In the datablock, the following information is given (see fig. 1). The first line of text is the identification of the aircraft; the second line gives the current altitude, and after that a character (either 'v' or '^') which indicates whether the aircraft is ascending or descending (this is blank if the aircraft is level). The third line of the datablock contains a two letter destination code (not used here, but included for completeness) and the altitude the aircraft has been given permission to change to. This is blank if the aircraft has not been instructed to change its altitude.

Figure 1 therefore shows an aircraft whose identification is BA102, flying at an altitude of 110 and descending, which has been authorised to descend to an altitude of 98, but no lower.

For examples of a conflict and a non-conflict situation, please look at figure 2. This shows two aircraft whose paths are crossing. Let us suppose that when they cross, the horizontal separation will be less than 3 miles. The situation shown in fig. 2a is *not* a conflict, because although JA124 is descending and crossing the path of BA1284, it has been cleared to level off just above it at an altitude of 110, which is 10 clear of BA1284 and so is safe.

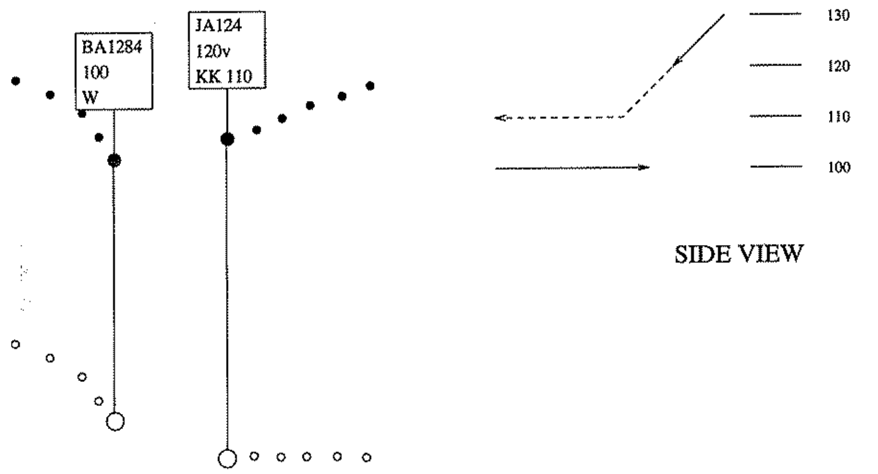
The situation in figure 2b, on the other hand, *is* a conflict, because JA124 has been cleared to descend *through* the level of BA1284 and will pass within 10 units vertically whilst it is less than 3 miles away laterally.

Note also that if JA124 in figure 2a *failed* to level off at its cleared level but carried on descending, this would also be a conflict.

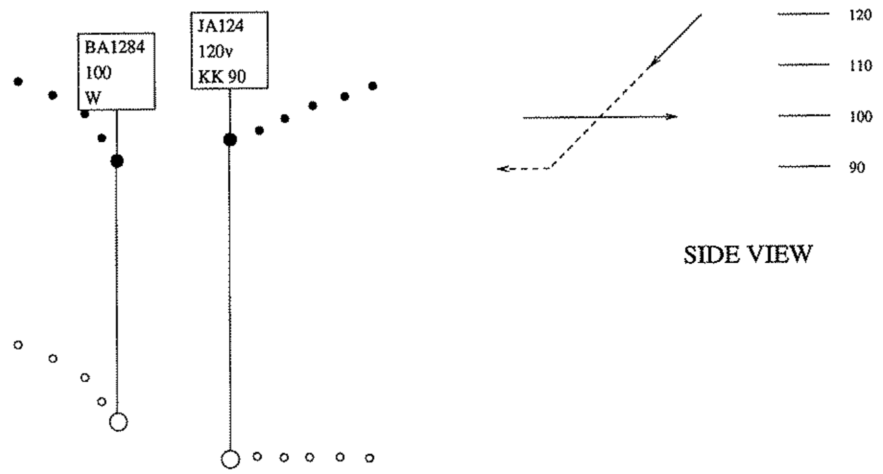
The Task

At the start of each scenario, a blank screen will be shown. When you are ready to start, please press the MIDDLE mouse button. The scenario will then start. Initially, there will be no trailing dots, but these will appear as time progresses. Your task will be simply to observe the scenario and to see whether or not a conflict will occur. *Note that scenarios do not necessarily contain a conflict.*

If you spot a conflict, i.e. a situation where loss of separation either has



(a) No conflict



(b) Conflict

Figure 2:

occurred *or will occur in the near future*, press the MIDDLE mouse button. As soon as you do this, the scenario will stop. Then click on the symbols of the two aircraft which you think are conflicting with the LEFT mouse button (as in task 2). Finally, tell the supervisor what you would have the aircraft do to prevent the conflict from occurring: for example, you could instruct one of the aircraft to level off (i.e. to stop changing altitude), or to turn. When you have selected both desired aircraft, press the MIDDLE mouse button to move to the next scenario.

Please note that once you have pressed the MIDDLE mouse button when you have spotted a conflict, the scenario will stop and you must make a selection by clicking on the two aircraft. If you detected a conflict my mistake (i.e. one which did not actually exist), click on the two aircraft which you thought might have conflicted anyway. You will *not* have the opportunity to reverse your decision.

You can still reveal obscured datablocks by clicking on them with the LEFT mouse button as before. You will be given a demonstration and practice before the task itself.

→ *Please tell the supervisor when you are ready to begin.*

B.12 Task 4 Questionnaire

Questionnaire 5 Task 4 Subject no.

1. How did you find the conflict detection task overall?

Very Easy: 1	
2	
3	
4	
5	
6	
Very Hard: 7	

2. Please try to describe how you went about doing the conflict detection task. For example, did you look at pairs of aircraft, first of all looking at ground position, then referring to height? Did you take speed into account? Please request additional sheets of paper if necessary.

1

3. Please write any comments which you have about the task, the display or anything else which you think worthy of note. Please request additional sheets of paper if necessary.

B.13 Task 5 Instruction Sheet

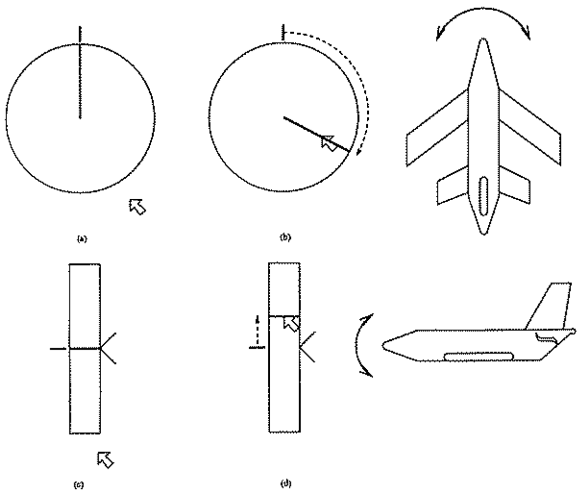
B.13.1 Two-Dimensional Display

Task 5: Chaser Task

In this, the last task, you will be required to control a "chaser" to "catch" a target. The task is divided into two stages.

First, a blank screen is shown. When you are ready to begin the task, press and release the MIDDLE mouse button. A static (non-moving) scene containing a number of aircraft will then be shown. You are required to select the target which is *closest* to a point on the ground at the middle of the displayed area by 'clicking' on it with the LEFT mouse button. (If two aircraft are the same horizontal distance away from the centre, you should therefore select the *lower* of the two.) There are 60 vertical units per mile.

When the selection is made, all aircraft apart from the one you selected will disappear, and a "chaser" (under your control) will appear and start to move. The target will also start to move. The task will now be to 'catch' the target by guiding the chaser within a certain distance of it (1 mile horizontally and 10 units vertically). To do this, you will be using the controls to the right of the screen (see figure below). Time to catch will be assessed, so please try to do this in the shortest possible time.



The circular control on the left controls the horizontal component of the chaser's path (i.e. its course—see (a) above). The outer pointer is where the chaser is actually heading, the long needle inside the circle is the control (where you want the chaser to go). To change the horizontal component of the chaser's path, move the red cursor inside the circle and press the LEFT mouse button. The needle will then "jump" to the tip of the cursor, and the outer pointer will start to move around to the selected course (there is a delay in the chaser's response), as shown in (b). If you keep the left mouse button pressed, the needle will 'follow' any movement of the cursor until the left mouse button is released.

The box on the right ((c) in the diagram) controls the vertical component of the chaser's path and is operated in similar fashion. The bar to the left of the box shows the current path angle of the chaser; the bar inside the box is the control. The arrow to the right of the box corresponds to bar position for level flight (i.e. neither ascending nor descending). Below the arrow, the flight path is increasingly in a *downwards* direction; if the bar is at the bottom of the box, the chaser is travelling vertically downwards. Above the arrow, the flight path is increasingly in an *upwards* direction; if the bar is at the top of the box, the chaser is travelling vertically upwards.

Moving the cursor into the box and pressing the LEFT mouse button causes the bar to "jump" to the tip of the cursor. The outer pointer will then start to move around to the selected flight path angle (see (d)). Again, the bar will follow the cursor so long as the left mouse button remains depressed.

The chaser will never be allowed to go to negative altitude (i.e. underground) even if the flight path is downwards.

You will be given an opportunity to practice this control before the task begins.

→ *Please tell the supervisor when you are ready to proceed.*

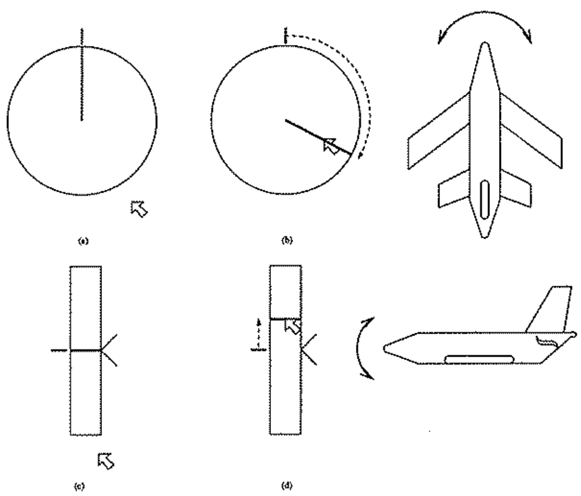
B.13.2 Three-Dimensional Display

Task 5: Chaser Task

In this, the last task, you will be required to control a "chaser" to "catch" a target. The task is divided into two stages.

First, a blank screen is shown. When you are ready to begin the task, press and release the MIDDLE mouse button. A static (non-moving) scene containing a number of aircraft will then be shown. You are required to select the target which is *closest* to a point on the ground at the middle of the displayed area by 'clicking' on it with the LEFT mouse button. (If two aircraft are the same horizontal distance away from the centre, you should therefore select the *lower* of the two.) There are 60 vertical units per mile, and remember that the vertical scale on the display is exaggerated by a factor of 2.

When the selection is made, all aircraft apart from the one you selected will disappear, and a "chaser" (under your control) will appear and start to move. The target will also start to move. The task will now be to 'catch' the target by guiding the chaser within a certain distance of it (1 mile horizontally and 10 units vertically). To do this, you will be using the controls to the right of the screen (see figure below). Time to catch will be assessed, so please try to do this in the shortest possible time.



The circular control on the left controls the horizontal component of the chaser's path (i.e. its course—see (a) above). The outer pointer is where the chaser is actually heading, the long needle inside the circle is the control (where you want the chaser to go). To change the horizontal component of the chaser's path, move the red cursor inside the circle and press the LEFT mouse button. The needle will then "jump" to the tip of the cursor, and the outer pointer will start to move around to the selected course (there is a delay in the chaser's response), as shown in (b). If you keep the left mouse button pressed, the needle will 'follow' any movement of the cursor until the left mouse button is released.

The box on the right ((c) in the diagram) controls the vertical component of the chaser's path and is operated in similar fashion. The bar to the left of the box shows the current path angle of the chaser; the bar inside the box is the control. The arrow to the right of the box corresponds to bar position for level flight (i.e. neither ascending nor descending). Below the arrow, the flight path is increasingly in a *downwards* direction; if the bar is at the bottom of the box, the chaser is travelling vertically downwards. Above the arrow, the flight path is increasingly in an *upwards* direction; if the bar is at the top of the box, the chaser is travelling vertically upwards.

Moving the cursor into the box and pressing the LEFT mouse button causes the bar to "jump" to the tip of the cursor. The outer pointer will then start to move around to the selected flight path angle (see (d)). Again, the bar will follow the cursor so long as the left mouse button remains depressed.

The chaser will never be allowed to go to negative altitude (i.e. underground) even if the flight path is downwards.

You will be given an opportunity to practice this control before the task begins.

→ *Please tell the supervisor when you are ready to proceed.*

B.14 Task 5 Questionnaire

Questionnaire 6	Task 5	Subject no.

1. How did you find the chaser task overall?

Very Easy:	1	
	2	
	3	
	4	
	5	
	6	
Very Hard:	7	

2. Please write any comments which you have about the task, the display or anything else which you think worthy of note. Please request additional sheets of paper if necessary.

B.15 Immersive Display Instructions

Virtual Reality Display Instructions for Use

The virtual reality (VR) machine uses a head-mounted display (HMD) to show a three-dimensional scene filling your view which will change as you move your head. The scene is similar to the 3D air traffic displays on the computer, but this time you will be "inside" it as opposed to looking at it "through the window" of a computer screen. Before the task, you will be given a short familiarisation session to get you used to it.

The HMD comprises a helmet containing two small television screens, one for each eye, with some wide-angle optics. Spectacles may be worn with this display. The helmet is connected to the VR machine by a cable at the back. Next to the cable is an power switch, and a nut which tightens or loosens a headband. Before you put the helmet on, please ensure that the nut is unscrewed. Then place the helmet on your head and tighten the nut so that the helmet is comfortable but will not fall off if you lean over.

Take a little time to look around you when you first become "immersed", to get used to the scene changing when you move your head. Try squatting down and tilting your head to the side, and notice the effect. Also, try turning on the spot, and again notice how the scene changes.

You will also be given a hand-held device on which there are several buttons: three on the top (left, centre and right) and two at the front (top and bottom). You will be able to see the position of the device as an arrow if you look at the hand holding it in the virtual world. Try moving your hand about and notice how the arrow changes direction with it.

Moving in the virtual "world" can be accomplished in two ways — you can either step in any direction or you can "fly". Walking anywhere is rather restrictive because of the cable. Flying is therefore the preferred method of moving.

Flying is accomplished with the buttons on the top of the hand-held device. Pressing the left-hand button on the top of the device moves you forward in the direction in which the arrow (i.e. your hand) is pointing. Pressing the right-hand button moves you backwards in this direction. Notice that you can look sideways whilst travelling—just turn your head in any direction whilst keeping the arrow pointing in the desired direction of travel. As an exercise, without taking a step, try to fly to the north of the displayed virtual area.

When you feel that you are familiar with the virtual environment, please tell the supervisor that you are ready to proceed.

Appendix C
Main Experiment Data Analysis

C.1 Introduction

This appendix contains tables of descriptive statistics and statistical analyses of the data from the main experiment.

In descriptive statistics tables, each cell entry is usually of the form '$\overline{X} \pm s$, N', where $\overline{X}$, s and N are the sample mean and standard deviation and the number of observations for each cell respectively.

Unless otherwise indicated, analyses were carried out at a significance level of 0.05.

C.2 Spatial Ability Test

Group	N	$\overline{X} \pm s$
N	14	70.1 ± 8.0
X	22	67.7 ± 13.6

Table C.1: Spatial Ability Scores: Descriptive Statistics

C.3 Task 1: Azimuth Angle and Relative Distance

Variable	Estimate	S.E.
const	-0.6631	0.7826
λ_A	0.9845	0.004815
$D = 2$	-2.125	1.078
$D = 3$	-1.685	1.147
B_x	-0.1489	0.06693
B_y	0.1625	0.06769
$D = 2.B_x$	0.3681	0.09292
$D = 3.B_x$	0.2910	0.09884
$D = 2.B_y$	-0.4485	0.09327
$D = 3.B_y$	-0.3533	0.09920

Correlation coeff. = 0.98

Table C.2: y_A Linear Model Estimates and Standard Errors

D	N	X
1	-2.1 ± 14.8, 120	1.0 ± 7.6, 140
2	-4.0 ± 16.5, 100	-3.7 ± 11.7, 180
3	-3.5 ± 10.7, 100	-3.2 ± 11.4, 140

Table C.3: Azimuth Error y_e: Descriptive Statistics

D	N	X
1a	13.7 ± 13.5, 6	8.5 ± 3.6, 7
1b	8.4 ± 4.5, 5	8.5 ± 3.6, 7
2	12.5 ± 5.0, 5	12.1 ± 3.9, 9
3	7.8 ± 2.5, 5	13.7 ± 4.8, 6

1a contains an outlying point, 1b omits it.

Table C.4: Subjective Azimuth Accuracy: Descriptive Statistics

Source	df	SS	MS	F
Display	2	16.615	8.307	0.196
Group	1	0.368	0.368	0.009
Interaction	2	182.017	91.008	2.149
Residual	32	1354.956	42.342	
Total	37	1553.629		

Table C.5: y_{S_A}: Two-factor ANOVA Analysis, all data

Source	df	SS	MS	F
Display	2	92.814	46.407	2.757
Group	1	25.930	25.930	1.541
Interaction	2	70.323	35.161	2.089
Residual	31	521.769	16.831	
Total	36	714.953		

Table C.6: y_{S_A}: Two-factor ANOVA Analysis, selected data

D	N	X
1	(5.0) 4.5 ± 1.8, 6	(2.0) 2.1 ± 0.7, 7
2	(3.0) 3.0 ± 1.9, 5	(3.0) 3.4 ± 1.6, 9
3	(2.0) 2.4 ± 0.9, 5	(3.5) 3.4 ± 1.3, 7

Entries in the form (median) $\overline{X} \pm s$, N

Table C.7: Subjective Azimuth Reading Difficulty: Descriptive Statistics

Source	df	SS	MS	F
Display	2	0.308	0.154	0.078
Group	1	0.814	0.814	0.413
Interaction	2	22.147	11.074	5.615*
Residual	32	63.113	1.972	
Total	37	86.382		

* significant at $\alpha = 0.01$

Table C.8: Subjective Azimuth Reading Difficulty: Two-factor ANOVA

Variable	Estimate	S.E.
const	-0.009477	0.02518
λ_D	0.9997	0.01279
$D = 2$	-0.04307	0.01973
$D = 3$	-0.03600	0.02099
B_x	-0.0008434	0.001218
$D = 2.B_x$	0.005362	0.001688
$D = 3.B_x$	0.002103	0.001795

Correlation coeff. = 0.89

Table C.9: $\ln y_D$ Linear Model Estimates and Standard Errors

D	N	X
1	1.42 ± 0.92, 6	1.03 ± 0.24, 7
2	1.30 ± 0.67, 5	1.30 ± 0.36, 9
3	1.76 ± 0.54, 5	1.57 ± 0.44, 6

Table C.10: Subjective Distance Reading Accuracy: Descriptive Statistics

Source	df	SS	MS	F
Display	2	1.273	0.636	2.137
Group	1	0.346	0.346	1.162
Interaction	2	0.243	0.121	0.408
Residual	32	9.528	0.298	
Total	37	11.418		

Table C.11: Subjective Distance Reading Accuracy: Two-factor ANOVA

D	N	X
1	(3.5) 4.2 ± 1.5, 6	(2.0) 2.6 ± 1.1, 7
2	(3.0) 3.4 ± 0.9, 5	(4.0) 3.8 ± 1.6, 9
3	(3.0) 3.4 ± 0.9, 5	(3.5) 3.4 ± 1.3, 6

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.12: Subjective Distance Reading Difficulty: Descriptive Statistics

C.4 Task 2: Information Extraction

C.4.1 Altitude Extraction

D	N	X
1	11.2 ± 4.5, 60	12.7 ± 4.4, 70
2	14.2 ± 8.1, 50	11.5 ± 4.3, 90
3	9.6 ± 2.3, 50	12.0 ± 5.0, 60

Table C.13: Altitude Extraction Time y_{AT}: Descriptive Statistics

Source	df	SS	MS	F
Display	2	146.097	73.048	2.980
Group	1	9.639	9.639	0.393
Interaction	2	457.331	228.666	9.330*
Residual	374	9166.585	24.510	
Total	379	9779.652		

* significant at $\alpha = 0.01$

Table C.14: Altitude Extraction Time y_{AT}: Two-factor ANOVA

Cells	df	F	p
G =N	2, 157	9.164*	< 0.01
G =X	2, 217	1.352	0.261
N1, N2	1, 108	6.072†	0.015
N2, N3	1, 98	14.793*	< 0.01
N1, N3	1, 108	5.013†	0.027
N1, X1	1, 128	3.839‡	0.052
N2, X2	1, 138	6.517*	0.012
N3, X3	1, 108	9.886*	< 0.01

* significant at $\alpha = 0.01$

† significant at $\alpha = 0.05$

‡ borderline significant at $\alpha = 0.05$

Table C.15: Altitude Extraction Time y_{AT}: Single-Factor ANOVA

D	N	X
1	(2.0) 2.3 ± 1.0, 6	(2.0) 2.3 ± 0.8, 7
2	(1.0) 1.6 ± 0.9, 5	(2.0) 2.0 ± 0.7, 9
3	(2.0) 2.4 ± 1.5, 5	(2.0) 2.0 ± 0.9, 6

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.16: Altitude Extraction Subjective Difficulty: Descriptive Statistics

C.4.2 Horizontal Proximity Extraction

D	N	X
1	6.2 ± 3.7, 60	6.9 ± 2.6, 70
2	11.7 ± 8.7, 50	10.2 ± 5.2, 90
3	8.9 ± 5.4, 50	9.1 ± 5.9, 60

Table C.17: Horizontal Proximity Extraction Time y_{HT}: Descriptive Statistics

Source	df	SS	MS	F
Display	2	1146.077	573.039	19.735*
Group	1	4.822	4.822	0.166
Interaction	2	91.835	45.918	1.581
Residual	374	10859.803	29.037	
Total	379	12102.537		

* significant at $\alpha = 0.01$

Table C.18: Horizontal Proximity Extraction Time: Two-factor ANOVA

D	N	X
1	(2.0) 2.7 ± 1.5, 6	(1.0) 1.3 ± 0.5, 7
2	(2.0) 2.0 ± 0.7, 5	(4.0) 3.9 ± 0.9, 9
3	(4.0) 4.0 ± 1.6, 5	(4.0) 3.8 ± 1.2, 6

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.19: Horizontal Proximity Extraction Subjective Difficulty: Descriptive Statistics

Source	df	SS	MS	F
Display	2	24.661	12.330	10.253*
Group	1	0.145	0.146	0.121
Interaction	2	17.560	8.780	7.301*
Residual	32	38.484	1.203	
Total	37	80.850		

* significant at $\alpha = 0.01$

Table C.20: Horizontal Proximity Extraction Subjective Difficulty: Two-factor ANOVA

Cells	df	F	p
$G = \text{N}$	2, 13	2.902	0.091
$G = \text{X}$	2, 19	19.942*	< 0.001
X1, X2	1, 14	44.912*	< 0.001
X2, X3	1, 13	0.011	0.920
X1, X3	1, 11	27.918*	< 0.001
N1, X1	1, 11	5.311†	0.042
N2, X2	1, 12	15.482*	< 0.001
N3, X3	1, 9	0.041	0.845

* significant at $\alpha = 0.01$

† significant at $\alpha = 0.05$

Table C.21: Horizontal Proximity Extraction Subjective Difficulty: Single-factor ANOVA

C.5 Task 3: Memory Recall

C.5.1 Recall Times

D	N	X
1	62.2 ± 23.8	54.5 ± 28.1
2	74.6 ± 30.4	63.7 ± 30.8
3	64.9 ± 33.0	60.7 ± 32.8

D & G

A	R	T
3	45.1 ± 19.7	36.1 ± 10.0
5	71.7 ± 28.3	71.1 ± 24.8
7	73.9 ± 31.8	63.7 ± 31.8
9	73.1 ± 31.9	68.4 ± 32.4

A & R

Table C.22: Memory Recall Times: Descriptive Statistics

Source	df	SS	MS	F
Display	2	6478.375	3239.188	3.725*
Group	1	4689.781	4689.781	5.393*
Interaction	2	200.438	100.219	0.115
Residual	296	257 414.130	869.642	
Total	301	268 782.724		

* significant at $\alpha = 0.05$

Table C.23: Memory Recall Times: Two-factor ANOVA between D & G

Source	df	SS	MS	F
A	3	52 671.938	17 557.313	24.403*
R	1	2 891.859	2 891.859	4.019†
Interaction	3	831.531	277.177	0.385
Residual	294	211 521.690	719.462	
Total	301	267 917.018		

* significant at $\alpha = 0.01$

† significant at $\alpha = 0.05$

Table C.24: Memory Recall Times: Two-factor ANOVA between A & R

C.5.2 Recall Subjective Difficulty

D	N	X
1	(6.0) 6.0 ± 0.9, 6	(6.0) 5.7 ± 0.8, 7
2	(7.0) 6.2 ± 1.3, 5	(7.0) 6.3 ± 0.9, 9
3	(6.0) 6.2 ± 0.4, 5	(7.0) 6.5 ± 0.8, 6

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.25: Memory Recall Overall Subjective Difficulty: Descriptive Statistics

D	N	X
1	(4.5) 4.2 ± 1.0, 6	(4.0) 4.0 ± 1.0, 7
2	(3.0) 3.0 ± 1.9, 5	(4.0) 3.9 ± 1.1, 9
3	(3.0) 3.0 ± 1.0, 5	(4.0) 3.8 ± 1.3, 6

Overall difficulty (all data): (4.0) 3.7 ± 1.2, 38

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.26: Memory Position Recall Subjective Difficulty: Descriptive Statistics

D	N	X
1	(5.5) 5.3 ± 1.6, 6	(6.0) 5.6 ± 1.4, 7
2	(7.0) 6.4 ± 0.9, 5	(6.0) 5.4 ± 1.0, 9
3	(6.0) 5.6 ± 1.7, 5	(5.5) 5.7 ± 1.2, 6

Overall difficulty (all data): (6.0) 5.6 ± 1.3, 38

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.27: Memory Ident Recall Subjective Difficulty: Descriptive Statistics

D	N	X
1	(6.0) 5.8 ± 0.8, 6	(6.0) 5.4 ± 0.8, 7
2	(4.0) 4.6 ± 1.8, 5	(5.0) 5.3 ± 1.5, 9
3	(5.0) 4.8 ± 0.8, 5	(6.0) 5.8 ± 1.2, 6

Overall difficulty (all data): (5.5) 5.3 ± 1.2, 38

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.28: Memory Height Recall Subjective Difficulty: Descriptive Statistics

C.5.3 Position Analysis

D	N	X
1	8.5 ± 9.4	7.5 ± 8.0
2	11.2 ± 8.8	13.9 ± 11.6
3	13.3 ± 11.8	9.4 ± 8.5

D & G

A	R	T
3	10.6 ± 9.6	3.7 ± 3.8
5	15.1 ± 11.3	10.5 ± 9.4

A & R

Table C.29: Recalled Position Error for Stimuli 1–4: Descriptive Statistics

Source	df	SS	MS	F
Display	2	2790.977	1395.488	14.36*
Group	1	43.750	43.750	0.45
Interaction	2	1074.934	537.467	5.53*
Residual	602	58 499.016	97.174	
Total	607	62 408.676		

* significant at $\alpha = 0.01$

Table C.30: Recalled Position Error: Two-factor ANOVA between D & G

Source	df	SS	MS	F
A	1	4514.313	4514.313	51.342*
R	1	4565.203	4565.203	51.921*
Interaction	1	182.668	182.668	2.078
Residual	604	53 107.789	87.927	
Total	607	62 369.973		

* significant at $\alpha = 0.01$

Table C.31: Recalled Position Error: Two-factor ANOVA between A & R

C.6 Task 4: Conflict Detection Task

Stim	N1	N2	N3	X1	X2	X3
A	6, 0	1, 4	2, 3	4, 3	4, 5	1, 5
B	4, 2	2, 3	3, 2	6, 1	6, 3	4, 2
C	6, 0	1, 4	1, 4	7, 0	8, 1	5, 1
D	6, 0	5, 0	5, 0	7, 0	7, 2	5, 1
E	6, 0	5, 0	5, 0	4, 3	6, 3	5, 1
F	5, 1	4, 1	4, 1	7, 0	9, 0	5, 1

Entries are in the form Ncorrect, Nincorrect

Table C.32: Conflict Detection: Number of Correct and Incorrect Responses

D	N	X
1	(2.5) 2.3 ± 0.8, 6	(2.0) 2.0 ± 0.6, 7
2	(3.0) 3.0 ± 1.6, 5	(3.0) 3.0 ± 1.2, 9
3	(3.0) 2.8 ± 1.5, 5	(3.0) 3.5 ± 1.6, 6

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.33: Conflict Detection Subjective Difficulty: Descriptive Statistics

D	N	X
1	51.4 ± 18.7, 16	40.6 ± 21.7, 23
2	47.1 ± 14.9, 12	47.6 ± 18.5, 30
3	48.4 ± 21.5, 15	32.7 ± 20.0, 21

Table C.34: Conflict Detection Time: Descriptive Statistics

Source	df	SS	MS	F
Display	2	1686.531	843.266	2.205
Group	1	2036.871	2036.871	5.328*
Interaction	2	1196.121	598.061	1.564
Residual	111	42 442.777	382.367	
Total	116	47 362.301		

* significant at $\alpha = 0.05$

Table C.35: Conflict Detection Time: Two-factor ANOVA

C.7 Task 5: Chaser Task

D	N	X
1	10.7 ± 11.0, 30	8.0 ± 6.5, 35
2	15.6 ± 14.4, 25	8.4 ± 5.7, 45
3	8.7 ± 6.5, 25	11.2 ± 7.0, 30

Table C.36: Task 5 Selection Time: Descriptive Statistics

Source	df	SS	MS	F
Display	2	1.339	0.670	1.41
Group	1	3.277	3.277	6.90*
Interaction	2	5.330	2.665	5.61*
Residual	184	87.40	0.475	
Total	189	97.346		

* significant at $\alpha = 0.01$

Note: gamma distribution used

Table C.37: Task 5 Selection Time: Two-factor ANOVA

D	N	X
1	2.0 ± 1.2, 30	2.3 ± 2.1, 35
2	3.4 ± 3.3, 24	3.0 ± 1.9, 44
3	2.4 ± 1.3, 25	3.1 ± 2.2, 29

Table C.38: Normalised Interception Time: Descriptive Statistics

Source	df	SS	MS	F
Display	2	4.762	2.381	6.23*
Group	1	0.149	0.149	0.39
Interaction	2	1.224	0.612	1.60
Residual	181	69.190	0.382	
Total	186	75.325		

* significant at $\alpha = 0.01$

Note: gamma distribution used

Table C.39: Task 5 Normalised Interception Time: Two-factor ANOVA

D	N	X
1	(2.0) 1.8 ± 0.8, 6	(2.0) 2.6 ± 0.8, 7
2	(2.0) 2.6 ± 1.5, 5	(2.0) 2.4 ± 0.5, 9
3	(2.0) 2.2 ± 0.8, 5	(3.0) 3.0 ± 1.4, 6

Entries are in the form (median) $\overline{X} \pm s$, N

Table C.40: Chaster Task Subjective Difficulty: Descriptive Statistics

Appendix D
Task 1 Subject Comments

D.1 Distance Reading

These are responses to the question '*How did you find reading* distances *from the display?*'.

1 N1 **D.1** *Gaze-directed centering of rings would be nice!*

2 N2 **D.2** *The concentric circles [range rings] hindered as well as helped. More marks on the ground e.g. a mesh, would have helped.*

3 N2 **D.3** *Distances took time to figure out but got faster at the end.*

4 N1 **D.4** *You can add some polar lines [radii] (gray and slightly visible) using some points [along them at 0.5 mile intervals to aid distance judgment]. These lines also can help for accurate estimation of angle as well.*

4 N1 **D.5** *You can add a magnifying window which is invisible normally. But if the left most button [on the mouse] is pressed on the spot that window make the area bigger and the system will give the accurate distance.*

5 N2 **D.6** *Maybe it would be easier if the concentric circles are closer (let's say 2.5 miles). It depends on how exact we want the measurement to be. It could be useful if we have line connecting the 'planes' projections on earth. Then it would be easier to judge the angle and consequently the distances as well.*

8 N2 **D.7** *Short distances relatively easy. Longer distances increasingly harder.*

8 N2 **D.8** *Easier to judge if approx. radial.*

10 N2 **D.9** *When the 'planes were at 0°, 90°, 180° or 270° it was easier than at intermediate angles.*

10 N2 **D.10** *The very wide distances seemed harder.*

11 N1 **D.11** *Rather easy to use. I'm sure given practice once can become very accurate with it.*

12 N3 **D.12** *The concentric circles were very helpful.*

14 N1 **D.13** *I found the circles to be a bit misleading measuring distances from 2 points.*

15 N2 **D.14** *Lines from the centre at certain angles would have helped ('cos I was imagining them).*

16 N3 **D.15** *The slanting made it tougher.*

20 N1 **D.16** *The reading of the distance is easier to determine when the aircraft lie in a line at 0°, 90°, 180°, 270° or at angles of 45°, 135° etc., but as they drift from these angles, and the distance between them increases, it is much harder to estimate the distance accurately.*

21 N2 **D.17** *No difficulty interpreting the display.*

23 N2 **D.18** *Reading distances was easy when the lines between the aircraft were normal to the circles [range rings].*

23 N2 **D.19** *This becomes harder as the separation increased (and also when they were not on the circle.)*

24 N1 **D.20** *Cross-hairs might have been useful?*

25 X2 **D.21** *Because you're looking at 4 points between contacts instead of 2, there's a tendency to think the distances are greater than perhaps they are.*

26 X1 **D.22** *Greater distances apart — margin of error would increase.*

26 X1 **D.23** *An inexperienced eye may try to measure distances between labels.*

26 X1 **D.24** *Range marks essential to the task.*

27 X2 **D.25** *Similar to what I use at work.*

28 X3 **D.26** *Fairly similar distance wise to reading them off a 2D display. Maybe slightly harder because the display seems quite strange to start with.*

29 X1 **D.27** *The display is good it is just my ability to be confident about my answers which is difficult!*

30 X2 **D.28** *Distances N–S or E–W were "easy"—using the relevant range rings as comparison. It became harder the more the angle became 'diagonal' to the x and y coordinates [axes]. I felt the distances were true along the East/West and North/South lines respectively but the further I moved out from these known areas the harder it became to "factor in" a scaled up or down value.*

31 X3 **D.29** *North South compared to East West axis difference scale but after familiarisation fairly easy to balance the two.*

32 X1 **D.30** *Distances easier to estimate at top of display as it is at 90° to viewer; distances not so easy to estimate at bottom of display because they are viewed at an angle.*

33 X2 **D.31** *With some of the examples which were at an "angle" to the range rings and some distance from the centre, I found it difficult to estimate without some sort of reference.*

34 N2 **D.32** *Seemed easy enough.*

35 X3 **D.33** *Seemed easier to assess distance east–west. Some uncertainty to assess distance accurately north–south. More depth perception seems to be required. Otherwise fairly easy.*

36 X1 **D.34** *I thought the display was acceptable for reading distances with the addition of the 5 mile range markers for reference.*

38 X3 **D.35** *I found in both cases of angle and distance more difficulty than assessing from a normal flat display.*

39 X1 **D.36** *Never easy when not radial. Extended runway centre lines with ranges do help when available.*

40 X2 **D.37** *The 3D element is obviously a new phenomenon. Perhaps I was a little wary about using my instinctive judgment to come up with my answers. However, I did complete each task [stimulus] fairly quickly.*

41 X3 **D.38** *The range rings seemed to vary in distance apart when looking at the display thus I found myself having to assess each distance gap individually whereas normally once using a radar display [PPI] I can more easily assess distance.*

42 X1 **D.39** *Easier because of familiarity.*

43 X2 **D.40** *Scale is obviously variable. Takes a bit of getting used to.*

43 X2 **D.41** *(Seems) easier to read distance when perpendicular to range rings.*

44 X3 **D.42** *When the positions of the aircraft are side by side, I found it much harder to judge the distance.*

45 X1 **D.43** *Neutral colours were restful on the eyes. Contrast was good.*

46 X2 **D.44** *It takes time to 'get used' to the oblique angle.*

46 X2 **D.45** *The method I used to judge distance was to transpose No. 2 in an arc to the vertical of No. 1 and imagine the distance [i.e. rotate the display mentally until the targets were along the y-axis and imagine the distance.] This takes time and obviously for different parts of the screen, all distances look different. It would take much experience & practice to judge distances for different parts of the screen quickly.*

47 X3 **D.46** *Getting used to the 3D effect was quite difficult especially when tangential to the circle.*

48 N1 **D.47** *Would be easier if the "centre spot" (i.e. the centre of the innermost circle [range ring]) was marked, I think.*

48 N1 **D.48** *Dots marking the 'aircraft' were a little big, making it harder to decide on the distance.*

49 X2 **D.49** *With the circles [range rings] being elliptical, it's harder to judge angled distances—if the two aircraft are in an area of narrow circles do you use that as the guide or the widest part. [i.e. does one use the part of the range rings along the x-axis (wide) or the y-axis (narrow)].*

D.2 Angle Reading

These are responses to the question '*How did you find reading angles from the display?*'.

2 N2 **D.50** *You have to remember there is a plane.*

5 N2 **D.51** *Maybe a line connecting the projections, as I said before [quote D.6], would be useful.*

7 N1 **D.52** *Maybe it would be nice to have a picture of compass, which would altogether ease orientation in space, especially when we talk about the angles.*

8 N2 **D.53** *"Slanted" angle of view suggests ease of reading angles would vary depending on what the angle was. Tends to "flatten" angles across display.*

10 N2 **D.54** *Judging the intermediate angles (i.e. not right angles) was more difficult.*

11 N1 **D.55** *OK, but not subjectively as easy as distance estimation.*

12 N3 **D.56** *Axis going through 'plane 1 [local axis indicator] would make things much easier (or maybe a horizontal grid).*

13 N1 **D.57** *An indication in the form of cross-wires [radii along the x and y axes and at 45/225° and 135/315°] may help.*

14 N1 **D.58** *No comments really. It was fairly straightforward.*

15 N2 **D.59** *Localised angle lines ... would have been helpful (click to appear) [i.e. compass rose around reference target.]. Distance vector as well!*

16 N3 **D.60** *This is easier than reading distances because you could think in 2-D.*

18 N3 **D.61** *If I could have drawn a temporary line between the 2 points, reading the angle & the distances would have been easier.*

20 N1 **D.62** *The estimation of angles is much easier when the aircraft are close to one another, but the further apart they are, the more difficult to accurately estimate the angle becomes.*

21 N2 **D.63** *Again, no difficulty reading the display.*

23 N2 **D.64** *This was generally OK.*

24 N1 **D.65** *Compass markings on the screen might have been useful.*

25 X2 **D.66** *The availability of a cursor would make it easier.*

26 X1 **D.67** *Although no cursor display, familiar airspace helped with reading angles.*

26 X1 **D.68** *Closer aircraft together greater error in reading angles.*

27 X2 **D.69** *Nothing on the display made angle reading more or less difficult.*

28 X3 **D.70** *Some of the angles seemed harder to read than others depending on where the two object[s] (aircraft) were relative to the centre of the display.*

29 X1 **D.71** *I have always worked with a Nth marker which I find helps my decision making. A compass rose around the outside of the tube is also helpful. Both the above are replaced by experience.*

30 X2 **D.72** *I estimated the angles as 2-dimensionally displayed and adjusted as I thought necessary for their positions on the display.*

30 X2 **D.73** *This task seemed harder than distance judgment.*

30 X2 **D.74** *I was looking for a compass rose and radial lines for assistance. Without these I found I was unsure in many instances.*

31 X3 **D.75** *Same as 2D radar display (I think!).*

33 X2 **D.76** *I found it difficult in that having the ground positions and air positions made the task harder.*

34 N2 **D.77** *Bit tricky depending on the position where the perspective of the display gets more acute.*

35 X3 **D.78** *Again [as with angle] easier east–west. Not so easy within sectors 330°–030° and 150°–210° because of 3D effect creating 'slant' effect.*

36 X1 **D.79** *This was not so easy [as angle] as there was no reference facility e.g. a compass rose around the display.*

37 X2 **D.80** *More difficult than the distances.*

39 X1 **D.81** *Known airspace boundaries on radar give a reference from which to judge. Likewise known tracks and final approach lines.*

40 X2 **D.82** *Same as before (quote D.37 above). 3D display a new concept to me.*

41 X3 **D.83** *Quite easy—as with most angular readings I guess my most inaccurate ones are when the symbols are closer together but I find that a problem with current displays I use.*

42 X1 **D.84** *Again, it is a familiar environment for an experienced radar controller.*

43 X2 **D.85** *Parts of the 'radar' were more difficult to read the angles, such as the corners. Easier to read on the axis (2) L to R and top to bottom through the centre.*

44 X3 **D.86** *Harder to estimate the angles when the aircraft were close together.*

45 X1 **D.87** *Good [display] resolution and neutral colours assisted judgment.*

46 X2 **D.88** *Obviously this should be easier [than distance reading] because the graduation of the oblique is consistent across the screen. Angles are therefore 'real'. Any inconsistencies are probably due to my lack of practice on the screens for quite a while!*

47 X3 **D.89** *I think I was quite inaccurate, again especially when the 'blips' were tangential to the circle [range ring].*

48 N1 **D.90** *A little difficult—no horizontal or vertical or crosswise guidelines (a couple of faint horizontal ones, but wasn't absolutely sure they* were *horizontal.*

48 N1 **D.91** *Harder when dots v. close together.*

49 X2 **D.92** *Again, it's a lot harder to see angles when viewing an elliptical display as opposed to a flat 2D screen.*

D.3 General Comments

These are responses to a question asking for any further comments.

2 N2 **D.93** *I am not sure about how perspective works in this display.*

3 N2 **D.94** *Is there a difference between ground and screen angles?*

12 N3 **D.95** *It was sometimes confusing when the 'planes were vertically above one another [along the same y-coordinate with drop lines overlapping].*

13 N1 **D.96** *Sometimes, I find that I might have made a mistake concerning where I read the angle from—that's not sure whether I read angle from aircraft 1 or not.*

15 N2 **D.97** *I think it was fairly easy.*

23 N2 **D.98** *The lack of markers in the environment i.e. off diagonal distances (nor normal) made the distance estimation task more difficult.*

25 X2 **D.99** *Under normal operational usage the calculation of the angle between 1 a/c and another is not something I would envisage trying to calculate.*

25 X2 **D.100** *The line between the air contact & ground position is a little unusual as present arrangements have the a/c's datablock (callsign, altitude) connected to the a/c's position on the radar [with a similar leader line].*

29 X1 **D.101** *The display is not cluttered when compared with real life!*

29 X1 **D.102** *The display and task were easy to interpret—the answers were the difficult part.*

30 X2 **D.103** *Practice at this task will make it easier. Also if guidelines were available as to adjustments necessary to my 2-D thinking for each sector of the display then learning to read 3-D would be quicker. Once in practice I feel this 3-D display could be as comfortable to use as the present 2-D system.*

35 X3 **D.104** *Would respond more easily to: [air] target–white (thinking 'in the sky' therefore illuminated), ground position–black (thinking 'in shadow').*

41 X3 **D.105** *The aircraft were easy to distinguish relative to each other but I found relating them to the range rings as a guide to distance not so easy.*

44 X3 **D.106** *As with any new display the problem is familiarity. More usage would increase my accuracy.*

Appendix E
Task 2 Subject Comments

This appendix contains comments made by subjects on the second task.

E.1 Altitude Extraction

These are responses to the instruction '*Please write any comments which you have about the ease of the display for picking out the highest and the lowest aircraft.*

1 N1 **E.1** *Just the speed makes doing this harder.*

2 N2 **E.2** *Numbers [datablock altitude readout] helped to make final decision.*

3 N2 **E.3** *Some feedback when I had mouse over selector [i.e. highlight aircraft in some way when mouse sensitive area is over the selectable point].*

5 N2 **E.4** *It might be easier if we have points on the height lines at distance 10 miles [1 000ft!], let's say. Then one could judge from the number of points. But I'm not sure, because too much information could be confusing.*

8 N2 **E.5** *Use lines first then numbers. Very easy to select lowest in this manner.*

10 N2 **E.6** *It was quicker to estimate the height by looking at the black lines between the dots and then checking by reading the nos [datablock altitude readout].*

12 N3 **E.7** *Maybe a colour coding for ranges of heights would make the reading easier.*

13 N2 **E.8** *Line [length] is initially used to choose highest. Use display [datablock altitude readout] for accuracy.*

14 N1 **E.9** *It was fairly simple to just look at the first digit [of datablock altitude readout] for most of them.*

15 N2 **E.10** *The drop lines were only really useful where there was a good spread of heights e.g. not so good if all bunch around 100 ft.*

15 N2 **E.11** *Good for quick overview & datablock resolved conflicts.*

16 N3 **E.12** *The height displays [datablock altitude readouts] made it too easy.*

17 N3 **E.13** *Very low aircraft were easy to spot without reading the numbers (e.g. between 10 and 49) but aircraft between 100 and 290 were harder to differentiate (I had to look at the numbers).*

18 N3 **E.14** *If the altitude displayed in the box was of different colours, e.g. yellow—it would have stood out more, allowing an easier double check—after having guessed/read the 2 a/c by judging their altitudes using the vertical line.*

19 N3 **E.15** *Very easy if the difference between highest and lowest [aircraft] was great. If all aircraft were about the same height then it was more difficult, had to read all the numbers to make a decision whereas before the drop line showed the differences better.*

20 N1 **E.16** *The way the data is displayed allows for easy determination of aircraft altitudes. However, if several aircraft are located close to one another, then determining their altitudes becomes more difficult since you have to select each aircraft to display its data [i.e. data block obscuration means that the blocks have to be 'selected' to make them transparent so that the block behind may be read].*

21 N2 **E.17** *No difficulty once I had broken the previous task "strategy". [i.e. some learning effect from previous task carried over.*

21 N2 **E.18** *Easy to read.*

23 N2 **E.19** *This was quite a straightforward task.*

23 N2 **E.20** *I had the strategy of selecting the lowest aircraft first (using the drop line) then searching for the highest using a combination of the drop line and actual height.*

24 N1 **E.21** *To distinguish heights (from other details), an improvement could be made by making altitudes a different coloured font.*

25 X2 **E.22** *Initially I looked at the heights [in the datablock] to find the highest/lowest. Towards the end of the exercise used the length of the line more which gave faster reaction and used height readouts to confirm decision.*

26 X1 **E.23** *A little more difficult than reading horizontal distances as more scanning involved i.e. reading digits.*

27 X2 **E.24** *I still tended to use data block height readout (through force of habit) rather than just the drop line.*

28 X3 **E.25** *Finding the lowest a/c easier because with a shorter line for the height a smaller difference was easier to distinguish than with the higher a/c where the same difference because harder to distinguish because it was a lower percentage of the overall line length.*

29 X1 **E.26** *No problems.*

30 X2 **E.27** *The drop lines make it easier than with a data block alone. I soon learnt to scan the drop lines to find the longest & shortest. This eliminated a number of*

aircraft, meaning less time was needed to compare data blocks on similar length drop lines.

34 N2 **E.28** *Straightforward.*

35 X3 **E.29** *Definitely using altitude readout rather than visual perception [referring to drop-line length]—afraid of error using latter.*

35 X3 **E.30** *Still want plan position black (shadow), air position white (in 'skylight').*

36 X1 **E.31** *I thought it was very clear which a/c was highest/lowest.*

37 X2 **E.32** *Done by numbers OK. But using the rods I would have been wildly inaccurate/conditioning of ht comes from lowest blip 2500 but 268 would only have come out as F140 by rods.*

37 X2 **E.33** *Not wish to do it by rods alone.*

37 X2 **E.34** *Ht code rods:*

F300
RED
F200
GRN
F100
BLUE

39 X1 **E.35** *No problems.*

40 X2 **E.36** *Fairly straightforward.*

41 X3 **E.37** *I found the lowest was easy to pick out due to the small line length, I used that a lot and only used figures in the TAB block when there were two similar ones which there didn't seem to be. The highest I found less clear, the line gave me some idea but I found myself reading the TAB block data far more and it took more time of my task.*

42 X1 **E.38** *It's necesssary in 2D to scan all the data blocks to gain the above info. Therefore it becomes a reading task and not one that can be perceived at a glance.*

43 X2 **E.39** *It's harder when similar sized down [drop] lines are further apart.*

43 X2 **E.40** *Also confusing when one of the highest ones you are considering is close to another one when you find the height similar to another one some distance away.*

44 X3 **E.41** *I did not find the length of the line to be of any help. I found scanning the picture much quicker.*

45 X1 **E.42** *Labels were easy to read.*

45 X1 **E.43** *Obviously easier when aircraft were at whole flight levels.*

46 X2 **E.44** *When poles close together, quite easy: but judgment across a screen is quite difficult, especially when poles are almost identical* $\pm 200ft$.

46 X2 **E.45** *You rely quite heavily on data block when you narrow your choices down to maybe 2 or 3.*

47 X3 **E.46** *Using the information block made this task quite easy.*

47 X3 **E.47** *Initially spotted the relative highest and lowest and confirming them using the block information.*

49 X2 **E.48** *Tended to use the mode C readout as opposed to the length of the line.*

E.2 Horizontal Proximity Extraction

These are responses to the instruction '*Please write any comments which you have about the ease of the display for picking out the horizontally closest aircraft.*

1 N1 **E.49** *Very difficult to distinguish between two pairs of aircraft with similar distances (unlike height).*

2 N2 **E.50** *Would have been easier in 2D.*

3 N2 **E.51** *Most problems were easily detectable. Only one where I had to check harder.*

4 N1 **E.52** *The system can select itself those two [closest aircraft] and draw a line between them showing those [that] are closest.*

8 N2 **E.53** *"Projection" effect due to oblique viewpoint seemed to introduce some distortions in the first exercise.*

10 N2 **E.54** *It was difficult when they are spaced more evenly.*

11 N1 **E.55** *It will be very easy except for a few pairs that appear equally close to each other*

11 N1 **E.56** *I find doing it faster then the previous task [altitude extraction].*

14 N1 **E.57** *Only one or 2 confusing cases, but otherwise straightforward.*

15 N2 **E.58** *It seemed easier than I thought it would be (and that worries me! :-))*

19 N3 **E.59** *V. difficult to judge distances, the range circles were a bit distracting.*

20 N1 **E.60** *Determining the closest horizontal aircraft is harder since they all lie at some vertical offset to one another, and to rapidly determine which is closer is not so easy.*

21 N2 **E.61** *No difficulty reading the display.*

23 N2 **E.62** *This was slightly more difficult than the heights because of the lack of information on the ground-plane.*

25 X2 **E.63** *Unless a/c were at same heights it's not something I would normally look for.*

26 X1 **E.64** *Very easy except when targets very close together.*

27 X2 **E.65** *Task became easier/more difficult depending on the relative positions compared with the range rings (i.e. easy if on a radial).*

27 X2 **E.66** *Drop lines did not clutter display or make the task harder.*

28 X3 **E.67** *This was a very similar principle to 2D display.*

29 X1 **E.68** *The aircraft were very prominent and the display was only referred to when I needed to refer to the range rings to confirm my decision.*

30 X2 **E.69** *This is similar to the present-day system, ignoring the data blocks & drop-lines, with the added difficulty of judging distance on a non-linear display.*

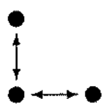
30 X2 **E.70** *Several instances occurred when 2 points were possibly the closest, and as they were aligned at different axis a comparison of distance by my 2-D brain was quite hard.*

30 X2 **E.71** *Once again I believe practice in judging these 3-D distances will improve my ability to achieve accurate distances in every case.*

33 X2 **E.72** *Similar comments to exercise 1 [task 1], trying to gauge distance when the subjects were at angles to my reference points i.e. the range rings.*

34 N2 **E.73** *Straightforward.*

35 X3 **E.74** *More difficult with similar distances yet 2 north-south with third east-west*

i.e.:  *same dist. in plan?*

36 X1 **E.75** *The display showed clearly the positions of a/c enabling one to easily assess proximity.*

37 X2 **E.76** *Visually used both the rods and the white blips.*

37 X2 **E.77** *Quicker than height definition [selection].*

38 X3 **E.78** *The oblique view requires more thought and is consequently slower to use.*

39 X1 **E.79** *Two were not too easy to decide upon. No different to real life [i.e. current ATC displays].*

40 X2 **E.80** *3D element exaggerates the complexity of the task as a/c are behind each other as opposed to left & rt.*

41 X3 **E.81** *Again I found the variant range rings not that helpful.*

41 X3 **E.82** *Although I was drawn easily to look at the white blobs I found that made me take the mouse with me and so I had to make another effort to then move up to and click the black dots—more time than recognising in some easy cases.*

42 X1 **E.83** *Again experience helps—but it a straight visual perception which can be seen at a glance.*

43 X2 **E.84** *Position on the radar again confuses judgment i.e. off axis or off centre*

44 X3 **E.85** *Once again when the aircraft were side by side I found it harder to judge.*

45 X1 **E.86** *Good contrast helped.*

45 X1 **E.87** *Aircraft near the centre of the display were easier to judge.*

46 X2 **E.88** *'Closest' ground spots on the screen not necessarily closest in reality. When ±2 miles it's quite a jiggle to get both evaluated against each other.*

47 X3 **E.89** *As with the first exercise [task 1] working out the distances when a/c were tangential [to range rings] was quite a problem and led to hesitation when choosing the closest pair.*

49 X2 **E.90** *Again, as the display is at an angle, it's not always easy to see who is close together when there are a couple very similar.*

E.3 Additional and Miscellaneous Comments

6 N1 **E.91** *The dots marking the aircraft could be larger. The current size makes selection fiddly.*

9 N1 **E.92** *Problems with mouse buttons due to being left handed.*

10 N2 **E.93** *It is a bit confusing in task b [proximity extraction] having to look at the white dots [ground plots] & click on the black ones [air plots].*

10 N2 **E.94** *The vertical lines sometimes mask a databox making the nos unreadable.*

11 N1 **E.95** *I guess finding it easy or hard depends on how fast you demand yourself to do it.*

12 N3 **E.96** *I think the problem is because of the different scale between the x and y axes on the ground.*

16 N3 **E.97** *It would be more fun if the land wasn't flat.*

17 N3 **E.98** *As the pointer was transparent I kept missing the black selection dots (the selection dots were quite small anyway.*

17 N3 **E.99** *The mouse was a bit sticky so was sometimes difficult to move properly.*

20 N1 **E.100** *Selecting the aircraft at speed is not easy since the aircraft "dots" you have to clock on are very small. Apart from this, I have no problems with the display or tasks.*

29 X1 **E.101** *I found the mouse very cumbersome because it is not moving as quickly or accurately as I would like it to.*

38 X3 **E.102** *Found difficulty in accurately positioning mouse for a/c selection.*

38 X3 **E.103** *Label overlap slows things down even further.*

45 X1 **E.104** *I tended to use my ATC experience in this task. I naturally would assume that aircraft near the airports would be lower than those further away.*

45 X1 **E.105** *I also tended to double check my answers before proceeding.*

47 X3 **E.106** *With the glasses working correctly it was easier but visualising/imagining the 3D effect was still not immediately easy.*

Appendix F
Task 3 Subject Comments

F.1 Memorisation

These are responses to the question '*Please try to describe how you went about memorising the scenes. For example, did you remember position by the patterns the groups formed, or because they resembled traffic patterns; did you remember the heights by visualising them in your head, or did you rely on memory of the numbers.*'.

1 N1 **F.1** *Shape of point arrangements.*

1 N1 **F.2** *Tried to fit outline of the shape to a height gradient (a 3D curve) to remember heights if only a few.*

1 N1 **F.3** *Some tags are easy to remember BA=British Airways, AF=Air France etc., but more than 4 say is impossible.*

1 N1 **F.4** *Number of tags [i.e. numeric part of flight codes] is very hard since there is no pattern.*

1 N1 **F.5** *On last couple I gave up remembering height gradient & tried to clump into 100, 200, 300, but failed miserably.*

2 N2 **F.6** *Looked at pattern on the ground.*

2 N2 **F.7** *Tried to group planes.*

2 N2 **F.8** *Tried to round off heights (to tens). Try to remember size of lines.*

2 N2 **F.9** *Tried to remember codes according to airlines/countries/airports e.g. LH = Heathrow AL = Alitalia.*

2 N2 **F.10** *Too many strategies to do at one go. I thought position & height were the important things to remember.*

3 N2 **F.11** *I visualized both dropline and visual position.*

3 N2 **F.12** *Had trouble remembering numbers. Kept switching between Dutch and English (Dutch no. 87 is spoken like 78, English other way around).*

4 N1 **F.13** *I started making pattern. Or doing some simple calculations, and considering similarities.*

4 N1 **F.14** *And making words with initial characters.*

5 N2 **F.15** *I remembered position by patterns on the ground.*

5 N2 **F.16** *I couldn't remember any numbers.*

5 N2 **F.17** *If the task was just to memorize the positions and relative heights (i.e. higher or lower) it would be much easier.*

6 N1 **F.18** *For dot position I used the relative position of the dots in relation to the back ground marking. For the rest I relied on number memory.*

7 N1 **F.19** *The easiest thing (if there was such) was to remember the position (I* hate *the numbers) memorising the pattern.*

7 N1 **F.20** *Identification: I tried to remember it as a word. (first part e.g. letters).*

7 N1 **F.21** *For any number I try to make some connection between neighboring two. Usually I remembered the first digit but the rest... not really.*

7 N1 **F.22** *The most difficult was to remember the second part in the identification (the number).*

7 N1 **F.23** *If I had, for example, 3 scenes with 7 aircrafts the first memorising (e.g. the memorising of the first scene) was horrible; last two were much better (I hope)—so, you need a time to accustom yourself.*

8 N2 **F.24** *Positions: Usually by map features. Occasionally by angle around centre. Combined with patterns of aircraft on ground as simple geometric shapes (lines, triangles etc).*

8 N2 **F.25** *Height: Purely numbers. Except very low aircraft which stuck in mind as short 'twigs'.*

9 N1 **F.26** *Tried to join the dots to make pictures*

9 N1 **F.27** *then id letters by mnemonics*

9 N1 **F.28** *tried to make patterns out of numbers.*

10 N2 **F.29** *I first tried to position the aircraft on the ground relative to the markings.*

10 N2 **F.30** *I then tried to create an ascending pattern of heights within each group, trying to remember the lowest & highest values.*

10 N2 **F.31** *Only when there were very few* < 4 *aircraft did I try to remember any identification.*

10 N2 **F.32** *Easier values seemed to be the low numbers grouped on the landing (shaded) area.*

11 N1 **F.33** *For position I use 2 methods:*

1. *remember the pattern (or clusters of patterns) of their positions.*
2. *remember their positions by the range rings & video map.*

11 N1 **F.34** *For aircraft ID, remember the airline (2 alphabets). If there is time, then remember the number.*

11 N1 **F.35** *For height, which I remember last, remember the number (if at all).*

11 N1 **F.36** *Also sometimes I tried to take a mental picture of the whole scene and try to replay them.*

12 N3 **F.37** *Position: Not idea of traffic pattern. Tried to remember how many in each circle.*

12 N3 **F.38** *Heights tried both:– visualise the height relation of the planes* and *tried to memorize the numbers.*

12 N3 **F.39** *For identification: characters were ok (sometimes) but rest impossible.*

13 N2 **F.40** *I cannot find any pattern.*

13 N2 **F.41** *I find that it's easier to memorize jet craft's id if I know the identification e.g. BA.*

13 N2 **F.42** *It's easier to memorize crafts that are on certain landmark.*

13 N2 **F.43** *I never train myself to memorize. (I do not believe in mental arithmetic, I always write things down).*

14 N1 **F.44** *I remembered positions first by locating near landmarks.*

14 N1 **F.45** *I tried memorising heights by visualisation but it didn't help too much.*

14 N1 **F.46** *ID number was near impossible to remember.*

15 N2 **F.47** *Initially brute force (e.g. 3 planes).*

15 N2 **F.48** >5 *planes—tried to spot patterns—notice some where flight take off routes & had a direction and had increasing height.*

15 N2 **F.49** *also tried to group planes into small blocks of 3. Therefore remember 3 planes position & relative height. Then remember the relative positions of blocks of planes to one another.*

15 N2 **F.50** *Note this didn't work for ID's 'cos to me they were a random bunch of numbers.*

15 N2 **F.51** *If I was familiar with BA as British Airways, SK as Scandinavian Airways it would have been easier.*

15 N2 **F.52** *Also counted no. of planes in scene just to check.*

16 N3 **F.53** *Positions by overall picture in my mind where objects were on the screen.*

16 N3 **F.54** *Together with planes next to each other that formed words (e.g. BA and SK).*

16 N3 **F.55** *Numbers got tricky at the end and so it was just 'photography' which in the beginning I was learning which one went with which plane.*

17 N3 **F.56** *When there were more than 4 objects I first counted the number of objects to ensure I would try to recall all the objects. Then I set about memorising the objects using the following priorities (1 is most important)*

17 N3 **F.57** *1. Position—try to remember general position on screen and any 'land marks' nearby, such as ring positions and straight line intersections.*

17 N3 **F.58** *2. Height—generally difficult to remember, just continually went round each object in the hope I would remember once I had the position drawn on the map. But, interesting neighbours such as 170, 180, 190 stuck in my mind more and so more easier to remember.*

17 N3 **F.59** *3. Airline. This is as with (2) except I seemed to remember BA batter as it caught my attention.*

17 N3 **F.60** *When drawing I tried to draw all the positions first (with any information that immediately sprang to mind) and then would look to see if the positions I had drawn triggered recall of any information (if so I wrote it down, obviously).*

18 N3 **F.61** *First I tried to memorise the no. of aircraft & their positions on the map by patterns.*

18 N3 **F.62** *Then their identifications, although it was much easier remembering the letters rather than the nos.*

18 N3 **F.63** *Finally their altitudes by memory of the nos., not their drop line heights.*

19 N3 **F.64** *Started by trying to remember positions in terms of groupings of aircraft and their relative positions to the grey 'blobs' on the screen.*

19 N3 **F.65** *Then tried to memorise heights by using numbers and attempting to associate these to the spatial groupings.*

19 N3 **F.66** *When differences in height were great tried to remember in terms of "3 aircraft over 200, 3 under 100" type way.*

19 N3 **F.67** *Identification no.s were virtually impossible to remember as they were so long. Could only remember letters (if anything) and usually only if the letters meant something to me (e.g. BA "British Airways"!!)*

20 N1 **F.68** *I memorised the positions by remembering where in relation to roads, and contrasting colours (light grey areas, shaded areas etc.), as well as (where possible) the pattern the aircraft made.*

20 N1 **F.69** *In the earlier displays, memorising the heights and ID's was just a matter of repeating them over and over, but as the number of aircraft increased, it became bad (in the time limit) to attempt to recall ID, heights and positions of all aircraft.*

20 N1 **F.70** *Instead I just tried to recall the positions of the aircraft, and if possible some of their heights.*

21 N2 **F.71** *Position was the easiest to recall as I used the 5 mile spaced lines to help me plus I counted how many aircraft in each 5 mile spacing, remembering the exact position at the beginning was not difficult.*

21 N2 **F.72** *The letters in flight numbers I recalled by association. BR—British Railway BM British Museum. The numbers in flight I recalled by (historical) date association.*

21 N2 **F.73** *I did not develop a good strategy for recalling the height numbers.*

21 N2 **F.74** *As the task became more difficult I did not bother memorising the numbers at all.*

22 N1 **F.75** *Using the circles for guidance I tried to remember how many were in each. Then I tried using the other map features to place them more accurately—but when I worked on paper this actually confused me more.*

22 N1 **F.76** *Possibility of remembering the alpha codes of the aircraft but not the numbers*

22 N1 **F.77** *Tried to remember the heights in relationship to the chars but as soon as the screen blanked out generally forgot.*

22 N1 **F.78** *It became more difficult as the numbers of aircraft increased.*

23 N2 **F.79** *I tried to visualise all aspects of the task—position, height, id. As the number of aircrafts got larger this proved very difficult. I therefore found it better just to try to remember the* relative *positions.*

23 N2 **F.80** *I also tried to relate (and find a pattern) between the heights. I tried to make use of the concentric lines.*

23 N2 **F.81** *The id was difficult to remember—I tried this in two ways—first memorising the characters then the numbers (per aircraft).*

23 N2 **F.82** *I think this task becomes very difficult (that is memorising all aspects) after 3 aircrafts.*

24 N1 **F.83** *Memory of overall numbers of a/c.*

24 N1 **F.84** *ID codes by airline code, the associating height with code.*

24 N1 **F.85** *Position of a/c using markings as reference.*

25 X2 **F.86** *Position: because they resembled traffic patterns.*

25 X2 **F.87** *Flt/no: by memorising in patterns.*

25 X2 **F.88** *Hts: virtually impossible in the later exercises to remember heights.*

26 X1 **F.89** *With a few tracks identification/height easier.*

26 X1 **F.90** *Later on in rests relied on groupings of aircraft i.e. no. in group rotating clockwise through compass points.*

26 X1 **F.91** *Also (sometimes) counted no. of BA aircraft also tried to use no. of aircraft above a certain level.*

26 X1 **F.92** *In a real ATC environment a/c would start 1 000′ vert. sep. therefore retention of levels would be easier. The more random levels the more difficult the exercise.*

26 X1 **F.93** *Some 'backtracking' to previous exercises was evident.*

26 X1 **F.94** *Also relied on no. of a/c around a given range ring.*

26 X1 **F.95** *Counted no. of a/c in an exercise.*

26 X1 **F.96** *The later exercises gave me a sense of failure as skill level dropped!*

27 X2 **F.97** *Attempted use of mnemonics.*

27 X2 **F.98** *Airlines I knew (fly[ing] into LHR) were easier.*

27 X2 **F.99** *Tried to remember patterns formed.*

27 X2 **F.100** *I found drop lines and data tag for a particular a/c confusing when (because of height) the data block positions were opposite to geographical positions of a/c.*

27 X2 **F.101** *All numbers remembered (or probably not!) in my head.*

28 X3 **F.102** *Most of the aircraft positions I tried to memorise using normal traffic patterns for example left [h]and circuits for [runway] 26 at Gatwick. Easterlies Heathrow Lambourn hold.*

28 X3 **F.103** *I tried to remember heights by seeing if there was anything unusual in the heights of two a/c relative to their positions.*

28 X3 **F.104** *When I knew I would have difficulty remembering everything I reduced it towards two letter designators, positions and relative heights to each other.*

29 X1 **F.105** *I found it very difficult—the letters and the numbers were difficult to memorise.*

29 X1 **F.106** *The heights were easier to memorise if the aircraft were close.*

29 X1 **F.107** *The more aircraft I had to remember the harder the task.*

29 X1 **F.108** *I tried to remember the positions by their relation to the map.*

29 X1 **F.109** *I just found the 'callsign' difficult to memorise at all.*

30 X2 **F.110** *I used traffic patterns where possible to remember position.*

30 X2 **F.111** *Heights were memorised (using the traffic patterns as a guide). When no patterns existed I used pairs of similarly high aircraft where possible.*

30 X2 **F.112** *As a back-up I counted the number of aircraft and attempted to memorise their [geometric] pattern.*

31 X3 **F.113** *Position by patterns–yes. Resembled traffic patterns—partially. Position also by reference to a common datum mainly the centre of the screen.*

31 X3 **F.114** *Heights—memory.*

32 X1 **F.115** *Position relative to video map and adjacent aircraft.*

32 X1 **F.116** *Heights by memorising numbers & in number order if possible.*

33 X2 **F.117** *I found the scenes which related to Heathrow traffic patterns the easier (easy being a relative term!!). If the callsigns had been 'regular' ones then I think I'd have fared better with the positions. When first faced with a number of 'unfamiliar' callsigns I found the task harder.*

33 X2 **F.118** *With the more random scenes I tried to establish positions by reference to the video map rather than the range rings. Callsigns and height purely by memory (not very well!). To go back to comment (1) (Quote F.117) if the callsigns had been 'familiar' and in a familiar traffic pattern then I'm sure I could have recalled more. Also I might well have remembered more height information as I wouldn't have had to keep trying to remember the callsigns.*

34 N2 **F.119** *Position: By relating to landmarks & distance circles and by grouping where applicable.*

34 N2 **F.120** *Height: Generalising for aircraft at similar heights. Number memory (difficult for 3 figure number.*

34 N2 **F.121** *Flight numbers: Letter pattern by associative memory BA≡British Airways etc. Number by memory only made easier if configuration memorable e.g. 468, 137 etc.*

35 X3 **F.122** *Related ground positions to known geographical locations rather than range [and] bearing.*

35 X3 **F.123** *Visualised a/c in company logo in sky. Considered (unwittingly) flt. no. less important.*

35 X3 **F.124** *Tried to match height to real life traffic pattern. If this tallied it made memorising heights much more easy (one ex. [example] relates to Gatwick in particular). Heights then remembered if poss.*

35 X3 **F.125** *Callsigns remembered more so if realistic in realistic location or if a numeric 'pattern' e.g.:* IB633
63

35 X3 **F.126** *Easier to remember 'lone' a/c.*

35 X3 **F.127** *Easier to remember whole heights e.g. 330, 370, 390. Other visualised in head 'in sky' and began to guess.*

35 X3 **F.128** *Conflicts more likely to be remembered.*

35 X3 **F.129** *Traffic patterns more likely to be remembered.*

35 X3 **F.130** *Scattered and unrealistic [random]* very *difficult.*

36 X1 **F.131** *With 3 a/c on the display I was able to photographically memorise the screen. With more a/c this was not possible. I began to memorise position which was the easiest, by trying to visualise the positions in my head, followed by the heights, picking out the heights which to me were the easiest to remember (i.e. those next to easiest to recall positions/lower numbers).*

36 X1 **F.132** *The callsigns were very difficult to recall for more than 3 a/c on the display.*

37 X2 **F.133** *Conflicting heights come first. Familiar traffic patterns/allied to those heights.*

37 X2 **F.134** *Random patterns hopeless.*

37 X2 **F.135** *Height rods of no use whatsoever.*

37 X2 **F.136** *A real callsign is easier IB633/BA324 than a phantom. Qualified ATCO, not new boy.*

38 X3 **F.137** *Tried to use callsign and positions relative to places on the ground.*

38 X3 **F.138** *Absolutely lost without some other form of reference (strips).*

38 X3 **F.139** *In real life the whole control sequence, RTF [radio telephony], strips, future plan, known procedures and set routes all help to form 'the picture'.*

39 X1 **F.140** *Position in relation to airspace boundaries. The rest had to rely on raw memory. Not easy.*

40 X2 **F.141** *Initially with few a/c constant scanning of the 3 components [position, callsign, height] for each a/c helped to make 'it stick'. With more a/c initially no discipline—distracted & race against clock. Finally decided to look & concentrate on position info. only on principle [that it is] better to get something right than nothing.*

40 X2 **F.142** *Perhaps if traffic was live you are more able to retain info.—if routine expected traffic then certain elements i.e. callsign become second nature.*

41 X3 **F.143** *I first identified a/c positions, relative to range rings, working out from the centre. If two were close together I took a range & bearing of one from another. Sometimes I used the airspace pattern to work out where all the aircraft were.*

41 X3 **F.144** *For heights I tried to relate it to the actual job I do where it bore some resemblance to what I would call my normal working radar image. Otherwise I tried to group a/c at similar levels.*

41 X3 **F.145** *The callsigns were difficult when more than 4 a/c present. I occasionally related the airlines easily, sometimes using where 2 BA a/c were found but found callsign numbers too much to take on board accurately.*

42 X1 **F.146** *Position resembled traffic pattern in some cases (some were more obvious than others). Familiarity with the airspace helped here.*

42 X1 **F.147** *Height remembered by a combination of visualising & memory. Again those traffic patterns which made more sense to me were easier to remember.*

43 X2 **F.148** *Working L to R bottom corner to top corner.*

43 X2 **F.149** *After initial scan working out if any callsigns etc. were familiar to me—such as BA324 Heathrow–Paris.*

43 X2 **F.150** *Also easier to memorise levels that were familiar numbers.*

43 X2 **F.151** *Tried to relate unknown callsign to made up name e.g. DM Deutsch Mark, DA Duck's ... etc.*

43 X2 **F.152** *One pairing was a CO callsign Continental [Airlines]. I remember reading of an airmiss at the level he was at including in Continental aircraft.*

43 X2 **F.153** *If none of the above worked tried to relate pairs i.e. level 35 and level 85. Also a couple of levels were just a few numbers different.*

43 X2 **F.154** *Found that unfamiliar letters/numbers even though I tried to put them into memory were quickly forgotten.*

initial scan working out if any callsigns etc. were familiar to me—

44 X3 **F.155** *I tried to find similarities or trends in any of the callsigns or altitudes e.g. IB633 63 or for altitude 70, 80, 90 etc.*

44 X3 **F.156** *I tried to remember positions by where they were on the map e.g. I would concentrate on a position [near a boundary] rather than one which was 35 miles out.*

45 X1 **F.157** *It was much easier as an approach controller to remember and visualise traffic in an approach stream and also traffic that was in a logical sequence of climbs and descents near an aerodrome.*

45 X1 **F.158** *I also tended to focus on the callsigns I recognised and on the flight levels rather than the idents.*

45 X1 **F.159** *I felt that the range rings were not bright enough for quick judgment.*

45 X1 **F.160** *I am used to dealing with aircraft below FL150 so aircraft above that level were more likely to be ignored.*

45 X1 **F.161** *Tended to try to remember numbers by their positions on the screen. Therefore half memory/half visualisation.*

46 X2 **F.162** *Position was measured relative to video map. i.e. points on zones, airways etc. (some reference to range rings). Then some were remembered as patterns of aircraft relative to each other.*

46 X2 **F.163** *Company signs were remembered as aircraft idents—remembering which callsign was where.*

46 X2 **F.164** *Heights were remembered as numbers & little or no use was made of the 'drop line' apart from extreme cases of very short or very long.*

46 X2 **F.165** *The few to start were relatively easy—to remember callsigns, heights etc.—but as more and more arrived, positions took overall priority with level, and farther on, positions and callsigns were left.*

47 X3 **F.166** *The first thing I did was count the number of a/c on the screen. Then I tried to recognise any obvious traffic patterns or approach sequencing and their relative levels.*

47 X3 **F.167** *Memorising the company designator letters was reasonably easy but when even slightly 'overworked' the callsign number was ignored.*

47 X3 **F.168** *Only relative heights of a/c close or abeam each other could be recalled, any periphery a/c i.e. beyond 40 miles were also generally the first to be 'ignored'.*

47 X3 **F.169** *I found this exercise very difficult and quite annoying!*

48 N1 **F.170** *Position: tried to remember patterns.break them down into pairs/threes.*

48 N1 **F.171** *I.D.: by memory of no.s; letters by sounds/mnemonics by stringing them together.*

48 N1 **F.172** *Height: by memory of no.s.*

48 N1 **F.173** *Tried to remember all that was presented—easy for 3 or 4 aircraft, but got much harder with 6, 7,...*

48 N1 **F.174** *Had difficulty transferring pos'ns back to the paper (because of different scale? black & white?).*

48 N1 **F.175** *Different sets of numbers tended to get muddled in my head... Also, got harder as they were more spread out (I suspect this was largely psychological!...).*

49 X2 **F.176** *I tried to remember position patterns, followed by height and then callsign.*

49 X2 **F.177** *Only in the first couple of exercises [the 3-aircraft scenarios] did I have time to look at all the information and even then when the screen went blank some of the information also did.*

49 X2 **F.178** *For an ATCO's point of view if the callsigns had been familiar then the task would have been also.*

49 X2 **F.179** *In a working environment with other information on the radar display along with regular traffic patterns i.e. SIDs/STARs and reporting points the task would have been slightly easier.*

F.2 General Comments

These are responses to a question asking for any further comments.

4 N1 **F.180** *It is difficult for a person who has been busy for doing some brainy work. It needs relaxation and deep concentration.*

5 N2 **F.181** *I think the numbers just can't be remembered. It is possible to remember the approximate positions and relative heights.*

8 N2 **F.182** *Found conversion from oblique view to planar (paper) view very difficult.*

8 N2 **F.183** *Abandoned aircraft numbers completely once* > 3 *aircraft!*

10 N2 **F.184** *There is again a problem of masking of the display.*

10 N2 **F.185** *The flight heights seemed particularly important.*

10 N2 **F.186** *It is hard to remember much when the no of aircraft is* > 5.

12 N3 **F.187** *Maybe some color coding would make memorising easier.*

15 N2 **F.188** Hard!

16 N3 **F.189** *I was continuously trying to find a pattern (i.e. numbers adding up or some sort of code that would work most of the time). P.S.: I don't think there is one.*

17 N3 **F.190** *I would like a graphic indicator of the amount of time left as sometimes I was looking closely at something and the screen went black which was quite shocking (or maybe the screen could fade out).*

17 N3 **F.191** *I did find it very difficult to remember the flight numbers, especially when there were lots of aircraft.*

19 N3 **F.192** *The information thing at the top sometimes obscured the drop line and having to click to make it transparent was a bit tedious.*

19 N3 **F.193** *Sometimes it was not possible to make the thing transparent in order to see the drop line.*

26 X1 **F.194** *Familiar callsigns helped with remembering the scenes.*

26 X1 **F.195** *Worthy of note most company designators are now three letters i.e. BAW.*

26 X1 **F.196** *Colour banding may have helped in remembering rough levels.*

26 X1 **F.197** *Quite a demanding exercise as ATCO's rely on aide-memoirs of mistakes happen!*

27 X2 **F.198** *V. difficult.*

27 X2 **F.199** *I may have sometimes used 3 letter codes in my answers (e.g. BAW instead of BA) as this is what I'm used to.*

28 X3 **F.200** *I hope there were no potential incidents.*

29 X1 **F.201** *Very difficult.*

30 X2 **F.202** *Where there was an obvious traffic pattern it was easier to memorise 'the picture'. 3–4 aircraft displayed at random were memorable but more than this number*

caused me difficulty. The traffic pattern examples allowed me to retain a picture of more aircraft.

31 X3 **F.203** *Fairly easy with low numbers of traffic. Very hard with more. Displays & operation no problem.*

34 N2 **F.204** *Three planes reasonably easy to remember.* > 5 *much harder, only position & general heights could be grouped.*

35 X3 **F.205** *Became difficult beyond 4 a/c.*

41 X3 **F.206** *Not helpful having to reveal data block as memory of those near each other I found very poor when recollecting.*

42 X1 **F.207** *My brain only has a limited capacity therefore information was filtered out that was considered not essential to the ATC picture i.e. the callsigns of the aircraft are not immediately relevant unless you have to make a transmission to that a/c, although the operating company was remembered (?) as it helped with the picture.*

42 X1 **F.208** *Long practice at accepting handover from one ATCO to another in busy ATC positions dictates that the traffic picture is gained as rapidly as possible.*

43 X2 **F.209** *I quite often do GMC (Ground Movement Control) working 10–15 aircraft without much reference to my Flight Data Strip because I am always looking out of the window—it's a visual 'out'. Most of the callsigns when doing this are familiar and are usually retrieved from memory quite quickly due to the speed of the task.*

44 X3 **F.210** *Far too many numbers to remember—slightly confusing as I am used to 3 figure [letter] callsigns not two so would have to concentrate on the 2 figures [letters] as well. Hard work!*

45 X1 **F.211** *I usually have good memory recall but it seemed to desert me today. I would normally use visualisation to solve the problems and also try to link similar levels and callsigns by their position on the screen. I am predominantly an aerodrome controller nowadays; perhaps some of my radar skills are deserting me!*

47 X3 **F.212** *This truly highlights how bad my short term memory is and how small my capacity for remembering things that would normally be beyond my field of control area.*

Appendix G
Task 4 Subject Comments

G.1 Conflict Detection

These are responses to the question '*Please try to describe how you went about doing the conflict detection task. For example, did you look at pairs of aircraft, first of all looking at ground position, then referring to height? Did you take speed into account?*'.

This question was added to the instruction sheet only after subject 6 (subjects 1–10 were pilot subjects) so comments are unavailable for these subjects.

7 N1 **G.1** *I tried to detect the closest ones looking at ground position, then I did refer to the height. The speed was the last to take into account.*

Tracking [i.e. trailing histories] helped a lot.

8 N2 **G.2** *First check all FL's for possible problem. Then look at ground position for close pairs. Then check climbing/descending aircraft. By then sufficient trailing dots to be useful—so survey whole scene for anything else.*

9 N1 **G.3** *Firstly identified any aircraft within 3 miles of each other.*

Then checked heights and looked for trends i.e. going up or down. Tried to identify those going beyond clearances.

10 N2 **G.4** *After identifying potential pairs, I first of all checked heights. I found that when the aircraft were well spaced I used the height of the black line but when they were close I had to use the no's given.*

When aircraft looked at the same height I checked their ground separation & the direction of travel. The previous position dots were very useful for this.

11 N1 **G.5** *First before the aircraft begin moving, I quickly look at positions of aircraft on the screen and try to check altitude of those close to each other.*

Then as they begin to move, I try to anticipate their paths, and look out for those crossing. Then I check their altitudes and whether they are ascending or descending, and whether a conflict could occur. I'll check speed after this.

12 N3 **G.6** *I didn't take speed very much into account. First I looked at the heights at the given time and then at their intended heights. Then a better look at their horizontal*

distances and their directions.

13 N2 **G.7**

1. *1st pay attention to aircrafts close together.*
2. *Then check their height.*
3. *Monitor the changes.*

14 N1 **G.8** *1st check all the heights, and look for potential conflicts. Then check position and the direction/velocity.*

15 N2 **G.9**

- *Got approx. positions from ground dot.*
- *Checked out current height & heading from dot trail on ground.*
- *Tried to find potential conflicts.*
- *Then checked if one or other plane was ascending/descending into other's flight path.*

16 N3 **G.10** *I didn't remember that speed was involved. I just checked which 'white' paths [i.e. ground projections of aircraft histories] looked like they were going to cross, and then cross-referenced with the height. So only if it was close enough horizontally did I check for vertical distance. (And vice versa i.e. if they were all close then I went for height.)*

17 N3 **G.11** *First I looked for similar heighted pairs of aircraft.*

If these aircraft were close together I checked if they were going up or down, if they appeared to be converging I would stop the simulator.

If there were no initial problems I watched the simulation to attempt to determine where aircraft would go based on the trails left by the aircraft and the ↑ or ↓ status, and their cleared ceilings, and if they would meet any other aircraft . If I thought they would meet I stopped the simulator.

18 N3 **G.12** *First I tried to check which a/c were within 3 miles of each other, then their heights → whether they were likely to conflict depending on '↑' or '↓' & their cleared altitudes. If the trails left showed there might be a conflict, then I stopped the screen.*

19 N3 **G.13** *Check on positioning first to rule out aircraft immediately. e.g. one aircraft way out on its own could be ignored for the time being.*

19 N3 **G.14** *Then determine which aircraft were ascending, descending or flying level. Determine from positioning & direction possible conflicts i.e. paths going to cross in future, then look at heights of aircraft involved & whether they were ascending or descending to determine a possible conflict.*

20 N1 **G.15** *First I examined the current heights, and intended new heights. Then I noted their direction and velocity. From this it was easy to spot whereby the aircraft may have "lost separation" from one another.*

21 N2 **G.16** *I tended to look first at speeds first then ascending/descending/level aircraft, then pairs of aircraft that were/seemed to be closing. I disregarded aircraft in clear airspace first then concentrated on the others. I learned to read speeds, height of aircraft* together *v. quickly.*

22 N1 **G.17** *Looked for closely positioned dots checked altitude then watched the direction of travel. Didn't really notice the speed issue consulting only the convergence of the dots.*

23 N2 **G.18** *I first eliminated those aircraft who were completely clear i.e. by separation. At the beginning of the task I was initially reading the heights to start with, but towards the end I started looking at the ground plane paths. When the aircraft were close the overlap of the datablock was off-putting.*

24 N1 **G.19**

- *Look for aircraft ascending/descending*
- *Look at current height of a/c where headings cross*
- *Where height conflicts may occur, observe relative approach speeds*

25 X2 **G.20**

1. *Looked at air position of a/c closest to one another.*
2. *Checked heights.*
3. *As trail dots appeared checked track.*
4. *No info on a/c type given so speed was not considered.*

26 X1 **G.21** *Scanned for vertical separations. Slight delay whilst tracking information built up. Exercise made slightly more difficult whilst checking for intentions of tracks which overlapped.*

27 X2 **G.22**

1. *Look at heights for similarities (datablock).*
2. *Observe tracks/speeds of possible conflictions noticed in (1).*

28 X3 **G.23** *First looked at all heights to spot potential confliction. Then assessed position, track, speed to see if potential confliction existed as real confliction.*

29 X1 **G.24** *I looked at all the aircraft to ascertain their height. I then looked at the aircraft close together to ascertain their planned descent or ascent and spotted the*

conflictions.

30 X2 **G.25** *I looked at heights to build a mental picture, with the displayed plan information. As the trail dots appeared I projected the tracks, using speed information, to foresee any conflicts. As track conflicts appeared I reviewed the height data, looking for actual conflicts. I continually scanned each aircraft for any change in height/course. The labels became a problem towards the end, as denser traffic appeared.*

31 X3 **G.26**

1. *Pairs—yes*
2. *Position–yes*
3. *Height–yes*

} *in that order*

Speed quite hard to judge until trail dots appeared displaying headings of targets.

32 X1 **G.27**

1. *Look at a/c with similar heights.*
2. *Observe display to determine their relative direction of travel.*
3. *Continue to monitor display to see if there were any changes of height (i.e. if one commenced climb or descent.).*

33 X2 **G.28** *I looked at ground positions and then at the most adjacent aircraft height block. I found it difficult to take speed into account with these displays. After my initial 'scan' of the picture I then turned my attention to what the aircraft were going to be doing in the next few minutes referring to level and heading. Those which were 'safe' and would continue to be I 'discarded' for the present, those which had the potential of conflict by virtue of heading or altitude I assessed before making a decision and then I returned to a scan of the complete picture to ensure that the aircraft previously thought of as 'safe' were still 'safe'.*

34 N2 **G.29** *Checked heights & positions. Looked for similar heights where there were groups of aircraft. Waited for while to judge speeds. Kept eye on potential conflicts. Assessed what the aircraft doing i.e. going up or down to whatever height. Made sure they were doing what they were meant to be doing.*

35 X3 **G.30**

- *1st, look at alts, actual level*
- *2nd look at cleared level*
- *3rd assess track & rate of closure*
- *4th action to deconflict*

36 X1 **G.31** *The display was not immediately clear—the* highlighted *information (i.e. dat-*

ablocks) is distracting. It would be easier to read if not highlighted.

36 X1 **G.32** *First of all I looked at heights doing a quick scan of the screen—obvious if 2 a/c were close together I looked at their heights first. Then I scanned again to look at changes in height.*

37 X2 **G.33** *Routes, i.e. track ground position aided by route code. Flight level. Climbing & descending. (But I am conditioned to 'collision corners'.*

38 X3 **G.34** *Used present and cleared levels first, then assessed track to determine conflict. Only thought of speed in the final stage but this was difficult to differentiate.*

39 X1 **G.35** *Look at closest in plan, check vertical, then look onward from there, taking the next closest in plan.*

40 X2 **G.36**

1. *Quick examining of current levels.*
2. *Quick look at whether* $\uparrow$ $\downarrow$ $=$
3. *Discern heading.*

41 X3 **G.37** *Firstly I would check the levels at the start, if there were any less than 1 000 ft then I would check the distances, if close then what speed—i.e. eroding or increasing separation. Having established this I then checked for a/c climbing or descending and who was in the projected flight path (say for about 10/20 miles)—and final conflictions there. Then if nothing obvious monitor the display for changes which may produce a conflict.*

42 X1 **G.38** *Assessed overall traffic situation. Looked for a/c in close proximity where one of the pair climbing or descending. Read the info on the data block. Speed taken into account where the two a/c were converging at the same level.*

43 X2 **G.39** *Check levels in the first 6–12 secs. Check tracks after a couple of dots. Work out who could be discounted from a confliction with the others. Then monitor pairs who could be a problem for speed differentials, changing tracks (turns) [and] level busts.*

43 X2 **G.40** *Although the level bust was difficult to spot (as always) because I trusted the cleared level.*

44 X3 **G.41** *I look at height info first of all to see whether they were climbing, descending or level. As soon as I spotted one which may be* $\uparrow$ *or* $\downarrow$ *through another aircraft I looked at the projections in their path. Speed was obvious in a few cases.*

45 X1 **G.42**

1. *Looked at position.*
2. *Looked at direction of flight.*
3. *Decided if aircraft were ↑ or ↓ or in level flt.*
4. *Grouped aircraft by height bands to eliminate possible conflictions.*
5. *Judged speed with trail dots.*
6. *Slight uncertainty of the trails of some a/c.*

46 X2 **G.43** *Look at the whole picture—find the a/c that are safe against each other—then 'disregard' them!*

Look closely at ↑ & ↓ a/c, and levels assigned. It does become an array and blur of black and white dots quickly & as we saw, data tags can become confused etc.

Putting assigned level in data block can confuse & clutter display—i.e. I refer to putting the level on a 'strip' and manually detecting conflictions—just on a data block is not enough. It makes speeds difficult to see also.

='l speeds easy to judge.

47 X3 **G.44** *Ground position and direction of flight was taken into account once the trail dots were established.*

A/c at high levels 'outbound' were the first to be disregarded as clean i.e. not in conflict with anything.

A/c in level flight were relatively easy to spot and relate to. Also a/c in an approach sequence were more obvious.

47 X3 **G.45** *The conflictions in the climbing and descending were difficult to spot when their close proximity meant garbled labels as switching from one label to another using the mouse was not always effective in spotting the conflict precisely.*

48 N2 **G.46** *Watched for pairs of 2D posns and potentially ≤ 3 miles & kept an eye on heights. Speed—not much... (height & posns seemed to be sufficient, mostly).*

49 X2 **G.47** *Quick look at what the aircraft were doing, i.e. descending or climbing, followed by where they were and where they were heading.*

49 X2 **G.48** *As some aircraft were turning, not knowing whether they were going to roll out on a safe heading or not doesn't help. In the "real world" as you would have more idea of what the aircraft* should *be doing the task would have been slightly easier.*

49 X2 **G.49** *The labels were very cluttered on a couple of the displays and a lot of time was spent "undoing" the labels etc.*

G.2 Additional Comments

3 N2 **G.50** *Rather click on box to bring it forward than to make it invisible.*

8 N2 **G.51** *Tried several times to click on information I wanted rather than that in the way. Might be useful to be able to do this via some hierarchy or similar.*

10 N2 **G.52** *Again when the aircraft are very close the data points are obscured.*

If the aircraft are directly N/S of one another the heights can get muddled.

12 N3 **G.53** *The 3D glasses are sometimes tiring. Not too bad just a bit tiring.*

15 N2 **G.54** *Vertical distance was easier to judge than horizontal.*

Vertical conflicts were the main ones detected.

19 N3 **G.55** *Data tag things obscured the display quite badly.*

20 N1 **G.56** *The display was clear and very easy to follow and use. The only times it becomes difficult is where several aircraft are close to one another, and their information tags obscure the other aircraft "dots".*

23 N2 **G.57** *The frame rate made it difficult to assess a clear path i.e. would be better if faster.*

25 X2 **G.58**

1. *Info would normally be available on route a/c following.*
2. *Wouldn't clear a/c to ↑ to altitude above 130.*
3. *Would control a/c at high level as well as a/c at lower levels.*
4. *Overlap is a problem without aide memoir of paper strips.*

29 X1 **G.59** *The display is hard to decipher when the datablocks are overlapping. The task is enjoyable and challenging.*

30 X2 **G.60** *Larger scale would reduce label overlap. This task feels 'false' to me as I known the area when is divided into many areas, each controlled by a different person. This division is both vertical and lateral. Also, I was watching for changes in aircraft heading and/or height/cleared level. At work these instructions would be issued by me and I would only be looking for non-compliance with instructions, or someone else's aircraft wandering into my bit of airspace.*

31 X3 **G.61**

1. Cleared *levels in whole thousands.*
2. *Labels cluttered in intense areas of activity.*

35 X3 **G.62** *Still this desire to reverse colour code of air/ground (plan) position.*

41 X3 **G.63** *I found display overlap a major problem. The clicking to clear displays wasn't easy with holding situations, and I didn't like an a/c label to be totally invisible to me as I couldn't tell any changes in level that might have occurred.*

41 X3 **G.64** *Further instructions given to a/c during the simulation run weren't obvious to me and although scanning the display I had previously discounted such a/c as low priority hazards to any form of confliction.*

42 X1 **G.65** *The reading of the data blocks was perhaps the slowest part of the task. Once involved in the control task & therefore issuing instructions one would be more involved in the loop rather[?] than the monitoring situation demonstrated.*

43 X2 **G.66** *Some scenarios had garbled labels and even worse the position dots and trail dots were obscured by other* labels.

44 X3 **G.67** *Too much information was displayed making the conflictions difficult to spot. Also the garbling did not help either. I found the ground position display confusing at times.*

44 X3 **G.68** *There was nothing in the display that aided conflict detection.*

45 X1 **G.69** *Initially missed having back up of flight progress strips.*

Appendix H
Task 5 Subject Comments

1 N1 **H.1** *Direction is relatively easy since it's on absolute bearing, but height is tricky since there is no feedback on the slider so you can level out easily.*

4 N1 **H.2** *If you can do both angle and height adjustment with one shape [control] that will be more efficient.*

7 N1 **H.3** *First I tried to direct the chasers and then to see how to adjust the height. It was just a bit difficult to catch the height (to see what it the height) when the info blocks would overlap.*

8 N2 **H.4** *Used ground marks for heading, numbers for altitude.*

9 N1 **H.5** *It was fun! Again problems with mouse buttons due to left handedness.*

10 N2 **H.6** *Difficult to check altitudes when they were close due to obscuring of the data panels. Good fun.*

11 N1 **H.7** *I think there is still an element of luck when I "caught" the aircraft. It seems easier to judge/control the direction than the flight angle.*

12 N3 **H.8** *It takes a few tries to get used to it but after that it's fun.*

13 N2 **H.9** *The height is difficult to control due to the time lag.*

14 N1 **H.10** *Fairly easy to do. No problems.*

15 N2 **H.11** *Cool! It was easier to see a heading in which you would collide with a place than vary your height.*

16 N3 **H.12** *Forgot the scaling (vertical and horizontal scales) so I just chased the one that looked closest. The steering was not a problem.*

17 N3 **H.13** *I found that the hardest part was the altitude control.*

19 N3 **H.14** *Difficult to judge the direction that the target and the chaser were going in. It was difficult to visualise the space as 3 dimensional to be able to control altitude and direction together.*

20 N1 **H.15** *It is quite difficult to 'fly' the chaser and select the databox of the target aircraft if both are close, and you cannot see the target's altitude etc. Apart from this, the controls are very easy to get to grips with.*

20 N1 **H.16** *One other problem is that of level flight. It is not too easy to get your chaser to fly at a constant level.*

21 N2 **H.17** *I tended to set a direction first and then try to adjust the height, which I stayed with to try and get as close as possible, then moved back to the direction, returning to make adjustments to the height, unless there were significant changes in height. Direction was the easier of the two factors to guide and control.*

22 N1 **H.18** *Check altitudes of the 2 planes. Set chaser in general direction of the plane, adjust altitudes and then try to follow plane.*

23 N2 **H.19** *Provided the aircraft maintained the same height the task was relatively easy. The problem comes when both variables are changing.*

26 X1 **H.20** *Again tracking information took time to build. Controls were sensitive in vertical mode. Interception wasn't too difficult if the approach to target was stable. If a miss occurred close-in (in vertical sense) this then became a more difficult task.*

27 X2 **H.21** *I'm afraid I still used data block for height calculations. When only two a/c are involved it is easier to do this (and you get an* exact *figure) than it is to look at the drop lines and then compare them.*

28 X3 **H.22** *I realised at the end that when I couldn't read the data block I could use the height lines but not very accurately.*

29 X1 **H.23** *Very difficult to correlate the display and the controls on the right. I had to concentrate on the height control more than the heading control. The task got easier with familiarity.*

30 X2 **H.24** *As the exercises progressed I developed a system. 1st try to match the heights—if the target was high I used a high rate of climb. 2nd set course towards the target. Once on course adjust rate of climb as levels become equal. Then adjust heading & climb rate to follow target. If the height was wrong as the target and chaser became adjacent the label overlap problem arose again, requiring rapid and accurate cursor movements to checks chaser's height and return to controls to adjust them as necessary.*

By the end I was thinking about a 'play-the-system' solution whereby I would put the ground symbols together and then go for max. climb/descent to achieve the proximity parameters. This would make the label overlap irrelevant.

I was happy I chose the closest aircraft to the center each time.

33 X2 **H.25** *Level change very difficult (for me) to control.*

36 X1 **H.26** *The only real difficult I found was the use of the mouse and the overlapping of the displays [datablocks].*

37 X2 **H.27** *Label overlap is a problem.*

38 X3 **H.28** *Pitch control far too sensitive.*

40 X2 **H.29** *Good fun. The height was the most difficult aspect to match!*

41 X3 **H.30** *I tended to choose high a/c near the radar head as climb performance was so good. Horizontal distance was easy to assess but vertical readout was obscured so much that I was unable to ascertain targets height easily and mouse workload was so high with a/c control that I was unable to provide time/capacity to click off the labels.*

42 X1 **H.31** *Mouse control of a computer is never a very easy way of controlling/coordinating.*

43 X2 **H.32** *Controlling the chasers level rather erratic. And also watching for any change in the target course & level challenging.*

44 X3 **H.33** *I found the ground position display very confusing. I did not realise that because the aircraft was climbing the label length increasedm this drew my attention away from the position of the aircraft in relation to the target. There was too great a distance between the display & the controls on the right hand side—made it quite difficult to keep up to date.*

45 X1 **H.34** *↑ control was quite sensitive otherwise exercise was fairly straightforward.*

45 X1 **H.35** *Tended to set direction before climbing chaser.*

46 X2 **H.36** *Level control most difficult. Direction quite easy to control. This is where the drop bars came in quite useful. Instead of having to watch the levels closely, the bars could be manipulated to the point they needed to go: also, a varying of the speed of the chaser (except climbing or ↓) would be useful, but maybe quite difficult to control.*

47 X3 **H.37** *The only main problem was when the blocks overlapped and seeing the level required by the chaser to go to was not seen.*

48 N1 **H.38** *Pitch was sensitive (might be helpful to have a way of levelling off the "chaser" exactly.)*

Appendix I
VR Subjective Data

I.1 General

33 **I.1** *Found it difficult to start with to relate the movement of the "screen" to hand as finger movements on the "gun". Practice would solve this. ... once a person is completely happy with the operation of the "gun" then the possibilities would widen. ... we'd all need "education" as to the operation of it all in particular those working on the new Heathrow tower to get maximum benefit.*

36 **I.2** *The computer screen was a lot clearer than the scene through the helmet.*

37 **I.3** *Somebody with more foresight & future vision than I is needed. One day I'm sure we'll use bits of it.*

38 **I.4** *Resolution poor.*

38 **I.5** *I found moving using buttons difficult to control. Moving hand up, down, left, right would be easier.*

40 **I.6** *Obviously crude & heavy.*

42 **I.7** *A lot of pressure is taken off controlling by working in a familiar environment—ask any ATCO who has to move to new equipment.*

42 **I.8** *Personally I wouldn't like to have my head encased in a VR helmet.*

43 **I.9** *Headset/optical set would have to be lighter for practical application.*

46 **I.10** *Obviously lighter headset and quicker update rate [are needed].*

46 **I.11** *Needs to be able to let you work as part of a team and not 'immerse' the user into his 'own world'.*

49 **I.12** *Certainly very interesting but until the display is no more than a pair of glasses then ATCOs will not use it. Just the few minutes I had the display on were enough!*

I.2 Radar Visualisation Demo

30 **I.13** *Radar requires the controller to 'sit back' to a certain extent, to enable him*

to see the whole picture. The ability to 'fly' into the VR world leaves him open to missing an important item out of his field of view.

30 **I.14** *I cannot see much need for this in the present radar environment, other than the ability to fly round your area, and view things from a different angle or perspective.*

32 **I.15** *Application in radar environment is not so clear other than as an alternative view of the basic situation.*

36 **I.16** *Taking into account my unfamiliarity with the equipment, I found it very disorientating.*

36 **I.17** *Height differences were readily apparent but it was not easy to gauge rates of descent/climb.*

37 **I.18** *Rod size gives better height indication than 3D display.*

37 **I.19** *Distance more difficult to estimate.*

37 **I.20** *Climb & descent flags come over better than ^ or* **v**.

37 **I.21** *Doubtful!*

39 **I.22** *Not a replacement for radar. Don't envisage as being so even with much better technology.*

40 **I.23** *Interesting concept... The system demands concentration & is claustrophobic. Whether physically or from a physiological point of view the system is acceptable I don't know.*

41 **I.24** *The headset would not I believe be suitable for radar orientated work as your constantly changing position relative to your traffic is alien to the average ATCO's working environment where the need for a fixed image against which to assess movement and location is important in conflict alerting.*

42 **I.25** *For the purposes of a display the fact that you can move around the display & view it from different angles* could *be a serious disadvantage.*

42 **I.26** *Maybe one could get used to it & find it useful but from the radar environment [point of view] I didn't.*

46 **I.27** *The user could easily become lost around the a/c—but judging the paths could be easier in the vertical by looking along the a/c flight path.*

47 **I.28** *No obvious use in ATC operationally springs to mind...*

49 **I.29** *As far as an ATC application I feel it would be limited as if you cannot physically see an aircraft then you would have to use radar and I think a standard*

2D display or TV '3D' display is always going to be enough without the need for being 'in there' with the aircraft.

I.3 Kitchen Demo and Training

30 **I.30** *Training in the VCR could be helped by simulating aircraft activity, provided the trainee controllers instructions could be relayed accurately to the VR aircraft.*

32 **I.31** *Initial impressions would suggest that the most suitable application for ATC would be aerodrome control, enabling the student to gain a full 3D perspective of the traffic situation from the control tower. (It is a perennial problem getting students to lift their eyes up from their strips!)*

32 **I.32** *Additionally, it would enable the working environment to be modified prior to any real world changes (e.g. installation of new control desks).*

33 **I.33** *Useful aid for airspace planning and training ATCOs as to what 3D would be like.*

33 **I.34** *Very good for plans for the layout of new rooms, line of sight, access etc. particularly the new VCR.*

35 **I.35** *Simulation of ATC visual control room, basic ATC training, tower emergency training. Azimuth 360°, elevation say* $-10°$ *[to]* $+40°$.

35 **I.36** *ATC furniture & ergonomic simulations/trials prior to installation. Include a/c outside, colour schemes.*

36 **I.37** *I certainly felt this could be useful in designing, for example the new ATC twr so that errors in size are not made and more importantly 'lines of sight'.*

38 **I.38** *Certainly control room design/layout review would be a possibility however I feel that the technology is not yet far enough advanced to allow this application.*

39 **I.39** *In time, will be very useful for modeling the visual element of control towers—moving desks re line of sight. Adding people, again re line of sight.*

36 **I.40** *VCR training could be made more realistic by using VR technology than the present 2-D computer screens.*

41 **I.41** *Useful for planning of new VCRs, gives us the ability to assess viewpoints out of the tower, help recognise the 'cone of blindness' inherent with the new type of tall stalk mounted VCR cabs.*

41 **I.42** *Training etc. isn't really suited to the headset due to the large amount of writing on strips etc. involved in the job.*

42 **I.43** *For simulation purposes on a limited scale i.e. one controller & some a/c I could envisage a use, but to act as part of a team is important at Heathrow though I feel a VCR mockup with computer displays on the windows (as in flight simulation) could be more realistic.*

43 **I.44** *It gives an opportunity to combine actual pictures with simulated data through perhaps a lightweight spectacle/projector arrangement.*

45 **I.45** *My initial thoughts are that VR is better suited to the VCR environment than for radar applications. There is an ongoing requirement for VCR simulators that replicate the environment faithfully. The ability to simulate a visual display from a VCR would be extremely welcome especially where there is a need for a 360° view. Also there would be only a limited requirement for interaction with the environment; attention should be directed outside the VCR to virtual aircraft, vehicles and taxiways. The GMC application is the most exciting with a 360° view of manœuvering aircraft possible. Much work needs to be done but at the moment all VCR simulators are very limited. The technology is available on aircraft simulators so let's exploit it! VR certainly seems to be a cost effective option.*

47 **I.46** *... a ground/tower simulator would be a very useful training aid to provide all round vision in simulation. The use of a tower simulation using a virtual reality idea has a great deal of potential in the future.*

49 **I.47** *I think with modification and progress the system could be used to harness controllers' views on a new control tower before it is designed and built, not only for what the tower would be like inside, but more importantly, the line of sight views one would get. If new stands, terminals etc. were being planned then the system could be used to see what affect the building would have on vision from the VCR.*

49 **I.48** *There may be an opportunity for training ATCOs for the VCR. However as most of the 'real' training an ATCO would do is live 'on the job' training I feel this would be limited.*

I.4 Proposed Application

30 **I.49** *The visual control room (VCR) could benefit from VR, particularly at night or in fog, when a see-through image could help in monitoring the progress of otherwise non-visible aircraft.*

33 **I.50** *... 'live' with computer input during fog or at bnight.*

35 **I.51** *Application to low vis operations, headset as if NVG.*

38 **I.52** *'Head-up' display may be useful but again technology would need to allow much smaller, lighter equipment.*

40 **I.53** *I can forsee uses—perhaps no longer requiring controllers to be in tall towers actually at the aerodrome. Some way off though.*

40 **I.54** *VCR work can be demanding, but seldom challenges the eyes & mind in so narrow and intense way.*

41 **I.55** *The headset may have limited use in a VCR where conditions of low visibility due to fog or night conditions might enable callsigns or images to be shown to the ATCO who would otherwise see only grey mist or limited a/c lighting at night.*

43 **I.56** *Has potential for the future. Could combine with radar data (SMR data) for low visibility uses or even HUD information for controller.*

46 **I.57** *Only real advantage for VCR would be in fog: but would the positions [of computer-generated aircraft symbols] really reflect the 'actual position' of the a/c well enough to allow it to be used.*

I.5 Miscellaneous

35 **I.58** *Application to flight training in partic[ular] basic flying training (visual circuits) with data overlay—height, heading, speed/vert speed etc. Tactical weapons training, military fast jet. Probably suited to single seat/tandem environment.*

35 **I.59** *Formation flight in cloud.*

49 **I.60** *There may well be a use for a much improved version on board an aircraft for use when landing in low visibility but with today's modern avionics it might be a case of a solution looking for a problem which has already been solved.*

Glossary

2D Two-dimensional

3D Three-dimensional

ATC Air Traffic Control

ATCO Air Traffic Control Officer

CGI Computer-Generated Imagery

COP Centre of Projection

COR Centre of Regard

CRT Cathode Ray Tube

CTDI Cockpit Display of Traffic Information

DOP Direction of Projection

FDP Flight Data Processing

GFOV Geometric Field of View

GUI Graphical User Interface

HMD Head-Mounted Display

IG Image Generator

LATCC London Air Traffic Control Centre

LCOP Left Centre of Projection

LOS Line of Sight

LVPN Left View Plane Normal

mode-A Octal transponder code, set by the pilot and transmitted in response to an interrogation signal from a secondary surveillance radar.

mode-C Barometric altitude information transmitted by an altitude-encoding transponder in response to an interrogation signal from a secondary surveillance radar.

MTI Moving Target Indicator

NATS National Air Traffic Services

POR Point of Regard

PPI Plan Position Indicator

R/T Radio Telephone

RCOP Right Centre of Projection

RD Regard Distance

RDP Radar Data Processing

RVPN Right View Plane Normal

SP Station Point

SSR Secondary Surveillance Radar

TTW Through the Window

VA Visual Angle

VC Viewing Coordinates

VCR Visual Control Room. The 'glasshouse' at the top of the control tower.

VD Viewing Distance

VE Virtual Environment

VEOS Virtual Environment Operating System

VFR Visual Flight Rules

VP View Plane

VPN View Plane Normal

VR Virtual Reality

WC World Coordinates

Bibliography

[Bad82] Alan Baddeley. *Your Memory: A User's Guide.* Penguin, 1982. ISBN 0-14-013683-5.

[BB91] Meridyth Svensa Burnett and Woodrow Barfield. Perspective versus plan view air traffic control displays: Survey and empirical results. In *Proc. Human Factors Soc. 35th Annual Meeting*, pages 87–91, 1991.

[BHRR92] Mark Brown, David Hawes, Chris Rigby, and Peter Rosner. A graphics library to meet the challenge of changing hardware. In *Computing and Control Division Colloquium on Computer Graphics Systems.* Institute of Electrical Engineers, 1992. Digest No. 1992/010.

[BLR90] Woodrow Barfield, Rafael Lim, and Craig Rosenberg. Visual enhancements and geometric field of view as factors in the design of a three-dimensional perspective display. In *Proc. Human Factors Soc. 34th Annual Meeting*, pages 1470–1473, 1990.

[Bur91] Meridyth Svensa Burnett. An analysis of the effects of multiple colors, information density, and traffic complexity on plan view and perspective air traffic control radar displays. Master's thesis, Dept. of Industrial Engineering, University of Washington, Dept. of Technical Communication, College of Engineering Human Factors Lab., Dept. of Industrial Engineering, University of Washington, Seattle WA 98195, 1991.

[Civ94] Civil Aviation Authority, London. *CAP 413 Radiotelephony Manual,* April 1994. ISBN 0 86039 199 X.

[Duk92] Graham Duke. *Air Traffic Control.* Ian Allan, 4th edition, 1992. ISBN 1-7110-2008-6.

[EH90] Stephen R. Ellis and Selim S. Hacisalihzade. Symbolic enhancement of perspective displays. In *Proc. Human Factors Soc. 34th Annual Meeting*, pages 1465–1469, 1990.

[Ell89] Stephen R. Ellis. Vision of visualization aids: Design philosophy and observations. *SPIE*, 1083:220–227, 1989.

[EMH84] Stephen R. Ellis, Michael W. McGreevy, and Robert J. Hitchcock. Influence of a perspective cockpit traffic display format on pilot avoidance maneuvers. In *AGARD Conference Proceedings*, pages 16–1–16–9, 1984.

[ETGS91] Stephen R. Ellis, Gregory K. Tharp, Arthur J. Grunwald, and Stephen Smith. Exocentric judgements in real environments and stereoscopic displays. In *Proc. Human Factors Soc. 35th Annual Meeting*, pages 1442–1473, 1991.

[FVFH90] James D. Foley, Andries Van Dam, Steven K. Feiner, and John F. Hughes. *Computer Graphics: principles and practice.* Addison Wesley, 2nd edition, 1990. ISBN 0-201-12110-7.

[Gra89] David Graves. *United Kingdom Air Traffic Control.* Airlife Publishing Ltd., 1989. ISBN 1-85310-107-9.

[Gre90] R. L. Gregory. *Eye and Brain: The Psychology of Seeing.* Oxford University Press, 4th edition, 1990.

[HH93] R. H. R. Harper and John A. Hughes. What a f-ing system! send 'em all to the same place and then expect us to stop them hitting: Making technology work in air traffic control. In Graham Button, editor, *Technology in Working Order: Studies of work, interaction, and technology*, chapter 7, pages 127–144. Routlege, 1993.

[HHS91] R. R. Harper, J. A. Hughes, and D. Z. Shapiro. Harmonious working and CSCW: Computer technology and air traffic control. In J. M. Bowers and S. D. Benford, editors, *Studies in Computer Supported Cooperative Working*, pages 225–234. Elsevier Science, 1991.

[Hop92] V. David Hopkin. Personal communication, 1992. RAF Institite of Aviation Medicine.

[Hop94] V. David Hopkin. Personal communication, February 1994. RAF Institite of Aviation Medicine.

[How92] David C. Howell. *Statistical Methods for Psychology.* Duxbury, third edition, 1992. ISBN 0-534-92955-9.

[HRG84] David G. Haralson, John M. Reising, and Joseph Ghrayeb. Toward the panoramic cockpit, and 3-D cockpit displays. In *IEEE National Aerospace and Electronics Conference*, pages 80–85, 1984.

[HRS92] John A. Hughes, David Randall, and Dan Shapiro. Faltering from ethnography to design. In *CSCW '92 Proceedings*, pages 115–122. ACM, November 1992.

[Lan91] Robert C. Lansdale. Texture mapping and resampling for computer graphics. Master's thesis, Department of Electrical Engineering, University of Toronto, Toronto, Canada, January 1991.

[Lea94] David Learmount. Reading the way: what are the human-factors issues in air-traffic control by datalink? *Flight International*, 146(4447):26–27, 16-22 November 1994.

[LS84] D. V. Lindley and W. F. Scott. *New Cambridge Elementary Statistical Tables.* Cambridge University Press, 1984. ISBN 0-521-26922-9.

[Mal91] A. J. Maltezos. Computerisation of flight progress strips used by air-traffic controllers. Master's thesis, Dept. of Computer Science, QMW College, University of London, Dept. of Computer Science, Queen Mary & Westfield College, Mile End Road, London E1 4NS, 1991.

[ME86] Michael W. McGreevy and Stephen R. Ellis. The effect of perspective geometry on judged direction in spatial information instruments. *Human Factors*, 28:439–456, 1986.

[Mil84] Steve Miller. *Experimental Design and Statistics.* New Essential Psychology. Routledge, 2nd edition, 1984. ISBN 0–415–04011–6.

[MJ89] Ann Moir and David Jessel. *Brain Sex: The Real Differences between Men and Women.* Michael Joseph, 1989.

[MR90] Kim M. Mazur and John M. Reising. The relative effectiveness of three visual depth cues in a dynamic air situation display. In *Proc. Human Factors Soc. 34th Annual Meeting*, pages 16–20, 1990.

[MVF72] Humberto R. Maturana, Francisco G. Varela, and Samy G. Frenk. Size constancy and the problem of perceptual spaces. *Cognition — International Journal of Cognitive Psychology*, 1(1):94–107, 1972.

[NH85] C. S. Narborough-Hall. Recommendations for applying colour coding to air traffic control displays. *Displays*, pages 131–137, June 1985.

[Nol90] Michael S. Nolan. *Fundamentals of Air Traffic Control*. Wadsworth Publishing Company, 1990. ISBN 0-534-12246-9.

[PG93] Peter W. Pruyn and Donald P. Greenberg. Exploring 3D computer graphics in cockpit avionics. *IEEE Computer Graphics & Applications*, pages 28–35, May 1993.

[RM92] L. Reynolds and C. Metcalfe. The interim NATS standard for the use of colour on air traffic control displays. CS Report 9213, Issue 1.2, UK CAA, Chief Scientist's Division, National Air Traffic Services, 1992.

[RS86] J. M. Rolfe and K. J. Staples, editors. *Flight Simulation*. Cambridge Aerospace Series. Cambridge University Press, 1986. ISBN 0 512 35751 9 (paperback).

[SD90] Bruce A. Steiner and Diane A. Dotson. The use of 3-d stereo display of tactical information. In *Proc. Human Factors Soc. 34th Annual Meeting*, pages 36–40, 1990.

[SD91] Mel Slater and Allan Davison. Liberation from flatland: 3d interaction based on the desktop bat. In F. H. Post and W. Barth, editors, *Proc. Eurographics '91*. Elsevier Science Publishers B.V. (North Holland), September 1991.

[SHRH94] D. Shapiro, J. A. Hughes, D. Randall, and R. Harper. Visual re-representation of database information: the flight data strip in air traffic control. In Michael J. Tauber, Dirk E. Mahling, and Farah Arefi, editors, *Cognitive Aspects of Visual Languages and Visual Interfaces*, chapter 4, pages 349–376. Elsevier Science, 1994.

[Str91a] Bjarne Stroustrup. *The C++ Programming Language.* Addison Wesley, 2nd edition, 1991. ISBN 0-201-51459-1.

[Str91b] Geoffrey G. Strutt. The human computer interface in air traffic control: Feasibilty study for a three dimensional radar picture. Master's thesis, Middlesex Polytechnic, London, UK, 1991.

[Sut66] Ivan E. Sutherland. The ultimate display. In *Proceedings of the IFIP Congress*, volume 2, pages 506–508, 1966.

[SVBW88] J. L. Leeds S. V. Bemis and E. A. Wiener. Operator performance as a function of type of display: Conventional versus perspective. *Human Factors*, 30(2):163–169, 1988.

[SW88] Pauline Smith and Chris Whetton. *General Ability Tests.* ASE, A Division of NFER-NELSON, 1988.

[SWK90] Alan Stokes, Christopher Wickens, and Kirsten Kite, editors. *Display Technology—Human Factors Concepts.* Society of Automotive Engineers, 1990. ISBN 0-89883-530-5.

[Tam93] Benny Kai-cho Tam. Three dimensional display for air traffic control. Master's thesis, Department of Computer Science, University of Manchester, March 1993.

[Tol87] Philip Tolin. Maintaining the three-dimensional illusion. *Information Display*, pages 10–12, December 1987.

[Tos91] Toshiba Corporation. *TC8512AG/AY High Speed Shading Processor (HSP) datasheet (tentative)*, 1991.

[Way88] Thomas C. Way. Steropsis in cockpit display—a part-task test. In *Proc. Human Factors Soc. 32nd Annual Meeting*, pages 58–62, 1988.

[Way89] Thomas C. Way. 3-d in pictorial formats for aircraft cockpits. In *Proc. Human Factors Soc. 33rd Annual Meeting*, pages 23–27, 1989.

[WF88] Christopher D. Wickens and John M. Flach. Information processing. In Earl L. Wiener and David C. Nagel, editors, *Human Factors in Aviation*, chapter 5, pages 111–155. Academic Press, Inc., 1988.

[WJ82] D. Whitfield and A. Jackson. The air traffic controller's "picture" as an example of a mental model. In *Proceedings of the IFAC/IFIP/IFORS/IEA Conference on Analysis, Design and Evaluation of Man-Machine Systems.* Baden-Baden, 1982.

[WT90] Christopher D. Wickens and Steven Todd. Three-dimensional display technology for aerospace and visualisation. In *Proc. Human Factors Soc. 34th Annual Meeting*, pages 1479–1483, 1990.

[WTS89] Christopher D. Wickens, Steven Todd, and Karen Seidler. Three-dimensional displays: Perception, implementation and applications. Technical report, Aviation Research Laboratory, Institute of Aviation, University of Illinois at Urbana-Champaign, 1989. Subcontract for AAMRL/HEL/CSERIAC Wright-Patterson Air Force Base, Contract DLA900-88-D-0393.

[YA85] John P. Yorchak and James E. Allison. Utilizing computer graphics display technology. In *IEEE NAECON*, pages 937–942, 1985.

[ZRWB88] John P. Zenyuh, John M. Reising, Scott Walchli, and David Biers. A comparison of a stereographic 3-d display versus a 2-d display using an advanced air-to-air format. In *Proc. Human Factors Soc. 32nd Annual Meeting*, pages 53–57, 1988.