3D Displays: A Human-Centred Review

The investigation which is the subject of this Report was initiated by the Dstl Programme Office.

Reference .... HFIDTCPV_T2_2_2

Version 1

Date ....24 January 2012

Prepared by BAE Systems (Operations) Limited
under Contract No. C/HCB/N03751 (DSTLX 1000041669) dated 04/2010.
The rights of the Ministry of Defence of reproduction and use are subject to the terms of DEFCON90 (Edn 11/06)

Security

Any person, other than the authorised holder, upon obtaining possession of the document should forward it, together with his name and address in a sealed envelope to:

The Security Officer
BAE Systems
Alvington
Yeovil
Somerset
BA22 8UZ

Telephone .... (01935) 443000

Facsimile... (01935) 443111

ALL RECIPIENTS OF THIS REPORT ARE ADVISED THAT IT MUST NOT BE COPIED IN WHOLE OR IN PART OR BE GIVEN FURTHER DISTRIBUTION OUTSIDE THE AUTHORITY WITHOUT THE WRITTEN APPROVAL OF THE DSTL PROGRAMME OFFICE – HUMAN DIMENSION AND MEDICAL SCIENCES DOMAIN.

Approvals

Approval .... Dr Karen Lane
Title .... HFI DTC Director

Authorisation .... Professor Bob Stone
Title .... Work Package Lead

Distribution

Dr Colin Corbridge	Capability Advisor, Dstl
Fiona Cotter	Dstl Programme Office
Neal Smith	Dstl Programme Office
Julian Starkey	HQ LF, FDT
Paul Elrick	HQ LF, FDT
Dr Dawn Johansen	ARTD
BAE Systems, Yeovil	File
Consortium Members	HFI DTC

Amendment Page

Version	Amendment Date	Change Reference	Author/ Amended By	Remarks
1	24 Jan 12	n/a	Professor Bob Stone	Original Issue
1b	08 Feb 12	01	Professor Bob Stone	Corrections Following First Proof-Read

Authors

Professor Robert J. Stone **University of Birmingham**

Dr James F. Knight **University of Birmingham**

Contents

Acronyms & Abbreviations

2D	Two Dimensions / Dimensional
3D	Three Dimensions / Dimensional
AEA	Atomic Energy Authority
AR	Augmented Reality
ATC/M	Air Traffic Control/Management
C^2	Command and Control
CAVE	Cave Automatic Virtual Environment
CAD	Computer-Aided Design
CCTV	Closed-Circuit Television
CD	Compact Disc
CGI	Computer-Generated Imagery
C-IED	Counter Improvised Explosive Device(s)
COTS	Commercial Off-The-Shelf
COVE	COmputerised Virtual Environment (in effect, a small CAVE)
CRT	Cathode Ray Tube
CSERIAC	Crew Systems Ergonomics Research Information Analysis Center
CURV	Cable Controlled Underwater Recovery Vehicle
Dstl	Defence Science & Technology Laboratory
DVD	Digital Video Disc
EOD	Explosive Ordnance Device/Disposal
FD^2	Frisby-Davis Distance (Stereo Vision Test)
FPR	Film Pattern Retarding
HD	High Definition

HFI DTC	Human Factors Integration Defence Technology Centre
HMD	Head-Mounted Display
HMMWV	High-Mobility Multi-Wheeled Vehicle
HQ LF	Headquarters Land Forces
i3D	Interactive 3D
IED	Improvised Explosive Device
IPD	Inter-Pupillary Distance
LCD	Liquid Crystal Display
LED	Light-Emitting Diode
LEEP	Large Expanse Extra Perspective
MIS	Minimally Invasive Surgery
MIT	Massachusetts Institute of Technology
MV	Monoscopic Vision
N/A	Not Applicable
NASA	National Aeronautics & Space Administration
PC	Personal Computer
PhD	Doctor of Philosophy
PVD	Plan View Display
RB2	Reality Built for Two
ROV	Remotely Operated Vehicle
SCAT	Submersible Cable-Actuated Teleoperator
SV	Stereoscopic Vision
TNO	Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek
TRL	Technology Readiness Level
TV	Television
UAV	Unmanned ("Uninhabited") Air Vehicle

UK	United Kingdom
US(A)	United States (of America)
VCASS	Visually-Coupled Airborne Systems Simulator
VE	Virtual Environment
VR	Virtual Reality

1 Executive Summary

History has shown how human interface technologies, be they early prototypes or well-established market products, quite regularly find their way into mainstream, real-world applications with very little (if any) attention being paid to the capabilities, limitations and learning requirements of their end users. Despite the unique developmental history surrounding the development of 3D and stereoscopic displays and associated viewing devices, such technology also now falls into this "premature adoption risk" category, irrespective of very recent developments in the commercial exploitation of stereoscopic cinematic, home entertainment and gaming products. Of equal concern is the fact that stereoscopic and 3D viewing technologies have been under investigation by non-entertainment-based organisations for nearly five decades, yet their widespread adoption has still to be witnessed.

This report sets out by reviewing historical, current and near-term future developments in 3D and stereoscopic viewing, especially as they relate to the interests and applications opportunities offered by the defence community. The aim of providing a brief history of stereoscopic viewing developments is to show how 3D technologies, like many other similar "high-tech" developments (and despite frequent periods of waxing and waning), have found niche applications within the defence sector. Yet, only a handful of devices and techniques have actually made a real difference to real-world visualisation and training applications. A review of the current main technologies involved in the generation and display of stereoscopic and 3D imagery is also included, presented from a Human Factors perspective and emphasising the pros and cons of different viewing techniques, including active and passive stereo, autostereo ("glasses-less") systems, volumetric devices and head-mounted displays. Throughout the report, the terms "3D" and "stereoscopic" are used interchangeably, although the majority of the research has focused on the display of stereoscopic or binocular visual information. Nevertheless, where relevant (and especially in the cases of Air Traffic Management and Command & Control), research relating to the use of 3D features in the generation of graphical displays is considered.

It was not the intention of this report to present a detailed overview of human 3D or binocular vision, as these topics are more than adequately covered in standard academic texts and within numerous online sources. However, it is important to be aware of some of the limitations and dysfunctions inherent in the stereoscopic characteristics of some individuals when considering the use of 3D display adoption. These issues have been summarised in a short section of the report, supplemented with a similarly short review of the tests that are available to support the screening of individuals with mild to severe binocular dysfunctions. The earlier review of the different stereoscopic technologies available highlights a number of shortcomings with the different display techniques. The effects of these shortcomings on human visual comfort and general well-being could be amplified when the displays are exposed to observers with certain stereoscopic dysfunctions.

Following a summary of the generic benefits of stereoscopic or binocular viewing, the main section of the report deals with applications case studies. Evidence for and against the use of various technologies is presented, covering the areas of Air Traffic Management, Aircrew Operations, Teleoperation, Surgery and Command & Control. It becomes very obvious from this section that, whilst there are a handful of relatively clear-cut examples of successful implementations in the stereoscopic and 3D viewing literature, in the main there is no conclusive evidence nor any conclusive guidelines relating to the adoption of the technology in any of the domains reviewed. It is also evident that there is a “disconnect” between laboratory studies and real-world domains and experiences.

The study concludes that, regardless of the application domain, ANY opportunity to deploy 3D or stereoscopic display systems MUST be preceded by a Human Factors study of the tasks required of the end users and the context in which such technology might be deployed. Some suggested points of interest and concern are presented for discussion. Another conclusion is that work should be undertaken to develop a Human Factors methodology that supports the analysis of tasks best suited to implementation in 3D synthetic or stereoscopic form, the design of appropriate 3D content, the selection of appropriate display hardware, the analysis of the environments into which a 3D display is to be introduced and the development of an appropriate set of subjective and objective metrics for the evaluation of different systems.

2 Introduction

2.1 Background

The contents of this report present a human-centred review of historical, current and near-term future developments in 3D and stereoscopic viewing, especially as they relate to the interests and applications opportunities offered by the defence community. The aim of providing a brief history of stereoscopic viewing developments is to show how 3D technologies, like many other similar "high-tech" developments have found niche applications within the defence sector. Yet, only a handful of devices and techniques have actually made a real difference to real-world visualisation and training applications. A review of the current main technologies involved in the generation and display of stereoscopic and 3D imagery is also included, presented from a Human Factors perspective and emphasising the pros and cons of different viewing techniques, including active and passive stereo, autostereoscopic ("glasses-less") systems and volumetric displays. Throughout the report, the terms "3D" and "stereoscopic" are used interchangeably, although the majority of the research has focused on the display of stereoscopic or binocular visual information, with the occasional reference to synthetic or virtual 3D data display using binocular or monocular presentational methods.

2.2 A Brief Historical Review

Binocular, or stereoscopic, vision was first described by Charles Wheatstone in 1838, for which he was awarded the Royal Society's Royal Medal two years later, after demonstrating his research into the perception of solid (or *stereo*) objects through geometric drawings and the first reflecting mirror stereoscope. Nearly a decade later, the Scottish scientist David Brewster developed what was to become the "standard" design for box stereoscopes and, by 1856, he claimed that he had sold over half a million units – hence the popularity during the Victorian era for stereoscopic photography and viewing. The general public's passion for stereoscopic imaging declined in the early 1900s, as a result of the growth of interest in early silent films. Nevertheless, stereoscopic imagery had already been identified as a means of obtaining information of strategic military interest when, during the Franco-Prussian War in the early 1870s, a military unit equipped with stereo cameras was deployed to map the fortifications at Strasbourg. Long before the use of stereo aerial photographs and associated interpretation processes in World War 2, techniques for the capture of aerial stereo imagery from two balloons was actually patented by American C.B. Adams in 1893, building on much earlier (1858) tests of terrestrial photography from cameras deployed on a string of kites (the first recognised example of aerial photogrammetry). Despite early interest in Adams' developments by the US Army Signal Corps, the technique was not actually used until 1908.

In the late 1930s, the UK's passion for 3D imagery as a form of entertainment was rekindled, partly as a result of Logie Baird's early experiments with stereoscopic television. These experiments continued well into the 1940s, although the first commercial 3D TV broadcast did not take place until 1980 in the USA. Even more of a stimulus to the resurrection of interest in 3D was the result of a partnership between Americans William Gruber and Harold Graves (the latter employed by the US company Sawyer's Photo Services). Exploiting Kodak's new and flexible 35mm film, Gruber and Graves developed a Bakelite stereoscopic viewer that accepted cardboard disks containing seven stereo image pairs, and the first *View-Master* was born. The *View-Master* (Figure 1, top image) was launched by Sawyers as a product at the New York World Fair in 1939 and is still available today.

As with much earlier stereoscopic imaging techniques, the US military adopted the *View-Master* for a wide variety of personnel training régimes, including anti-aircraft range estimation, Air Force "cones of fire" estimation and ship-to-ship identification. Some 100,000 viewing units and nearly six million stereo image disks were procured for military use from 1942 to the end of World War 2. A competitor to Sawyer's, Tru-Vue Inc., was founded in 1931 and, during World War 2, produced a stereo film strip for the company's viewer (Figure 1, bottom image) entitled "Keep 'em Flying". This featured black & white stereo images of aircraft models to support recognition training for civilians. Tru-Vue's assets – in particular its lucrative Disney Studio contracts – were acquired by Sawyers in 1952, two years after the company had introduced colour into its film strips to compete with the *View-Master* product.

Figure 1: *View-Master* (Top Image) and *Tru-Vue Viewer* (Bottom Image)
Sources: http://www.007collector.com and www.etsy.com

With the strong uptake of television in the US in the 1950s, the cinema's answer to rekindling the public's interest in film presentations was to attempt to provide rudimentary 3D movies – the titles and reviews of the earliest offerings can be found within numerous Internet sites. However, none of these ever reached the levels of success demonstrated by the more impressive wide-field movie features offered by *Cinemascope*, created by 20th Century-Fox in 1953 and *Cinerama*, first demonstrated to the public in 1952. Indeed, *Cinerama* became one of the inspirations for the early work of the late Morton Heilig. In 1955, Heilig's essay "The Cinema of the Future" outlined his vision of a theatre capable of delivering multi-sensory experiences (3D imagery, sound, motion and smell) to large theatre audiences. Yet, Heilig's greatest achievement was a solution that actually moved away from the large audience experience, and attempted to deliver the "Cinema of the Future" to the individual user. Heilig's solution was *Sensorama* (Figure 2) – a multi-sensory arcade machine with wide-angle stereoscopic film images which – much later – was to confirm the inventor as the "Father of Virtual Reality". Interestingly, in his patent for the *Sensorama* (invented in 1957 and awarded in 1962), Heilig made specific mention of the potential of the system to support US military instructors, who, to use his words "must instruct men in the operation and maintenance of extremely complicated and potentially dangerous equipment, and it is desirable to educate the men with the least possible danger to their lives and to possible damage to costly equipment". Adding to his credits as the "Father of Virtual Reality", Heilig also invented the *Telesphere* mask, which was awarded a patent in 1960, entitled "Stereoscopic Television Apparatus for Individual Use". The *Telesphere* concept was, in effect, a head-mounted display (HMD) implementation of *Sensorama* and, as such, pre-dated the efforts of other pioneers normally accredited with launching the Virtual Reality (VR) community. These included Philco Inc., with their *Headsight* head-mounted display, developed in 1961 for the purposes of remote viewing of hazardous environments, Ivan Sutherland's *Sword of Damocles* HMD (Sutherland, 1965), and the HMDs developed by NASA and the US company VPL Inc. in the late 1980s, including the commercial *EyePhone* product.

Figure 2: Promotional Image of *Sensorama* Source: Authors' Archives

It was only a matter of time before researchers in safety-critical industries began to appreciate the potential of stereoscopic viewing for operations that demanded the human operator to be located at a remote and safe workstation. The history of modern *teleoperation* began at the end of the 1940s when the first master-slave manipulator was developed in the Argonne National Laboratory in Illinois for chemical and nuclear material handling (Vertut & Coiffet, 1985). Following the pioneering use by the Royal Navy of the *Cutlet* remotely operated vehicle (ROV) for torpedo recovery in the 1950s, the domain of subsea teleoperation witnessed a slow pace of technological developments, lagging behind its nuclear counterparts. It was not until the 1960s and 70s that the complexity of ROV technologies accelerated, stimulated in part by the highly successful US hydrogen bomb and manned submersible recovery missions with the *CURV* (Cable Controlled Underwater Recovery Vehicle) platforms. On the UK side of the Atlantic, high-tech "workhorse" submersibles also came under the spotlight. The British Aircraft Corporation's ROV *CONSUB 2* was chosen as the flagship submersible for the newly formed Sub Sea Services offshore engineering company, following its appearance in a 1976 Birmingham ROV

Exhibition. Indeed, the *CONSUB* ROVs were equipped with one of the earliest underwater stereoscopic camera systems, as well as other advanced sensors, and these capabilities prompted the UK Department of Energy to launch its *Bondi Initiative* in the late 1970s, the aim of which was to undertake research into technologies capable of replacing the human from hazardous underwater environments. In the US, another advanced ROV, *SCAT* (Submersible Cable-Actuated Teleoperator), was also built in the early 1970s, specifically to investigate subsea stereoscopic television controlled by a head-coupled human interface.

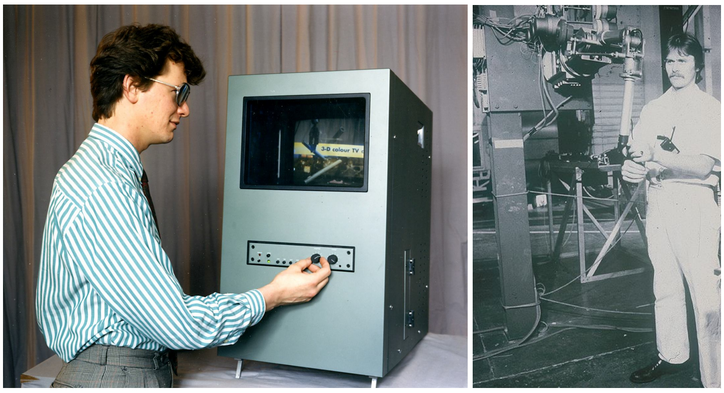


Figure 3: The AEA TV^3 Display (Left) and MA-23M Manipulator (Right) Source: Authors' Archives

In the UK, one of the most successful stereoscopic remote viewing developments in the 1980s was undertaken by the Atomic Energy Authority (AEA) at Harwell (e.g. Dumbreck *et al.*, 1991). The AEA TV^3 Display (Figure 3, left image) consisted of two monitors and an optical arrangement which combined the two remote CCTV pictures using a semi-reflective mirror, viewable by means of lightweight polarising spectacles. The original TV^3 monitor sizes were 15 inches for black and white and 16 inches for colour (diagonal screen dimensions). To provide a more compact display unit, 12-inch black and white and 14-inch colour monitors were made available later in that decade. One slight problem in using twin monitors and semi-reflective mirrors is that care had to be taken to avoid veiling glare on the optical surfaces. Not only could veiling glare be visually disturbing, it would also destroy the stereoscopic effect. Nevertheless, experience of using the display with the La Calhène MA-23M Master-Slave Manipulator at Harwell (Figure 3, right image) suggested that, when compared with conventional 2D TV systems, the TV^3 device improved users' performance (e.g. completion time and accuracy) on basic remote handling tasks (e.g. "peg-in-hole") by 27% (Stone & Mason, 1986).

Nuclear industry developments notwithstanding (with subsea, space and defence applications very much playing a "catch-up" role, in an historical sense), the main driver behind research and commercial interests in stereoscopic displays came, without doubt, from early developments within the early Virtual Reality (VR) community. Many of the formative developments were being spearheaded in the US, notably through the *SuperCockpit* programme at the US Wright Patterson Air Force Base in Dayton, Ohio, and as a result of pioneering Human Factors research into intuitive control techniques for space robots ("telepresence") at NASA Ames. The goal of the *SuperCockpit* work was to develop advanced avionics and cockpit management systems to enhance information transfer and to protect pilots' eyes from the dazzle threat posed by laser weapons and nuclear airbursts. Instead, virtual representations of the environment external to the cockpit were generated from airborne sensors and reconnaissance data for presentation using large screens or sophisticated stereoscopic HMDs, such as the *VCASS* System (Visually-Coupled Airborne Systems Simulator; Harvey, 1987).

So-called "immersive" technologies – designed to create a believable illusion of "presence" within a computer-generated virtual environment, or at a remote real worksite – were first commercially demonstrated in June 1989 when the VPL Inc. launched their *RB2* system ("Reality Built for 2"), many components of which were exploited as a result of earlier NASA research. *RB2* featured the *EyePhone* HMD, mentioned above, which comprised a pair of liquid crystal displays (effectively cannibalised pocket televisions) mounted behind a special stereo-optical assembly of lenses, all mounted within a cumbersome diver mask-like headset. The lens system, a proprietary wide-angle viewing product called *LEEP* (Large Expanse Extra Perspective), was developed by a Massachusetts-based company and was, for many years, the optical system of choice for nearly all HMD products.

In the UK, the other country accredited with major innovations in the field, VR first came to the notice of the British public late in the late 1980s. Commercial research teams who had been involved in developing the technology presented their work for the first time at the 1990 London Computer Graphics Conference. Even before then, VR projects had been under way, the first notable instance being in the early 1980s. The *Spatial Workstation*, literally a trolley-mounted television, displayed simple 3D wire frame images to the wearer of shuttered glasses ("active stereo" – see Section 3.2). The demonstrator was developed as part of a PhD research project at Loughborough University by Jonathan Waldern, who went on to establish the VR games company W Industries (later Virtuality plc, the developers of the *Visette* stereoscopic HMD, shown in Figure 4).


Figure 4: The Virtuality *Visette* HMD
Source: Authors' Archives

Another attempt at achieving a credible sense of immersion, avoiding the need to don cumbersome HMDs was the CAVE (Cave Automatic Virtual Environment), developed in 1992 by the Electronic Visualization Laboratory at University of Illinois at Chicago. The CAVE is, in effect, a small room within which a small number of users are surrounded by whole-wall displays onto which the virtual images are back-projected using high quality video projectors. CAVE users are often seen wearing liquid crystal "shutter" glasses, synchronised with the projectors, so that each alternate scan line of the display triggers one of the shutters, presenting left-eye or right-eye images only, thus creating a 3D effect (again, an example of "active stereo" – see Section 3.2). Since this time, there have been many variations on the theme of the CAVE, including desktop versions (or "COVEs") and even spheres, providing a form of omni-directional treadmill, allowing users to move with some degree of freedom through a virtual scene.

Technologies capable of presenting VR users with stereoscopic visual information have undergone a wide range of transformations since these early attempts. However, it is fair to say that, despite the emergence of new product ranges in recent years, VR technologies have not changed significantly in their capabilities since the early 1980s and the "Holy Grail" of immersion remains as elusive today as it did then (see also Stone 2011).

As well as the highly publicised developments in stereoscopic technologies for cinemas and domestic TV, also being targeted for the "3D treatment" are personal and mobile computers and digital imaging and communication technologies. This trend is, no doubt, being spearheaded by the manufacturers to encourage and sustain widespread adoption by the future mass consumer electronics market. Low-cost stereoscopic lenses and viewers are now available for digital cameras (Figure 5). Laptops equipped with 3D-ready displays are also becoming available that are capable of exploiting real-time left-eye/right-eye image generation from the increasing number of advanced stereo-ready graphics cards. However, many of these incur significant weight penalties, due to the display technologies used, some reaching up to twice the weight of similar high-performance gaming laptops.


Figure 5: Loreo 3D Lens for Digital Cameras.
Source: www.denverslair.co.uk

Figure 6: Nintendo *3DS* Console
Source: www.bestnintendo3dsgames.com

Handheld gaming devices, such as Nintendo's *3DS* (Figure 6) now feature versions of the parallax barrier technology used in domestic TVs (see Section 3.3). Even the iPhone, iPod and iPad products are unable to escape the attention of developers who believe that 3D information display is the way ahead. One form of glasses-less 3D for the Apple product range uses a prototype (non-stereoscopic) technique called *head-coupled perspective*, which exploits face-tracking data generated via the integral camera to render perspective-rich on-screen images (Figure 7).

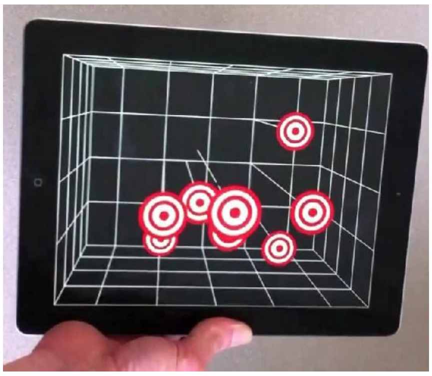


Figure 7: Head-Coupled 3D Concept Demonstration for iPad and iPod
Source: www.ipadjailbreak.com

Another technique, demonstrated by a Japanese company, makes use of a lenticular lens sheet called *Pic3D* which, it is claimed, delivers a 120° field-of-view 3D effect at an image brightness penalty of only 10%. Perhaps the clearest indication to date of technology developed based on historically proven stereoscopic techniques is Hasbro's *my3D* viewer for the iPhone/iPod (Figure 8) – in effect a 21st Century version of the popular *View-Master*, described earlier (Figure 1).


Figure 8: *my3D* Viewer for the iPhone/iPod
Source: www.hasbro.com

But are these devices really the answer to the dreams of gamers or casual users, or indispensable features demanded by future TV viewers? Judging by the many online forums and sites supporting reviews, and public feedback relating to electronic products, there appear to be as many comments against 3D technology as there are in favour, with a significant (and growing) number relating to the incidence of eyestrain and headaches. And what of adoption of 3D and stereoscopic technologies for applications other than entertainment, such as telerobotics, command and control or air traffic management? As will be seen in the remainder of this report, practical interest in the exploitation of stereoscopic viewing for real applications is certainly evident, but a "disconnect" between the scientific community and the real world exists as much today as it did in the early years of the VR community. Potential real-world adopters often fall foul of the blind-faith expectation that, if it works for the masses, then it "must" work for specialist, or non-entertainment applications. Academic and scientific papers addressing the psychophysiological, mathematical and theoretical bases of stereoscopic vision and 3D displays are in huge abundance and have been growing steadily since the early 1960s. More recently, international research appears to have focused on "automated" 3D or stereo vision systems supporting the development of advanced robots. Nevertheless, research into new and effective means of delivering usable 3D information to the human user (and this includes real-time *interaction* with such information) continues unabated. Given the opportunities offered by the ever-increasing range of technologies appearing on the entertainment market, this trend looks set to continue well into the foreseeable future.

3 3D & Stereoscopic Display Technologies

As with reviews of human 3D vision, articles relating to display technologies are in abundance in the literature and especially from online sources, given the regular developments announced by the TV industry. This section of the report provides just basic summaries of the main technologies in existence, together with some of the human-centred pros and cons relating to their potential use in future applications.

3.1 Passive Stereo

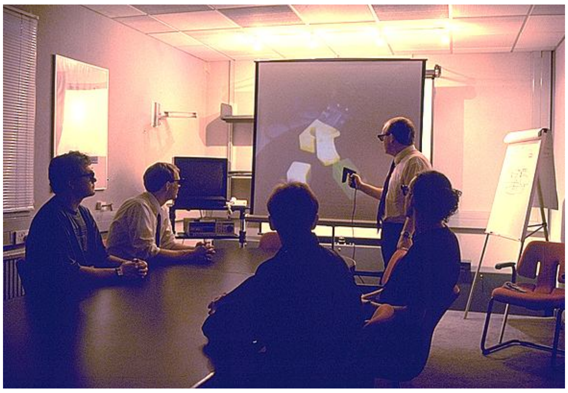


Figure 9: Multi-Observer Projection 3D System Using Polarising Glasses
Image Source: Authors' Archives

The term "passive" stereo refers to any form of technology which does *not* require some form of electronically switchable element to generate a left-eye, right-eye image separation. In the main, this category relates to the use of polarising filters, mounted on the image projection system (display screen, projector, etc.) and on glasses worn by the observer(s) (Figure 9). Polarising filters achieve a stereo effect by blocking light that is similarly polarised at the image source. Filters can be linear or circular (circular filters use clockwise and anticlockwise spiral filtration). Linear filters tend to be cheaper than circular, although circular filters allow for a degree of head-tilting on the part of the observer, without degradation of the stereo image. Passive stereo viewing technologies for TV, desktop computer and laptop displays have been evolving quite rapidly in the past few years, despite not being the 3D method of choice for the TV market until very recently (with active stereo systems dominating early products).

One of these passive stereo solutions uses two precisely-located LCD stacks, whereby the front stack controls the polarisation angle and left-/right-eye image exclusion and the back stack generates the full-colour left-/right-eye composite video signal. Another solution uses a matrix of micro-polarising optics, bonded to a flat-panel display. In some recent TVs, this technique of "film pattern retarding" (FPR) has been used to good effect, as the film is organised into alternating horizontal strips of circular polarisers, each the size of an individual pixel, thereby creating an interlaced polarising effect. Manufacturers claim this technique delivers a brighter screen with less crosstalk, less ghosting and no flickering. *Anaglyph* passive stereo glasses, which use chromatically opposite coloured filters – red-cyan or red-blue/green – are used to view images with offset colour layers. This technique is no longer widely used (due mainly to the colour distortion effects of the viewed images and incidences of binocular rivalry – the "switching" of visual consciousness between two dissimilar images presented simultaneously to each eye), but can still be found demonstrating 3D effects online, in promotional or educational media (CDs, DVDs, etc.) and in magazines.

Pros:

- In contrast to active stereo eyewear, polarising glasses are less expensive (and do not require power).
- Polarisation stereo is effective for large groups of observers.
- Polarisation stereo is less susceptible to colour or picture distortion.
- Generally, polarisation stereo places less visual strain on the wearer, which means that the 3D imagery can be viewed for longer periods of time (again when compared to active eyewear).

Cons:

- Silver (aluminium-coated, "non-depolarising") screens are required for front- or back-projection stereo to preserve polarisation - these screens can be costly and fragile. If damaged or set up without significant care and attention, these screens can produce distracting visual artefacts.
- Polarisation techniques can reduce the brightness and contrast of the final fused 3D image.
- Polarisation techniques can also reduce the vertical resolution in order to show the images for each eye. A passive stereo display system with a normal (2D) screen resolution of 1920 x 1080 will, as a result, only have a resolution of 1920 x 540 when in 3D mode. New large (84-inch) passive stereo TV displays are now becoming available such that standard HD vertical resolution will be achievable when in 3D mode.
- Polarisation techniques typically require lower ambient light environments to deliver optimum stereoscopic imagery to the observer.
- Linear polarisation glasses cannot be used with circular polarisation systems and vice versa.
- Polarisation stereo is prone to ghosting (incomplete isolation of the left and right image channels from screen and glasses-mounted filters).

3.2 Active Stereo

The term "active" stereo refers the electronic generation of binocular images for left- and right-eye viewing by means of rapid alternate eye occlusion, for example through the use of LCD shutter glasses (Figure 10). Also referred to as "field" or "frame sequential" stereo, the shutter glasses consist of a pair of LCD eyepieces which become opaque when an electric current is applied. The eyepiece switching is synchronised with a timing signal from the display source (via cable, infrared, radio signal or other means of transmission) and instructs the left or right eyepiece to turn transparent or opaque.

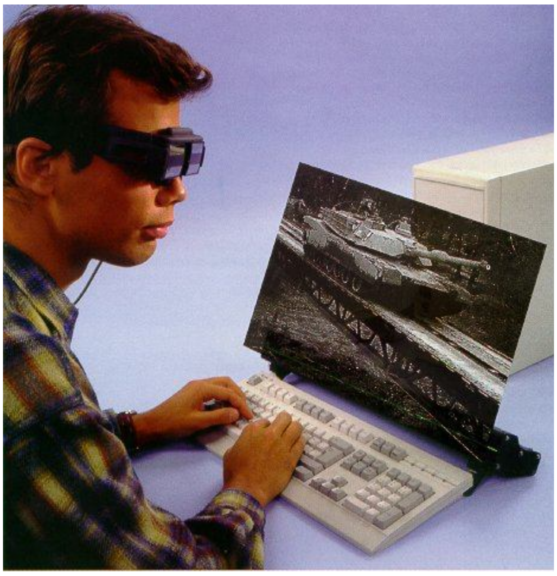


Figure 10: Active Stereo Display Example Using Shutter Glasses
Source: Authors' Archives

Historically this synchronisation was linked to the raster scanning pattern of a cathode ray tube (CRT), such that alternate scan lines synchronised with the left- and right-eye shutter. Today's LCD technology is not usually rated by frames per second but rather the time it takes for the LCDs themselves to make the transition from darkness to brightness and back to darkness. The refresh rate delivered by the stereo monitor should be 120Hz (now recognised as a "minimum" to avoid perceptible flicker), as the effect of wearing shuttered glasses is to effectively reduce this rate by 50%. A refresh rate of anything less than 60Hz will undoubtedly lead to serious perceptual issues with flicker, as was witnessed in the early days of desktop active stereo, where products were only capable of 30Hz (or worse). Even at 60Hz, flicker is not eradicated for some end users. Early LCDs (TVs, PC and laptop screens) were unable to refresh at equivalent CRT rates, which made active stereo almost impossible. In order to match the 120Hz refresh rates of CRTs, each pixel in a flat panel or LCD display must be capable of making the transition from dark to light to dark in 8.3msecs or less. Companies such as Samsung and RealD (who supply the largest number of circular stereo projection units for cinemas worldwide) have developed shutter glass technology for integration via an active circular polarisation LCD panel which sits in front of the source display, as opposed to integrated with the eyewear. The companies claim that the RealD technology is capable of presenting each eye with a full high-definition image, something that is not achievable using, for example, FPR passive 3D TVs (described above).

"Crosstalk" still remains a problem however. Crosstalk can occur when the active stereo system fails to achieve an adequate speed of right-eye/left-eye image switching which can result in image ghosting in both eyes. Crosstalk can also result from data compression and transmission distortions, resulting in a reduction in image quality and visual comfort and an increase in perceived workload (Tsirlin *et al.*, 2011). Tsirlin and his colleagues also conducted a range of experiments to assess the effect of crosstalk on depth perception. They showed, with a direct depth estimation task, that the amount of perceived depth decreases in the presence of crosstalk. For all disparities perceived depth was reduced by about 20% at crosstalk level of 8%. Beyond 8% depth was reduced at increasing rates especially for larger disparities.

Pros:

- Active stereo systems can work with a single projector.
- Active stereo is associated with less (but not totally eliminated) image ghosting.

- Full resolution images; an active eye-glass type 3DTV will have (1920 x 1080) for each eye (compare this with the resolution reduction for polarising techniques above).

- Active stereo systems are colour neutral (unlike anaglyph images, for example).

- Active 3D glasses allow full colour and picture information since both images are not overlayed.

Cons:

- Active stereo glasses can be expensive and require batteries or recharging.

- Active stereo glasses are, in the main, incompatible between different manufacturers' products. However, the European company X6D (marketers of various 3D viewing products under the brand name XpanD) has recently announced a line of "universal shutter glasses products", capable of automatically recognising a particular active stereo PC, TV or even cinema display system and configuring the glasses accordingly. The list of compatible display systems is growing steadily[1].

- Cable-less synchronisation techniques can be sensitive to interference and can experience periods of temporary and distracting inactivity if the wearer's head movement prevents the sensor from receiving the synchronisation transmission signal.

- Active stereo glasses effectively prevent light reaching the eyes for half of the time.

- Some LCD eyepieces are actually polarised (when not blocking light), which makes the scene being viewed slightly darker, even when the shutter effect is inactive.

- Some users notice flickering even at a display refresh rate of 120Hz (60Hz image presentation). Others can also be sensitive to subliminal display flicker.

  Adverse flickering effects can be experienced in the periphery of the wearer's vision, especially if the environment in which the system is being used contains bright light sources or other forms of display.

3.3 Autostereoscopic Displays

Like many of its counterparts in the stereoscopic viewing arena, the concept behind the autostereoscopic display is by no means new. The idea of generating binocular images by dividing them into narrow vertical stripes and then viewing them – without glasses of other form of eyewear or headgear – through a fine grating (or "parallax barrier") was evident in the early 1900s. Another technique, based on placing a series of lenses at the surface of a picture, was also proposed around this time, although the lenses were not lenticular (as is

---

[1] Note that in March 2011, Panasonic and X6D/XpanD announced the "M-3DI Standard", the aim of which is to encourage manufacturers of active shutter stereo glasses to make their products compatible across a wide range of 3D-capable products (TVs, projectors, PCs, laptops, etc.).

evident today) but more spherical in nature, producing a "fly-eye" effect. Lenticular techniques, based on sheets of fine cylindrical lenses, flat on the back (i.e. the focal plane), were identified later as having the potential to generate stereoscopic effects by being placed over an image source to refract the rays of light from alternate left-eye, right-eye image columns to each eye. Holographic optical elements may be used instead of lenticular lens arrays. Today's autostereoscopic 3D displays exploit either of these optical solutions to produce an output such that at least two illuminated regions, or "windows", in space can be viewed by each eye of an observer (Figure 11). If these regions actually form a stereoscopic pair, then binocular 3D will be perceived. In the lenticular lens, an array of cylindrical lenses directs light from alternate pixel columns to the two viewing regions, allowing each eye to receive a different image at an optimum distance. In the parallax barrier technique (which is used to generate the 3D effects in the Nintendo *3DS* hand-held games console), a mask is placed over the LCD display which directs light from alternate pixel columns to each eye. In the case of some displays, the barrier takes the form of a switchable LCD. When the parallax barrier LCD is switched off, the image generation LCD displays standard 2D images. When the parallax barrier LCD is switched on, alters the placement and/or width of the crystals in the barrier, sending a different set of images to each of the observer's eyes. This form of display has already found favour in transport security checkpoints, where, it is claimed that the switchable quality of the parallax barrier technology reduces the likelihood of false alarms and improves detection rates are improved. As well as the type of technology generating the binocular imagery, there are, broadly speaking, three types of autostereoscopic display techniques. They are:

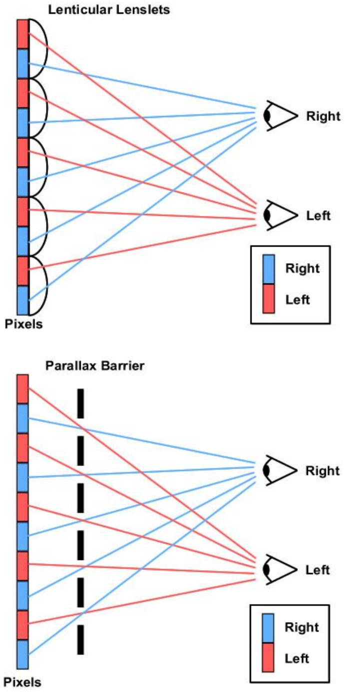


Figure 11: Lenticular and Parallax Barrier Light "Guides" Used in Autostereoscopic Displays
Source: http://www.hometheater.com/content/3d-glasses-free-last

- ***Two-View/Window Autostereoscopic Display*** – This type of autostereoscopic display system operates as described above, for a single viewer. The two windows that provide the left- and right-eye images for the observer to fuse are primarily visible from a central and relatively constrained viewing position. Indeed, the observer may have only 20 or 30 mm of head motion "freedom" around this central viewing position, beyond which the 3D effect deteriorates sharply. Typically this type of display delivers high resolution viewing and is of a relatively low cost.

- ***Multi-View/Window Autostereoscopic Display*** – Multi-view autostereoscopic (sometimes referred to as "automultiscopic") displays generate more than two views simultaneously, in multiple viewing windows. Viewing any two of these simultaneously generates a strong stereo image. Multi-view displays support a much wider lateral viewing zone than the two-view technique, which also means that they can deliver simultaneous stereo images to multiple viewers. Autostereoscopic displays of this type have lower resolution than their two-view counterparts, as the underlying display has to be divided into multiple views. Nevertheless, they, too, are obtainable at a relatively low cost. Because of the wider viewing zone, 2D-3D switching is a rare feature of multi-view displays.

- ***Tracked Two-View/Window Autostereoscopic Display*** – Tracked two-view displays aim to deliver the best features of the two techniques described above – the higher resolution views of two-view display and the wider viewing zone of the multi-view display. To achieve this, the two views have to be “steered” to follow the position of the viewer’s head or eyes (typically this is achieved using a webcam-like add-on to track the viewer and an optical steering mechanism subsystem within the display. These displays provide high resolution and wide viewing freedom but incur additional cost to implement the tracking and steering mechanisms. Currently these displays do not possess a 2D/3D switching ability, nor can they cope with multiple viewers.

Both the two-view and the multi-view technique require the viewer to sit at a specific “sweet spot” in order to ensure that each of the stereo images is pointing at the correct eye. Any sort of movement away from these sweet spots – including head tilting – will cause immediate blurring of the 3D image. Prolonged viewing in this manner could lead to severe eye strain. The requirement for the viewer(s) to remain more or less motionless at these sweet spots also makes the technology inappropriate for rooms where individuals need to move regularly between different zones. Camera-based eye tracking is likely to be the future solution to effective motion-tolerant autostereoscopic displays, but mature technologies could well be 5 to 10 years away.

3.4 Volumetric Displays

Volumetric displays (such as the concept shown in Figure 12, top image) present users with a 3D computer-generated or virtual object that possess three physical dimensions (x, y and z). In other words, the virtual object occupies an actual or real-world space and, being volumetric, supports viewing by a reasonably large number of observers from a wide range of angles. Furthermore, unlike its planar or screen-based counterparts, an object or scene presented using a volumetric display possesses consistent depth information. Volumetric displays are also autostereoscopic. Therefore instances of asthenopia (see Section 4.1), where discrepancies between human visual accommodation and convergence can cause perceptual and ophthalmic problems are, in theory, minimised. Volumetric displays, be they concept or actual, are often portrayed as a transparent sphere with 3D imagery “hovering” inside.

Figure 12: A concept C^2 Volumetric Display (Upper), Sony *Ray Modeler* (Middle) and Actuality *Perspecta* Volumetric Display (Lower)
Sources: Authors' Archives
and www.gizmodo.com

Broadly speaking, there are two classes of volumetric display – *static volume* and *swept volume*. Static volume displays operate on the basis of voxel (a "volume pixel") stimulation. Here, an addressable volume of space that has been created out of active elements (solid, liquid or gas) is opaque or luminous (or changes colour) when in the 'on' state (e.g. when stimulated by intersecting laser beams) and transparent when in the 'off' state. Swept volume displays use a rotating projection screen or mirror. As the projection surface sweeps through the display volume, it reflects or emits light synchronised to its location. If the volume is refreshed frequently enough (e.g. 20 volume sweeps per sec. minimum), then human persistence of vision can be exploited and the reflected or emitted light is fused into a single 3D image.

A commercial example of a swept volume volumetric display is Actuality Systems *Perspecta* system (Figure 12, bottom image; although the status of this device is, at the time of writing unclear, following the acquisition of Actuality's assets by Arlington-based Optics for Hire). Unlike the concept C^2 display shown in the top image of Figure 12, the *Perspecta* display only generates a 25-cm diameter spherical 3D volumetric image by sweeping a semi-transparent 2D image plane around the Y-axis. Each image (or "slice") consists of 768x768 pixels, and a total of 198 2D slices are uniformly displayed around the Y-axis, resulting in a total of 116 million voxels ("volumetric pixels"). The display's refresh rate is 24Hz. However, because the entire viewing volume is only being updated at 24 Hz (academic laboratory versions exist that are capable of much higher refresh rates), there is a noticeable flicker in the displayed image. There are two key issues with this type of display (although the extent to which these become evident in later products remains to be seen). Firstly, as the projector brightness is quite low, the room lighting levels also need to be kept very low if detail is to be discernable in the scanned image. Also, due to the way the projection screen is scanned, image slices that are separated by 180^{o} are mirror images of each other and, due to imperfect alignment (amplified by the scanning mechanics and the low refresh rate), appear to shimmer.

Sony's 360º autostereoscopic display, the *Ray Modeler*, is cylindrical in shape (Figure 12, middle image), 13 cms in diameter and 27 cms high (although the effective display space only occupies about a half of the height) and features LED light sources to enable viewing of full colour volumetric objects from all directions (360 different images in all directions, at 1 degree intervals). It also possesses a hand-activated motion sensor which supports simple gesture control of the display's orientation.

Examples of fully interactive volumetric displays are hard to find in the literature, and most seem to be dedicated to product advertising. Examples under investigation for real-world tasks are also hard to find, although, given the size (and price) of the commercially available systems, not to mention the lack of mature applications development toolkits, this is unsurprising. Another possibly limiting feature of volumetric displays is that the images they present "exist" within arm's reach of observers. This may be acceptable when displaying – and interacting with – engineering components, for example. However, exploded engineering or computer-aided design (CAD) views might well be too blurred to resolve visually, given the resolution and refresh rates available. Large-area virtual environments will require significant leaps in the technology before the fidelity becomes acceptable. It is interesting that, even before the technology is of a form that is capable of hosting credible real-world virtual environment applications, researchers are proposing different interactive techniques for individual and collaborative participation (e.g. Balakrishnan, *et al.*, 2001; Grossman & Balakrishnan, 2008).

Grossman & Balakrishnan (2006) also conducted research to assess the extent to which volumetric displays (and the *Perspecta* system in particular) presented superior 3D imagery to observers. The Actuality display was compared with three other display set-ups. For monoscopic image presentation, a 19-inch 120Hz CRT monitor was used (with 3D scenes displayed as perspective projections). For the planar stereoscopic image presentation, the same CRT monitor was used, but in conjunction with LCD shutter glasses (each eye receiving a 60Hz refresh rate). In an additional active stereo condition, the user's head was tracked in real time using a proprietary VR tracking system. This was set up in such a way that the 3D scene would fade if the head position deviated from a fixed starting position by more than 25 cm (note that the same viewing angles were used for all displays). Three tasks were used in the comparison of the display technologies – a depth ranking task (judging the depth in 3D space into computer-generated sphere), a path-tracing task (involving judgements relating to identifying minimal paths between a dense pattern of interconnected nodes) and a dynamic collision task (requiring participants to judge whether or not two approaching objects would collide or pass each other).

The results of the experiments showed that, for depth ranking, volumetric displays provide superior depth perception in comparison to stereoscopic and head-tracked stereoscopic displays. The volumetric display did not perform as well in path tracing (although this may be explained by the technology's refresh rate and related scanning artefacts). In the dynamic collision task, the volumetric display provided the best result, although the difference between this display and the stereoscopic display with head tracking was not significant. The volumetric display had a significantly lower error rate than the stereoscopic only display. The high error rate for the perspective (monoscopic) display was significantly different from all other displays confirming that the task was practically impossible without any stereo or motion cues.

3.5 Head-Mounted Displays (HMDs)

The following is an extract from a recent document published under the HFI DTC programme (Stone, 2011):

"Today, HMDs are experiencing something of a revival. The rapid growth of mainstream and serious games markets seems to have prompted "cottage" developers and large companies alike to – once again – make all manner of wearable technologies available to an uninformed end user community. However, it is fast becoming apparent that many of the commercially available products that are in existence today, not to mention those being touted for future release, suffer from the same Human Factors issues as their VR predecessors of the 1990s – low resolution, small fields of view, large visible areas of optical housing, inadequate provision for wearers of spectacles, poor build quality, lack of ruggedisation, and so on. The same revival is being driven by developments in Augmented Reality (AR), where the headset-based displays, when integrated with the output of miniature head-worn cameras, support the fusion of real-world and virtual imagery. The aim of this is to provide the wearer with real-time information superimposed on the real-world view that is normally invisible or can only be visualised using other forms of media and/or interactive display technologies".

It is not proposed to include a detailed overview of the stereoscopic capabilities and limitations of HMDs. Instead, the reader is referred to Section 3.4.2 of Stone (2011) for example uses of HMDs and cautionary comments based on previous case studies.

4 Human Stereo Vision – Dysfunctions and Viewing Problems

Some cautionary remarks need to be raised at this point relating to the effective use of the above technologies by human observers. During the course of the literature review supporting this study, it has become apparent that visual, or specifically stereoscopic "comfort", is an issue that features in a number of studies and is certainly becoming an online topic of discussion relating to the alleged acceleration of uptake of 3D TV, computer and gaming devices. As with many of the experimental and review papers relating to the effectiveness of 3D and stereoscopic displays, concrete agreements relating to the *actual* incidence of stereoscopic dysfunctions is elusive. Tam *et al.* (2011) point out that "visual comfort of stereoscopic images has been a long standing problem in stereoscopic research". They go on to state that "studies have found differences in terms of individual's tolerance of visual discomfort and fatigue. It is unclear whether these differences simply reflect normal differences in visual processes or are linked to some form of stereo-anomaly". Furthermore, the interrelationship between stereo discomfort, motion (vection) and simulator sickness is uncertain, although guidelines for simulator sickness (or *cybersickness*) certainly exist (see Kolasinski, 1995, cited in Stone, 2001, for example).

One of the key problems faced by the 3D community generally is that, whenever an apparently "new" technology is launched onto the market, then the warnings and stories of negative impacts on end users tend to be brought to the fore, via trade magazines, online articles and even academic institutions attempting to raise their profile, without a strong scientific understanding of the real human factors issues. Already, one university has suggested that prolonged 3D viewing could impair the distance judgements of the viewer and result in subsequent driving fatalities. This, too, was raised as an issue in the early 1990s, when the number of public VR arcades was growing significantly. It was even suggested that a high-speed car crash on the German Autobahn was caused by relatively short periods of VR headset-based gaming in one of these arcades (although this was never confirmed and grew rapidly to become an urban myth).

The "great stereoscopic defect debate" accelerated in 2010 when Samsung issued a comprehensive warning about possible health effects when watching 3D TV. Sony and Nintendo have done the same for the *Playstation 3* monitors and *3DS* products. Interestingly, this parallels a similar event in the 1990s when Sony, albeit some time after the device had been on the market, issued a similar health warning for its *Glasstron* HMD product in the late 1990s (a warning that occupied some six pages in the product's manual!). No doubt there will be similar warnings for the company's latest 3D Visor, which, despite delivering nothing significantly new in the commercial HMD viewing market, is being touted as a display system that will "change 3D forever".

At the time of writing, then, what is known about human perceptual issues with stereoscopic or 3D displays forms, in broad terms, two areas of concern – actual physiological issues (i.e. medical disorders that prevent the eyes focusing and/or aligning correctly) and the problems caused by poor design of content and of the delivery technologies. As will be seen later in this report with the 3D display literature itself, the extent of the problem varies considerably from paper to paper and reporter to reporter.

4.1 Stereovision Problems 1: Physiological Conditions

The general consensus within the published and online literature is that 4 to 10% of the population, when presented with 3D content, will suffer from asthenopia. This manifests itself through nonspecific symptoms including fatigue, pain in or around the eyes, blurred or dimmed vision, headache, nausea, dizziness and occasional double vision. Compare this with colour blindness, where the figures quoted are around 0.5% for females and 8% for males (note that most colour blind viewers will be able to see 3D effects but will experience a reduced colour spectrum).

Less than 5% of the population have *severe* visual disabilities which make seeing in 3D difficult or impossible. This group includes those with medical diagnoses of amblyopia ("lazy eye", or reduced visual acuity in one eye) or strabismus (where the eyes look in different directions – "crossed eyes" or "wandering eyes"). Individuals who fall under these categories either cope in the real world by exploiting motion parallax or monocular cues, such as shadowing, focal depth, texture gradient, geometric and aerial perspective and geometric overlap (or interposition). Added to the 4 to 10% of people with inability to see stereoscopic 3D, another 10 to 20% of individuals will present symptoms similar to motion or simulator sickness when exposed to static and dynamic 3D imagery, as is evident in Virtual Reality implementations (see Stone, 2011). Typically this form of sickness begins with eye strain, disorientation and a headache or general sense of fatigue, and can lead to nausea.

One particular concerning study (Montes-Mico, 2001) suggests that, whilst not severe (as defined above), 56% of the population who are between 18 and 38 years of age have one or more problems with binocular vision and therefore could – at times – have difficulty seeing 3D. Stereo viewing capabilities begin to diminish in the fourth decade of life, but, even here the literature is contradictory.

For these reasons, it is important to pre-test any users prior to their exposure to a 3D or stereoscopic display system, remembering that stereo blindness does not preclude individuals from using standard display techniques, nor does it compromise their ability to exploit other, monocular cues to depth and distance in real-world and virtual scenarios.

4.1.1 Testing for Stereo Blindness – Stereoacuity

"Stereo blindness tests" are available online, but, in the main, these tend to be based on simple diagrams and illustrations from popular media, as opposed to clinical tests (one interesting exception is an online large-scale stereovision test being undertaken by researchers at McGill University in the US (http://3d.mcgill.ca/)). One of the key concerns with many of the tests that are available (including those clinical devices summarised below) is their relevance to (a) measuring individuals' abilities to deal with 3D and stereoscopic images and information presented using a variety of 3D display technologies (such as those described above) and (b) to 3D perception – binocular and monocular in the real world.

Stereovision tests typically measure stereoacuity, in other words, the smallest depth difference (measured in arc seconds) an observer can perceive. There are, in essence, two groups of clinical tests used to measure stereopsis and stereoacuity – "random-dot" and "contour" or "displacement" (Fricke & Siderov, 1997). Research by Hofstetter & Bertsch (1976) suggests that the mean stereoacuity threshold for a population is 14.4 arc seconds. Their research suggests that 98% of the population should have a stereoacuity range between 2 and 38 arc seconds. Based on this, 40 arc seconds may be used as the general pass/fail cut-off for adults, even though some stereo tests set a 20 arc second target (which is the smallest measurable stereo acuity threshold a clinical test of the sorts described here can measure).


Figure 13: Randot Stereo Test
Source: www.sussexvision.co.uk

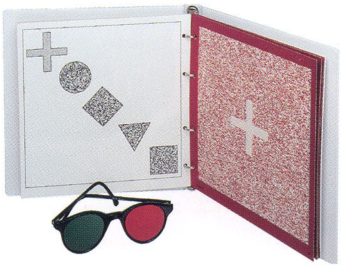


Figure 14: TNO Stereo Test
Source: www.sussexvision.co.uk


Figure 15: Titmus Stereo Test
Source: www.sussexvision.co.uk

Random-dot images (or "stereograms") were first used by Julesz in the late 1950s, as a precursor to autostereograms (made popular in the 1990s by the *Magic Eye* series of books). Random-dot stereograms show that, in the absence of familiar objects, perspective or any other form of monocular cue, binocular fusion can still occur and stereoscopic depth can still be perceived (Julesz, 1971). Examples of well-established random dot tests include:

The **Randot Test** (Figure 13), which requires the observer to wear polarised spectacles. Random dot patterns evaluate stereo depth perception by requiring participants to identify six geometric forms from random dot backgrounds (500 to 20 seconds of arc) at a distance of 40cm.

The **TNO Test** (Figure 14), which uses random dot stimuli with red-green anaglyph glasses to separate the images presented to each eye. With this test, monocular clues are absent and the target image is not outlined by monocularly visible contours. Evaluation trials have shown the TNO Test to be the most sensitive measure of amblyopia (Farvardin & Afarid, 2007; Ohlsson *et al.*, 2001).

Contour stereo tests evaluate stereovision capabilities by presenting individuals with two horizontal disparate stimuli. Examples of well-established contour or displacement stereovision tests include:

The "**Titmus**" Fly/Butterfly & Circle Test (Figure 15), which uses black, contoured stimuli together with polarised glasses to separate the image components for presentation to each eye. The stimuli are finely-separated circular patterns (40 to 800 seconds of arc) and grossly-separated components making up an image of a fly (where the wings appear to be closer to the viewer). On the Titmus Stereo Fly Test, if the observer can see 9/9 targets, then s/he has a stereoacuity of 40 arc seconds.

Note that versions of some of the tests listed above (e.g. Randot, Titmus) are available without the need for the observer to don polarising glasses. Instead, the tests have been recreated using a prismatic printing process to ensure that separate images are presented to each eye. Hatch & Richman (1994) found that there were no significant differences between the two types of test and concluded that the non-polarising versions are just as valid in measuring stereopsis as their traditional counterparts.

The Near-Field **Frisby** Test consists of three square transparent plastic plates (6, 3 and 1mm thicknesses) onto the front of which have been printed four similar patterns (Figure 16). In the central region of one of the four patterns is a circular area, which is printed on the reverse side of the plate and can, depending on the stereoacuity of the observer, appear in depth. By presenting each plate at different distances, retinal disparities of this circular area of between 600 and 7 seconds of arc can be achieved. Monocular parallax cues can be avoided by preventing observers from moving their heads during testing.

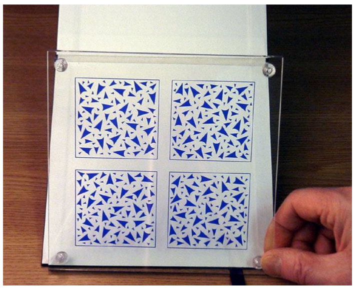


Figure 16: Near-Field Frisby Stereo Test
Source: www.frisbystereotest.co.uk

The **Frisby-Davis Distance** (FD^2) stereo test uses four small objects presented to the observer inside an open fronted box, located 6 metres away. The test relies on defining the threshold for detecting which of the four objects is at a different distance from the observer. This, the developers and others claim, is the closest a stereo test comes to being able to generate measures of ‘‘real world’’ stereoacuity. However, environmental changes can drastically affect real-world stereoacuity (including weather, lighting, dynamic activity and sophisticated forms of camouflage), not to mention the presence of other potentially conflicting monocular and illusory/transformational cues, including motion and associated blur, intervening media, such a cockpit or windscreen.

Most researchers agree that it seems improbable that stereoscopic ability in the real world can be predicted from static clinical stereoacuity tests. Therefore their use should be restricted to screening individuals with mild-to-severe stereoscopic dysfunctions, with the aim of excluding them from exposure to 3D/stereoscopic displays or limiting the tasks they have to perform with such displays.

4.2 Stereovision Problems 2: Poor Content and Delivery Design

As well as individual stereo vision dysfunctions, 3D and stereoscopic displays that attempt to exploit human binocular vision have been (and continue to be) accompanied by reports of asthenopia of varying degrees, brought about by the discrepancy between accommodation and convergence. The discrepancy arises because, whilst accommodation is fixed on the depth of the display surface (i.e. where the light is originating from and, sometimes, courtesy of the display’s physical mounting), the eyes are converging at a distance that is dictated by the perceived depth of the object being fixated. In real-world scenarios, accommodation and convergence always act together produce identical depth information.

This discrepancy is also known to be more pronounced the closer one is to a screen. Hence, fewer issues of eyestrain, fatigue and nausea, are reported by cinema audiences than those watching domestic 3D TVs or viewing virtual environments via HMDs. Emoto *et al*. (2004) measured viewers' fusional amplitudes[2], the interaction between convergence and accommodation, and subjective visual fatigue after 1 hour of viewing conventional 2D and stereoscopic 3D TVs. They found that stereoscopic viewing causes more serious visual fatigue than monoscopic viewing. They also found evidence of decreased fusional amplitude after viewing stereoscopic images, but the decrease recovered after a short period (around 10 minutes) of relaxation.

Some of the "extreme" effects one sees when viewing 3D films, such as objects frequently being "thrust" out of the screen, apparently close the audience's viewpoints, will contribute to viewers' malaise, since these objects will require the eyes to cross to keep the objects converged (whilst the focal point on the screen remains the same). Indeed, HMDs (and COTS devices in particular) add to the problems, courtesy of their face-enclosing properties, narrow fields of view, optical assemblies and low resolution displays. Emoto *et al*. (2004) point out that, since many HMDs employ convex lenses and small screens to achieve a "big screen" effect, this can place a considerable burden on the observer's visual accommodation system. In addition, inappropriate fitting of the HMD and prismatic effects caused by mismatches between the observer's interocular separation and the optical centres of the lens assembly may place additional strain on the convergence mechanism. Furthermore, any motion artefacts or lags that may be evident in the method by which the wearer's head is tracked will exacerbate feelings of malaise. Note that, with many COTS products, head tracking is not provided, in which case the HMD effectively becomes a static head-mounted viewing screen. This can cause even more problems of disorientation, especially as cinematic stereoscopic productions may not be ideally formatted for viewing with HMDs or other forms of eyewear.

In an excellent and highly readable overview entitled "Stereoscopic 3D Film and Animation – Getting it Right", Kenneth Wittlief (2007) outlines some of the key issues that are important when attempting to set up a stereoscopic viewing facility. Wittlief's "rules" help to avoid some of these convergence-accommodation mismatches and related perceptual artefacts. Whilst his recommendations focus on the cinematic environment, they are just as valid to other to other environments, from small rooms to C^2 centres, and to different media, film media to computer-generated imagery (CGI).

Rule 1: Define how the images are to be viewed and the environment in which they are to be viewed. Before creating the virtual space in front of the viewer(s), it is important to understand what tasks the observers are required to undertake so that, when capturing or

---

[2] Typically measured by changing the angular separation of two images, each of which consists of similar background scene, but each of which also contains a unique left-eye, right-eye object. If the observer sees the two unique objects, together with the background scene as one image, then the images are correctly fused. The angle between the images can then be increased or reduced until the patient is no longer able to see one image with both controls. This gives a measurement of the patient's fusional amplitude.

developing the 3D imagery, those images appear comfortably within the virtual space that will exist in front of the viewer.

Rule 2: Human 3D vision is effective out to ranges of around 180-200m (Braunstein, 1976). For large area or deep-range stereoscopic representations, *hyperstereo* effects are often employed by increasing the distances between the real or virtual camera views beyond that of the average viewer's inter-pupillary distance (IPD). This has the effect of re-introducing shape and depth into an otherwise flat scene background. However, hyperstero representations must be executed with care, in order to avoid the "doll's house effect", whereby the viewed scenes take on the look and feel of a model. Consideration again needs to be given to the nature of the human's task – different camera separations will be required depending on whether the task requires wide or narrow area imaging, ranging from small densely-packed objects that may represent the contents of a close-range IED (for example) to a terrain display that needs to be monitored for logistics simulation purposes or UAV status updates.

Rule 3: Design the 3D space and display technique to support the type of viewing the observers will be undertaking (e.g. search, focused attention, etc.). As pointed out by Wittlief (2007), observers new to the 3D or stereoscopic experience may well spend much of their early sessions visually exploring displayed scenes, as opposed to concentrating on the intended objects of interest. Wittleif's solution is to set the cameras to converge on the most distant objects in view, adjust the separation so that infinity is around 5 cm apart at the screen, and, as he puts it "let the foreground objects find their own place in that space". Wittleif continues "resist the temptation to converge ... cameras on the centre of attention ... to lock the viewer's attention on one area ... use a depth of focus effect to blur the rest of the image, so the viewer will not be inclined to look around ... at other things".

Figure 17: Typical CAD Display
Source: Authors' Archives

A related display issue to take into consideration is that of "frame" or "edge violation" – often seen when using poorly set-up stereoscopic projections or displays of complex computer-aided (CAD) databases, as might be the case with the image in Figure 17, if such a violation occurred. Frame violation occurs when objects that are being viewed under negative parallax conditions (i.e. the observer's eyes are converging on objects that appear to be in front of the screen – "crossed disparity") appear in front of the screen, but are truncated by the screen's frame itself. The truncation suggests that the image is "disappearing" behind the screen, but the negative parallax is generating a strong cue that it should be in front of the screen. Real-time CAD models of industrial plant (for example), when explored stereoscopically, can produce a wide range of confusing effects, especially as pipes and vessels appear to come closer to the observer and are then "clipped" by the stereo monitor's frame. One solution by filmmakers is to implement a "crop mask" on the left edge of the left image and the right edge of the right image, thereby creating an illusion of the screen frame being on top of the stereo image.

Rule 4: Ensure that the real or virtual material being developed for stereoscopic presentation takes into account the specifications and limitations – scanning/frame/refresh rate, for example – of the stereo/3D displays being used, or proposed for use. This is particularly important for moving images or objects (motion across or out of the screen). If the scanning rate is not high enough, then, given that the observer's focus is fixed on the screen itself and convergence cannot function effectively alone, the 3D effect can be lost (this was mentioned in the earlier discussions relating to volumetric displays, but applies equally to other devices).

5 The Pros and Cons of Human 3D/Stereo Vision

It is not the intention of this report to present a detailed overview of human 3D or binocular vision, as these topics are more than adequately covered in standard academic texts and within numerous online medical and educational sites. Some of the key issues have already been presented in Section 4, relating to stereovision dysfunctions and problems, as these are important issues to be aware of when considering the implementation of a 3D or stereoscopic display for any form of application, be it entertainment-focused or not.

There is no doubt that stereoscopic or binocular vision brings with it a range of benefits and advantages – this is, from an evolutionary perspective, irrefutable. However, the key issue is one that has been apparent for many years, and certainly since the first stereoscopic TV systems found their way into significant and often safety-critical domains, such as the nuclear industry. **No matter how much scientific and academic knowledge relating to binocular vision is developed in laboratories across the globe (and even a cursory literature search will expose literally hundreds of such references), as soon as 3D or stereo viewing devices make an appearance in real-world settings, individual, group and operational problems come to the fore and, more often than not, the technologies fail to gain adoption**. Partly, this is symptomatic of a "technology pull" attitude towards adoption. It is evident that few studies exist which describe the use of stereoscopic displays as the outcome of a strong and early Human Factors analysis, for example addressing tasks and contexts, and the appropriateness of 3D content and display hardware. Furthermore, very few integrated guidelines documents exist which support Human Factors specialists in their analysis of tasks and contexts in order to arrive at a reasoned set of judgements supporting the use of 3D and stereoscopic displays. This is an area that requires urgent attention if the benefits of 3D or stereoscopic vision, some of which are listed below, are to be exploited in real-world applications and settings.

5.1 Generic Benefits of Binocular/Stereoscopic Vision

The following generic "benefits" are culled from the literature search undertaken in support of this report (see Reference List at end of report). Specific applications and associated findings are presented in Section 6.

- Humans have a maximum horizontal field of view of approximately 200° with two eyes, approximately 114° to 120° of which makes up the binocular field of view (seen by both eyes), flanked by two monocular fields (seen by only one eye) of approximately 40°.
- Relative depth judgement. Within the 120° binocular field of view, the spatial relationship of objects in depth from the viewer can be judged directly.
- Accurate depth judgement. Stereoscopic vision supports fine or skilled activities that require accurate depth perception at close distances (e.g. surgical actions – suturing, specialised tool usage, key anatomical structure positions, general tool usage, etc.).

- Spatial localisation. The brain is able to concentrate on objects placed at a certain depth and ignore those at other depths using binocular vision.

- Some surface material and texture properties can be perceived as luminance or colour differences in each eye, thus supporting stereoscopic presentation for fine detail tasks, such as inspection or analysis (e.g. Thomas *et al*., 2002).

- Judgement of surface curvature can be interpreted more effectively with binocular vision.

- With complex images it has been found that stereopsis enhances 3D spatial judgments when monocular depth cues are ambiguous (e.g. lighting, shadowing and perspective ambiguities). Stereopsis appears to be a compelling depth cue except when in conflict with motion or occlusion.

- Stereoscopic viewing is important in conditions of static or dynamic visual interference, such as poor resolution, motion blur, display scanning artefacts, and so on.

- Related to the above, stereoscopic viewing is also an important feature in *binocular summation* (e.g. Campbell & Green, 1965; Blake & Fox, 1973), enhancing visual acuity and the detection of faint objects (slightly), improving light detection thresholds (especially under conditions of low contrast and reduced illumination - important for night driving or night operations) and even camouflaged object detection.

6 Application-Specific Examples – Evidence For and Against 3D/Stereo Displays

The following sections present the results of a number of experimental programmes and case studies, mainly conducted in the 1980s and '90s. The studies have been chosen from a literature review of some 200 papers, with a primary focus on defence and aerospace studies and a secondary focus on evidence both for and against the use of 3D/stereoscopic displays. The review consists of a diverse range of projects addressing both hardware and software issues as they relate to 3D and/or stereoscopic data display. Throughout the review, an attempt has been made to highlight the key findings, and these are presented in **bold blue text**.

6.1 Air Traffic Control / Management

In very basic terms, air traffic control or management (ATC/M) requires each controller to be responsible for a three dimensional sector of airspace. The controller's job is to guide aircraft that enter into this sector efficiently and safely. They ensure that there is no risk of collision with other aircraft by enforcing standards of separation which dictate minimum lateral, longitudinal and vertical distances between the aircraft. As ATC/M is clearly a dynamic 3D and spatial awareness problem, it comes as no surprise that the potential use of 3D displays has received considerable attention from the research and development community. The papers reviewed in this section are also of relevance to those summarised under Section 6.4, relating to Command and Control.

6.1.1 3D Perspective Displays

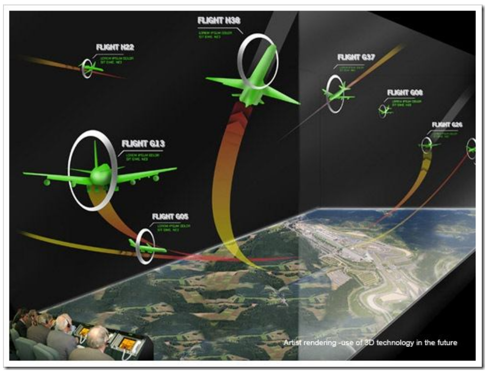


Figure 18: Artist's Rendition of a Possible Future ATC Perspective Display
(Source: http://eye-of-sky.com/)

In a report prepared for the US Federal Aviation Administration, Wickens (1995) described a series of six experiments to examine the effectiveness of 3D perspective displays for ATC use in such tasks as traffic flow management, conflict detection, evaluating pilot routing requests, terrain separation and weather. An artist's rendition of a possible future 3D perspective ATM system is shown in Figure 18. In all of the experiments, two displays were compared. The first was a conventional 2D "planar" display with a digital representation of altitude; the second a 3D perspective display which represented the airspace from a 45^{o} elevation. Wickens found that, overall, the results of the experiments showed **few differences in performance between the conventional 2Ddisplay and the 3D display**. Indeed, **where differences were found, they tended to favour the 2D display, principally for time performance measures rather than accuracy**. This is highlighted in Wickens *et al.*, (1995) in an ATC study where controllers were required to evaluate pilot requests for flight plan changes which might result in a mid-air conflict with another aircraft. This study found that the **2D displays supported faster approvals if the requests involved simultaneous vertical and lateral changes, but overall there was little difference in performance between 2D and 3D** displays.

Smallman *et al*., (2001) pointed out that both **2D and 3D display formats contain intrinsic deficiencies for representing three dimensions on a flat screen, due to *projective* or *line of sight* ambiguity**. In a 2D display, the altitude of an aircraft is ambiguous and has to be represented in another way, such as a digital readout located adjacent to the aircraft symbol. In 3D displays, all three dimensions are available and represented as actual distances on the display. This may spread ambiguity across all three dimensions and so designers therefore often add additional cues, such as shadows or drop lines to the ground to help with location uncertainty. Different ways of representing aircraft altitude and pitch in 2D and 3D displays were investigated by Smallman *et al*., (2001). Using a visual search task, they found that **performance, in terms of search time, with a 2D display was significantly better than with 3D**. They argued that, with a 3D display, altitude was confounded with distance, and pitch confounded with heading. Therefore, because of these ambiguities and distortions, a **2D display is preferred over 3D for any tasks that require precise spatial judgements**. Addressing the alleged superiority of 3D displays for appreciating the third dimension of scenes, they concluded that **3D format is less important than information availability, and that this benefit can be obtained from well-designed 2D displays**.

The key issue here is the use of the term “well-designed”, highlighting the fact that, even today, over ten years on from the Smallman *et al*. paper, very few human-centred design guidelines exist to provide support for the design of complex, dynamic traffic management displays.

Van Orden & Broyles (2000) compared the performance of air traffic controllers on several 2D and 3D display formats. During each trial, they completed altitude and speed judgement tasks, a vectoring task (similar to the transmission of bearing and altitude instructions from an ATC controller) and a collision avoidance task. Their results found that, **when using a 2D plan or side elevation view, the operators’ performance was as good or better for speed and altitude judgement tasks compared to a 3D perspective, 3D stereo, or laser based 3D volumetric display systems**. However, they found that **for collision avoidance the best display format was a 3D volumetric display**, as evidenced by higher accuracy and faster response times. The authors concluded that **users may benefit from 3D representations of data for tasks requiring integration and prediction of moving display elements within limited spatial areas**. When compared to 3D stereoscopic and perspective displays, the veridical[3] display of localised spatial information within a volumetric display may also provide high fidelity stereoscopic and parallax cues, improving human performance for some specific tasks. However, May (2000) cautions that using veridical displays may also introduce visual artefacts and unpredictable sources of information, possibly leading to erroneous interpretations on the part of the end user.

---

[3] A term often used to describe scenes that are identical to a natural scene filmed with a camera.

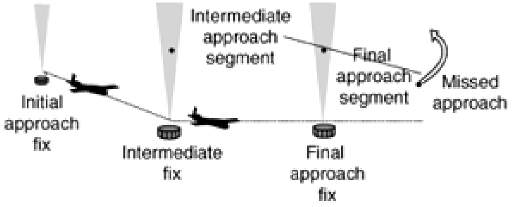


Figure 19: Final Approach Fix Point During Aircraft Descent
Source: www.answers.com

In a small-scale ATC study, Rozzi *et al.* (2007) interviewed four ATC controllers after they had performed a simulation task involving the management of aircraft in a holding stack and intercepting the final approach fix (just prior to intercepting the glide path – Figure 19). In this task the controller had access to a 2D radar display and a 3D display. The study showed that **operators managed the traffic referring mainly to 2D radar display**. However, on occasions the **controllers used the 3D displays during the management of the holding stack**. In these occasions the focus was on anticipating relative position either between aircraft within and outside the stack (e.g. aircraft approaching/leaving the stack), or aircraft flying in proximity of the stack. The authors concluded that these results suggest that **3D information displays appear to improve awareness of the relative position between aircraft, airspace and landmarks** and that **3D displays may improve performance in tasks that require integration and computation of 3D spatial temporal information**.

In a series of experiments, Brown & Slater (1997) conducted tests with three display types: 2D Plan View Display (PVD), 3D stereo and 3D ("pseudo-3D) monocular display. Performances by novice and experienced ("expert") ATC operators were evaluated over four tasks:

- Judgement of azimuth angle and relative distance ("Task 1"),
- Selection of highest and lowest aircraft ("Task 2"),
- Selection of aircraft pair with lowest lateral separation ("Task 3"), and
- Conflict detection ("Task 4").

Quoting directly from the Brown & Slater report:

Task 1: participants were able to "**judge angles and distances more accurately using the 2D display than either of the 3D formats** ... the stereo-3D display gave higher accuracy than the pseudo-3D display".

Task 2: "It was anticipated that the graphical visualisation of height in the 3D display would give faster task performance than the 2D display. The **novice group was found to perform the task fastest using the stereo-3D display**, but with the 2D PVD the next fastest and the pseudo-3D display the slowest, and **display type was found to have no significant influence on the performance of the expert group**".

Task 3: participants were able to "**determine the pair of aircraft with closest lateral proximity faster and more accurately using the 2D display than either of the 3D formats**, with the stereo-3D display again giving better performance than the pseudo-3D display".

Task 4: Participants were presented with animation sequences of 90-second lengths, each consisting of 7 aircraft, and were asked to report imminent conflict events. Conflict detection time analyses showed that experienced ATC operators were able to detect potential conflicts significantly faster than the novices but that **conflict detection time was not dependent on the type of display used**. Further analysis showed that significantly more aircraft were incorrectly perceived as being in conflict, in the cases where subjects were using one of the 3D display formats. The evidence suggested that ambiguity of position along the display depth axis may have been a contributing factor (ambiguity effects being emphasised in the work of Smallman *et al.*, (2001), mentioned earlier). The results also showed that "clutter" was another factor, partly resulting from a large display scale was too large and from the fact that the 3D displays had more symbology than the 2D displays.

Of significant relevance to the present report, Brown & Slater concluded that:

> **"These experiences show the need to consider the operator's task requirements in the choice and design of a display format, and that 3D displays have additional design parameters which can considerably complicate their design".**

Stereoscopic 3D systems have been proposed for ATC within the Eurocontrol[4] complex (e.g. Lange *et al.*, 2003), specifically for systems that could support traffic allocation within sector (i.e. balancing traffic between two or more sectors), weather and terrain representation, general training purposes (Dang, 2003) and conflict resolution (Lange *et al.*, 2006). Papers reporting on the evaluation of these systems have found that **controllers performed better in terms of reaction time without detriment to accuracy for judgement tasks with 3D stereoscopic displays as opposed to 2D displays** (Dang, 2003). Bourgois *et al.*, (2005) also reported that **identifying targets was quicker in 3D stereo display conditions** and, indeed, that controllers subjectively rated their performance to be better in the 3D condition. However, these were reports of preliminary studies (cited as PhD theses) and further evaluation was recommended. As part of the same work involving collaboration with Eurocontrol Experimental Centre, Dang *et al.* (2007, cited Dang *et al.*, 2009), conducted an informal study to present controllers with a conceptual 3D representation of ATC-related information. According to the report the controllers thought that **3D representations could be a very suitable option for critical weather information**.

---

[4] A series of papers were published describing 3D system developed in collaboration with Eurocontrol for training and evaluation – however, little detail of any formal evaluations of the system can be found.

6.1.2 ATM / ATC – Summary

Reviews of the ATC and traffic management literature highlight conflicting outcomes for performance with 2D or 3D displays. Smallman *et al.*, (2001) point out that "much of the complexity and inconsistency in the burgeoning literature on performance comparisons between 2D and 3D displays may well stem from simple interface and artificial enhancement differences rather than from the core differences between in the way that the displays depict the scene" (p.52).

Another reason for differences in performance outcomes between 2D and 3D displays in ATC/M may be due to the nature of the task. Wickens (1994) proposed that tasks which require integration of spatial dimensions benefit from 3D views, whereas tasks requiring focused attention on one or two dimensions benefit from 2D views. St. John *et al.*, (2001a) argue that 3D views are most useful for tasks that require a general understanding of shape of 3D objects or the layout of scenes; whereas 2D is most suitable for tasks that require judging the precise distances and angles between objects, as distortions in 3D hamper judging relative positions (see St. John *et al.* (2001b) paper in the Command and Control section of this review).

Alexander & Wickens (2005) commented that studies such as those described above serve to highlight the task dependency of 2D-3D trade-offs: "the 2D ... display supports tasks involving precise, spatial judgments. The 3D ... display, on the other hand, is superior for those tasks which involve judgments across all three axes (i.e., conflict avoidance manoeuvring; or involving a shape understanding component). It is important to note that the 3D display has an inherent advantage over a 2D display ... due to its analog representation of the vertical dimension. Aircraft altitude, for example, must be represented by digital datatags within a 2D ... display".

6.2 Teleoperation

As mentioned earlier in this report, remote, safety-critical operations in the nuclear and subsea domains were amongst the first serious explorers of the potential benefits of 3D or stereoscopic viewing, with the aim of enhancing the human operator's sense of "presence" during teleoperation. Teleoperation, to use the words of MIT's Tom Sheridan, is the "extension of a person's sensing and manipulating capability to a location remote from him" (Sheridan, 1987; 1992a). A development of teleoperation, *telepresence*, represents the "ultimate" in human-system interfaces for controlling a remotely-located robotic system (or "teleoperator"). Telepresence, again to quote from Sheridan (1992a), is an integrated input-output system that enables the human operator to receive "sufficient information about the teleoperator and the task environment displayed in a sufficiently natural way, that the operator feels physically present at the remote site". Yet, even with the extensive research conducted into teleoperation and telepresence in the 1980s, from a human factors perspective, the experimental support for the benefits of stereoscopic viewing has been consistently contradictory (e.g. Yorchak, 1986; Sheridan, 1987; Stone & Mason, 1987, *op cit.*; Wickens *et al.*, 1989).

A regularly-cited paper in the telerobotics and remote viewing arena is that of Draper *et al.*, (1991), in which the authors describe three experiments, each increasing in degrees of remoteness and manipulation complexity. The first two of their experiments incorporated a Fitts tapping protocol, which involves tapping between targets (after the classic experiments which led to *Fitts' Law* in the early 1950s). In the first experiment, participants undertook the task by directly moving a stylus (held like a pen) between the targets in 4 viewing conditions – direct view with both eyes, direct view with one eye covered, via a monoscopic TV display and via a stereoscopic TV[5]. Results of this experiment showed **significantly better performance in terms of movement time and miss rates for direct viewing, and no difference between the monoscopic TV and stereoscopic TV conditions**. The second experiment investigated task completion performance between monoscopic TV and stereoscopic TV conditions, again using a Fitts tapping task, but this time with a teleoperated device (an Oak Ridge National Laboratory Advanced Servomanipulator), under four different levels of difficulty (based on tapping target diameters and separation distances). The results showed that there was **no difference between the monoscopic and stereoscopic viewing conditions at the lower difficulty levels but at the highest difficulty level there was a significant improvement in time when using stereoscopic viewing, but no difference in errors**. In the final experiment, three operators used a Central Research Laboratories Model M-2 manipulator to reach into a box and to insert two 7-pin connectors into male sockets. The results showed that the **participants completed the coupling task faster in the stereoscopic TV condition than the monoscopic condition**. Note, however, that only three participants were included in this study, and these were the most experienced participants from the second experiment.

Draper and his colleagues also noted that "the failure to find powerful and consistent [stereoscopic] TV effects may stem from sub-optimal performance measurement. Rate variables (like time) may be relatively insensitive to the presence of [stereoscopic] TV because users are able to adapt to [monoscopic] TV by changing strategies, reducing the quality of performance, or changing their level of effort. Perhaps performance has not been adequately measured for precise definition of the effects of [stereoscopic] TV. Most studies have examined outcome variables like time to complete and errors, but few include process variables which describe differences that occur during task performance like changes in end-effector trajectories during target acquisition".

The selection or, indeed, design of appropriate metrics and tasks for teleoperation research, especially where the human interface components under investigation attempt to satisfy some sensory quality of the end user (haptic, stereo viewing, spatial sound, etc.) is a perennial problem and is highly worthy of further Human Factors research efforts.

Two years on from Draper's study, Drascic & Grodski (1993) put forward two key benefits for the adoption of stereoscopic vision technologies in teleoperation, namely:

---

[5] The stereoscopic TV system used for these experiments was an AEA Technology TV^3 Camera, as described in Section 2.2 of this report.

- that **a greater likelihood would exist that the telerobot will be used**. For explosive ordnance disposal (EOD) applications, if a stereoscopic vision capability allows the operator to use a telerobot more skilfully, operators may be more willing to use it.

- that **a greater range of tasks would become possible**. Telerobots and their human operators may be limited because of limitations of monoscopic camera views. For example, without stereoscopic vision, it may be very difficult to use a telerobot to lower an X-Ray plate behind an improvised explosive device (IED), so this task is rarely attempted remotely.

The authors also drew attention to some of the costs of implementing stereoscopic vision for teleoperation, including:

i) hardware costs,

ii) operational costs (e.g. maintenance, camera alignment and matching, lack of appropriate lenses),

iii) user costs (e.g. acceptance – users are used to using monoscopic vision, training, comfort, unforeseen side effects), and

iv) social costs (i.e. percentage of population with some form of stereoscopic defect being unable to use the technology efficiently – see Section 4).

The authors conducted two experiments using a telerobotic EOD-like task under two conditions (Stereoscopic Vision (SV) and Monoscopic Vision (MV); Drascic & Grodski (1993)).

The first experiment employed 8 munitions technicians with various degrees of training and experience. They attempted 3 tasks:

i) manipulation of 8 blocks,

ii) positioning an EOD weapon, and

iii) lowering an X-ray plate between two briefcases.

No significant differences between SV and MV were found for all tasks. In the second experiment, eight expert EOD operators were used. These participants completed an identification task (to examine eight IEDs on a table and to determine which four devices were complete and would function). They also completed a manipulation task, weapon positioning task and X-Ray sensor placement task. The results showed **significant improvements using SV for the manipulation task, weapon positioning and X-Ray plate positioning but no significant improvement in the preliminary identification task**. **Objective ratings for both experiments showed that both novice and expert users preferred SV over MV and agreed that SV was superior for tasks requiring precision operation**.

Figure 20: Teleoperated HMMWV Example
Source: www.torcrobotics.com

Under contract to the US Army Research Laboratory, Scribner & Gombash (1998) conducted experiments with remote driving operation of a high-mobility, multipurpose wheeled vehicle (HMMWV – similar to that shown in Figure 20). Their experiments compared monoscopic viewing with stereo viewing, together with two field-of-view conditions – narrow (55°) and wide (165°). Their results showed that there were **no differences in participants' remote driving performance times, but that there were fewer driving errors in the stereo viewing condition**. Interestingly, of the 26 cases observed, 12 preferred the wide field-of-view stereo condition, 8 the narrow field-of-view stereo condition and 3 both the narrow *and* wide field-of-view monoscopic conditions, even though self-reported stress ratings were higher in the SV condition and simulator sickness ratings were significantly increased for the wide field-of-view conditions. The authors go on to recommend that **stereoscopic viewing systems should be employed when remotely controlling platforms to traverse unfamiliar terrain**. They also recommend that "enhanced field-of-view" technologies, "based on trends in the data", should also be used, namely multiple overlapping camera views or head-slaved devices. This appears to be a somewhat radical recommendation, given the scope of their reported research, but they claim that "... the combination of these two technologies will provide the end user the greatest overall benefit for real-time, real-world use. The perception of a wide field-of-view through the use of a fast-response pan and tilt mechanism would provide essentially the same information as the overlapped camera views at one-third of the bandwidth "cost".

Building on some of the findings and recommendations from the Scribner & Gombash study, a review conducted by Chen *et al.*, (2007) addressed human performance issues and user interface design for teleoperated robots. They found that, for some applications, **stereoscopic displays may improve depth perception, obstacle avoidance, navigating difficult, unfamiliar and complex terrains and remote arm manipulation**. However, the review highlighted that such systems may induce motion sickness, specifically when using hyperstereo, which, by exaggerating the stereoscopic effect through the adoption of larger inter-camera/inter-lens separation distances may well have multiple negative effects on human teleoperation performance. The review suggests that military telerobotics users (primarily land-based systems) should be provided with an optional mode for complex terrains and remote manipulation tasks, although the cost of providing this (without hard, conclusive evidence) may well deter future procuring agencies from specifying such an option.

In a rather unique application of teleoperation technologies, Lim & Quek (2002) conducted an experiment addressing remote container landing using cranes. Setting up a scale model of the scenario, they tasked their participants to land a container vertically. Three viewing conditions were investigated:

i) direct view,

ii) 2D "quad" view (using four cameras mounted directly above the container grappling mechanism to relay a video-composited view of the four corners of the containers to the operator), and

iii) stereoscopic 3D view using polarising glasses.

Measuring the container landing forces, the results showed that the **landing force was lowest in the direct view condition and that stereoscopic 3D delivered lower landing forces than the 2D viewing condition**. However, when considering **overall task timings, stereoscopic 3D took the longest**. The study concluded that 3D stereo display showed promise for remote operation when direct views were not possible. However, this needs to be verified with real-world tests with more external validity, such as the inclusion of other cues which may affect the operator including shadows, sound, container dynamics and environmental conditions.

Crescenzio *et al.* (2009) developed and trialled a system for supervising unmanned air vehicle (UAV) missions using a 3D stereoscopic and augmented display where the user could switch between an internal (pilot) and external view point. Their paper presents the results of trials with three different levels of automation (manual, semi-automatic and automatic). The evaluation was solely subjective with measures of situation awareness, workload and interface design. **With regard to the 3D display, the authors note that it "was appreciated because of its realism", and rated well (approx 4.5 out of 5)**. However, there were no measures of actual task performance and no comparison with the augmented (or any 2D) display.

Park & Woldstad (2000) conducted an experiment comparing a multi-2D display (with 4 views: plan, right-side, left-side, front-side), a 3D monocular display and a 3D stereoscopic display in a **virtual simulation** of a robotic task, which involved controlling a robotic arm to pick and place a virtual object. The results showed that there was **no significant difference between the display types for completion time or distance travelled**. However, **there were significant effects found with respect to the number of errors and workload ratings where the multi-2D display condition performed best**. In addition, further investigations showed that adding certain visual enhancement cues, performance in the 3D display conditions became equivalent to the multi-2D display condition. These cues included:

- a solid reference line extending from the face of the robot gripper in the direction of any object to be grasped,
- a translucent reference cylinder surrounding the solid reference line and
- additional reference lines equally spaced around the solid reference line, which were intended to create a 3D volume to provide additional depth cues.

Kanduri *et al.*, (2005) conducted experiments showing how difficult it is to judge the heights of distant objects from monoscopic views of natural 3D scenes. This work has implications for remote mobile robotic systems used in field surveillance, search and rescue and scientific exploration (including planetary exploration). Two techniques were investigated. **In the first, involving direct height and distance estimation, observers overestimated by 190%**. **In the second technique, horizon analysis**, the observer indicates the position of the top and bottom of the object on the image, whereupon the height is calculated by measuring the visual angle between the theoretical horizon and points indicated. In this condition, **observers overestimated by 80%**. The results showed that even when provided with a rich set of supplementary and context information, operators have difficulty in perceiving the scale of distant objects.

As well as the earlier review by Chen *et al*. (2007) reported earlier, the researchers have also conducted experiments for the US Army Research Laboratory to investigate stereoscopic displays for real and virtual remote driving applications (Chen *et al*., 2010a; 2010b). The studies used two types of stereoscopic display types: field-sequential shutter glasses and passive polarised glasses. Both types of display system could also be used in a 2D mode, thereby enabling comparisons in human performance between 2D and 3D to be made. In brief, two scenarios were developed to conduct the 3D *vs*. 2D evaluation. For the first scenario, participants remotely operated a real Talon EOD robot through a course of cones on a grass terrain. In the second scenario, participants drove a remote vehicle through a number of simulated driving environments (developed using the VBS2 toolkit).

For the real-world remote driving task, **stereo vision resulted in faster course completion times than was found in the 2D viewing condition**. However, there were **no significant differences in driving error between the 3D and 2D conditions,** as measured by the number of cones on the course that were hit by the Talon robot.

For the simulated indirect-driving study, three tests – in effect virtual “driving courses” – were designed. The first course involved a set of “floating objects” in which participants were required to drive as quickly as possible along an enclosed, paved course, towards and around six sets of object pairs. The results of this test showed that **2D display of the virtual environment was accompanied by faster driving times. However, the 3D condition produced more accurate results in terms of the number of correct trials undertaken.**

The second simulation test took the form of an “obstacle course” in which participants were again required to drive as quickly as possible, this time around obstacles such as virtual rocks and shrubs. The results of this experiment showed **no difference in performance in terms of completion time and time off-course between the 2D and 3D conditions**.

The third simulation test was described as a “negative terrain course”. Here, participants were required to drive on an enclosed virtual course consisting of a variety of negative and positive terrain features (e.g. holes in the ground, drop-offs and hills). **In this test, 3D viewing was accompanied by faster driving performances than 2D, but no differences in driving accuracy were discovered as measured by time off-course**. Measures of workload showed no difference between 2D and 3D, although participants reported **more symptoms of simulator sickness after using the passive polarised stereo display**. In addition females reported significantly higher sickness severity and more oculomotor and disorientation-related symptoms than males.

From these experiments the authors concluded generally that "3D displays have advantages over traditional 2D displays, specifically for driving and robot teleoperation tasks". However, the results did not favour the use of 3D comprehensively. Of the four 2D *vs*. 3D tests, each with two measures of performance (time and error/accuracy), three measures showed better results in the 3D condition, one measure was better in the 2D condition and four measures showed no difference between 2D and 3D. As with many other studies, this raises questions regarding context specificity for using different types of display.

There have been anecdotal reports suggesting an interaction between stereoscopic video from remotely-operated systems and haptic (force/torque) feedback, particularly in cases where the environment being navigated (or the object being manipulated) exists under conditions of poor illumination and unfamiliar clutter. Unfortunately, the experimental evidence supporting this is also elusive, despite a number of articles describing the development of such integrated sensory systems (many of which present no user evaluation data whatsoever). Lee & Kim (2008) compared different set-ups of such an integrated system in a "tele-navigation" (remote driving) task. Haptic feedback – the generation of "collision-preventing forces" was computed from the range information delivered by a 16-sensor sonar array and relayed to the human operator by means of a 4D4M desktop haptic feedback system (a strange choice for remote driving, given the stylus-like nature of the device, which was similar in design and specification to the more popular *PHANToM* feedback product). Binocular imagery was provided by a desktop active stereo system. The task of the participants was to navigate the remote vehicle through a maze comprising walled areas and cylindrical obstacles.

Three independent variables were investigated, each with two levels: force feedback (off/on), display (mono/stereo) and screen resolution (low/high). The main (but not the only) dependent variables were: the number of collisions (the objective measure) and subjective measures of presence, realism and "embodiment" – ''the sensation of embodiment of an individual in a real life distant location'' (Paulos & Canny, 1997) – akin to the term "telepresence", coined by the likes of Sheridan (1992a;b). The subjective results indicated that each treatment condition increased the rating from the control condition. In other words, **presence, realism and embodiment ratings increased when force feedback was turned on, when the display was stereoscopic, and when the resolution was high**. This is not altogether surprising. However, for the objective performance measures, an interesting interaction emerged. In the absence of force feedback, a change from a monoscopic view to a stereoscopic view was accompanied by a significantly reduced number of collisions (i.e. **when no haptics, stereo is better than mono**). But when force feedback *was* active, there was no difference in the number of collisions between mono and stereo (i.e. **with haptics, mono and stereo are the same**).

Other papers (e.g. Basdogan *et al*., 2002; and Garcia-Robledo *et al*., 2009) are typical of the remaining literature, in that they describe remote systems possessing both stereo and haptic capabilities, but no user evaluations are conducted and no data are presented on performance.

6.2.1 Stereoscopic Systems for Remotely Operated (Subsea) Vehicles

Stereoscopic video feedback from remotely operated underwater vehicles (ROVs) was the focus of early studies during the *Bondi Initiative* of the late 1970s (mentioned in Section 2), although more attention was being paid at the time to features that might, today, appear to be very basic modifications to remote viewing systems, such as the introduction of high-resolution TV and colour (e.g. Westwood, 1981). By the time the Initiative had got under way, the use of stereoscopic TV for subsea exploration was still in its infancy (e.g. Berry, 1979). Even today, the use of remote stereoscopic viewing systems for subsea applications is quite rare, although some companies continue to pioneer the technologies, such as Kongsberg Maritime, who recently launched a “fourth generation” high definition (1920x1080 pixels) system, which, the company claims, is capable of interfacing with most COTS active and passive 3D TVs and PC monitors.

Of course, many of the binocular and monocular cues to depth and distance are missing or suppressed in the underwater environment, especially in conditions where an ROV is operating in conditions of minimal lighting and colour, combined with turbid water conditions, where the particles are either in static suspension or in motion as a result of thruster disturbance. Human Factors research conducted in the latter part of the *Bondi Initiative* (e.g. Stone, 1983) placed considerable doubt on the use of stereoscopic TV systems and their effect on ROV operator performance, especially in the poor water conditions around the coasts of the United Kingdom. At a time when most of the studies reporting positive results were conducted in ideal subsea scenarios – “blue water” conditions (such as the Gulf of Mexico) or in research tanks – the additional cost of deploying a stereoscopic remote viewing system simply could not be justified for many (if not all) of the companies operating “workhorse” classes of ROVs on this side of the Atlantic.

This early conclusion seems to have been borne out by other more recent (1990s) studies, specifically those from Australia and the Curtin University Centre for Marine Science and Technology. The work of this research team (e.g. Woods *et al.*, 1994 a,b; Woods, 1997) gives an initial positive impression of remote stereoscopic viewing for ROV operations (based on a polarising . However, on closer inspection of the results, it becomes clear that the laboratory-based experiments (with ROV operators but not using in-water viewing conditions) yield more objective data than the field trials using the same equipment, where the study authors place greater emphasis on subjective feedback from the offshore operators. Therefore, one can only really conclude that the suggested improvements put forward by the authors *may* be attainable by the use of stereoscopic viewing systems underwater, but only under ideal environmental conditions. These include:

- “easier” and faster positioning and alignment of manipulator-grasped objects (with less trial-and-error behaviours and errors than are evident in 2D viewing conditions),
- “better quality” viewing than is provided by 2D through suspended matter (fine particles) in the water (a signal-to-noise ratio improvement of about 3dB has been suggested),
- better spatial awareness of structured underwater worksites than 2D,

- assistance for ROV operators in the identification of unfamiliar or complex scenes (such as objects covered in marine growth),
- reduced workload and "frustration" when compared to 2D systems.

6.3 Surgery

Medical and surgical applications of stereoscopic or 3D viewing are included here, not because the domain is necessarily of immediate or direct relevance to defence applications, but that a considerable amount of research results from medical studies exist, most of which the authors of the present report deemed relevant to possible near-term and future specialised applications of defence, particularly counter-IED (C-IED) and EOD remote inspection and handling activities. It is not difficult to appreciate the close perceptual motor relationship between endoscopic or laparoscopic surgical procedures and EOD and telerobotic-mediated activities, and, given the advanced nature of some of the remote viewing and manipulation technologies now in use in various forms of surgery (including the successful *Da Vinci* minimally invasive surgical (MIS) system (Figure 21), one can expect the defence arena to benefit from the research findings, as EOD telerobotic systems become more and more sophisticated.

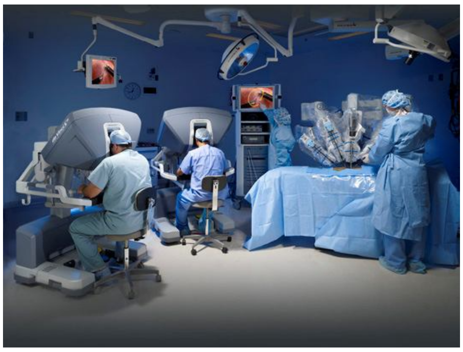


Figure 21: *Da Vinci* MIS Robot System
Source: www.islandoncology.com

6.3.1 The Rationale for the use of 3D in Surgery

A fine explanation supporting the aspiration to use 3D or stereoscopic displays in surgery (and, for that matter, teleoperated EOD activities) was put forward recently by Tabee *et al.*, (2009). To quote from their report:

"The major criticism and limitation of endoscopic surgery relates to the lack of depth perception of the 2D endoscopes. Depth perception is thought to be critical to precise motor movement. Two distinct aspects of the control of fine surgical movements have been described. The first involves initiation of a gross movement in the general desired direction. This is followed by multiple correctional movements that are modified based on a combination of visual cues. The number of required movements and accuracy of each movement are affected by the clarity of the visual feedback and experience of the surgeon. In endoscopic surgery, the lack of tactile cues and 2D visualization represent barriers to efficient and accurate movements. The acquisition of endoscopic skills inherently involves the ability to translate a 2D image into a mental 3D representation of a given area. This occurs partially through monocular cues including relative structure, size, texture gradients, linear perspectives along anatomic trajectories and motion parallax. Trained surgeons additionally learn to infer spatial relations from haptic cues and surgical movements. Despite these compensatory factors, 2D visualization does not match the depth perception gained by binocular cues including vergence, stereopsis, and vertical disparities.

6.3.2 Evidence For and Against – Experimental Studies

Munz *et al*. (2004) conducted an experiment to assess whether stereoscopic visualisation improves surgical performance on bench models using the *da Vinci* MIS robotic system. Four tasks were evaluated. The first task was a pick and place transfer task; the second a rope-passing task between left and right hand instruments; the third was a suturing task; and the fourth was similar to the first task (pick and place) but within a very limited space, requiring the movement of delicate foam balls (4mm in diameter). The tasks were attempted in two conditions using the *da Vinci* system – 2D viewing and in 3D stereoscopic mode. The results showed significant improvements for all tasks in the 3D condition, as measured by the time taken for task completion, the number of movements made, the total distance travelled and number of errors. The authors concluded that **in view of the lack of tactile feedback, robotic-assisted performance on bench models using 3D is more efficient using stereoscopic visualisation**.

Van Bergen *et al*. (1998) compared 2D and 3D vision systems for minimally invasive surgery. Using standardised surgical tasks, including suturing and knotting, surgeons involved in basic and advanced laparoscopic courses were assessed using 2D and 3D stereoscopic endoscopes (an example of the latter is shown in Figure 22). In addition, single-channel and bi-channel optics for stereoscopic endoscopes were compared. Single-channel optics consist of only one lens system, the image for which is split and then presented to the surgeon's left and right eyes (using, in this case, field-sequential shutter glasses) to generate the stereoscopic picture. In contrast, bi-channel optics consist of a two-lens system, each transferring one image.


Figure 22: Stereoscopic Endoscope
Source: his.anthropomatik.kit.edu

The results showed that **performance times were shorter, and fewer mistakes were made, in the 3D conditions**. In addition, all tasks were subjectively rated as easier under 3D. However, a number of subjective questions raised issues for the 3D systems. Rating on a five-point Likert scale ('no', 'very little', 'little', 'much', 'very much'), 17% of participants (out of 169) rated 'much' or 'very much' that **their perception of the operating environment was affected using the shutter glasses**. 36% (out of 167) rated 'much' or 'very much' when asked if they had become tired using the 3D system. 6% (out of 152) rated 'much' or 'very much' when asked if they had experienced a headache (30% rated at a "little"). A further 18% (out of 165) rated the reproduction of anatomic detail (i.e. the task detail) delivered by the 3D system to be 'bad' or 'very bad'. 36% (out of 159) rated the comfort of the shutter glasses as 'bad' or 'very bad' and 25% (out of 166) stated their impression of the brightness of the image as 'bad' or 'very bad'. Between the bi-channel and single- channel 3D systems, the authors found that for close-up work (1-3 cm) the single-channel optics were better and for distant work (4-6cm) the bi-channel system was better.

Falk *et al*. (2001) evaluated a series of tasks using the *da Vinci* MIS robotic system in three conditions: 2D, high-definition (HD) 2D and 3D. The results for a visual resolution task found that **the HD 2D system was associated with better resolution than the 2D, it also tended to be better than 3D, although not significantly**. For a distance evaluation task, there was no difference between the three viewing systems for determining the distance of pins relative to each other. **For a reach and touch task, 3D was associated with faster performance than either of the 2D viewing conditions. There were no significant difference between conditions for an object placement task in terms of speed and error, although the 3D conditions tended to be quicker. 3D was quicker than 2D for both a suturing task and a knot-tying task, but there was no difference with the addition of high definition (i.e. between 3D and HD 2D)**. The results show that 3D vision enhances surgical teleoperation performance for some surgically-related tasks, but that some of these improvements can also be made using high definition 2D.

Tabee *et al*. (2009) described a novel 3D stereo-endoscope for use in minimal access surgery. The system was used during neurosurgical interventions on 13 patients. The surgical outcomes were compared with a matched group who underwent similar procedures with 2D endoscopes. The results showed that, when using the 3D endoscope there were no interoperative complications. **Compared to 2D there was no significant difference in operation time or the extent of resection** (removal of organ or other significant body tissue). Subjectively, the surgeons who used the 3D endoscope reported improved depth perception. The study was the first to reported usage of a 3D endoscope for this type of surgery. The authors concluded that 3D endoscopes may become the standard tool for minimal access neurosurgery.

In current systems used for cardiac surgery, 3D volumetric data acquired from echocardiography scanners are projected on a conventional 2D displays where the depth of field is rendered by varying shades of grey. Vasilyev *et al*. (2008) compared this current method of display with one using stereoscopic images generated by a high- performance volume rendering software package. The task undertaken for this evaluation involved the surgical correction of an atrial defect within a pig's heart. The results showed that the **stereoscopic display system reduced surgical time and demonstrated greater navigational accuracy**. However, accuracy of actually placing the atrial anchors to correct the defect was no better than the normal 3D condition. The authors concluded that stereoscopic displays combined with 3D echocardiography ultrasound images may improve the safety of beating-heart intracardiac surgery. However, they highlighted that this experiment used a single participant highly experienced in endoscopic image-guided beating heart surgery, and that further studies should be carried out with individuals with various levels of surgical experience. The authors' conclusions should, however, be tempered by cautionary remarks made by Gronningsaeter *et al*. (2000) and Mueller-Richter *et al*. (2004), relating to the need to **ensure good image acquisition in order for 3D to be effective for surgical applications**. Specifically, Gronningsaeter *et al*. state that "stereoscopic display of ultrasound data is feasible when there is sufficient contrast between the objects of interest and the surrounding tissue". Mueller-Richter *et al*. (2004) stated that devices for 3D picture presentation/display are at a more advanced development than devices for 3D picture acquisition.

6.3.3 General Medical/Surgical Review Papers

Van Beurden *et al.* (2009) reviewed the use of stereoscopic displays for use in four distinct application areas in medical domains: diagnosis, pre-operative planning, minimally invasive surgery (MIS) and training/teaching. They found that, for diagnosis, **stereoscopic displays can augment the understanding of complex spatial structures and increase the detection of abnormalities**. **A stereoscopic presentation of noisy and transparent images when using 3D ultrasound results in better visualisation of the internal structures** (this confirms some of the binocular summation and noise reduction benefits of stereo, as listed earlier in this report). For MIS, stereoscopic displays can decrease surgery time and increase the accuracy of surgical procedures. They also suggest that **surgical procedures when using high resolution 2D displays generate similar levels of performance with lower resolution stereoscopic displays**. This suggests an image quality improvement for stereoscopic displays. They concluded that "... overall there is a clear need for more empirical evidence that quantifies the added value of stereoscopic displays in medical domains, such that the medical community will have ample basis to invest in stereoscopic displays in all or some of the described medical applications".

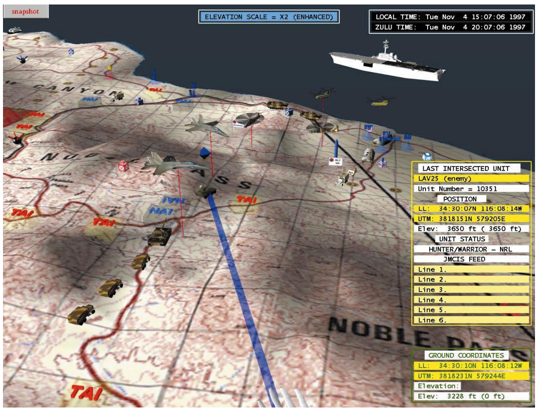
Hofmeister *et al.*, (2000) reviewed the use of 2D and stereoscopic display techniques in **endoscopic surgery**. Based on 15 studies, the authors concluded that **only about 50% found a significant benefit of stereoscopic systems**. They noted a number of experimental design problems with these studies, including conditions where the MIS camera is held stationary, therefore avoiding any complications with a moving camera, and the small sample sizes, as they often rely on surgeons with limited availability. Surgeon experience is also a key issue, and it may be more appropriate to involve medical students not previously exposed to either 2D or stereoscopic systems. The authors also noted some issues for improving performance in endoscopic surgery. They suggest that trainees may be tested for their ability to recover depth from pictures, which could be an important contribution to the selection of surgeons to use a 2D system, let alone a 3D system. The use of lighting was also highlighted, as shadows often play a critical role in determining shape and depth from a 2D scene. Ultimately, the authors state that "progress will result from a multidisciplinary approach, involving technological advance in the quality of the displayed image together with psychovisual motor and ergonomics research".

6.4 Command and Control

In many respects, the situation regarding 3D displays for Command and Control (C^2) applications parallels that for ATC/ATM, with significant cross-over between the papers uncovered for each domain. With C^2 applications, many papers focus on the visual characteristics of 3D content relevant to the tasks of detection, identification, perception of terrain topography and contouring, visualising trend data, and so on.

An early study investigating 3D representations for tactical displays by one of the authors (Stone, 1996) pointed out that "current [note date of study] communications and defence operations have led to increased information loads while reducing the time available for information processing and tactical decision-making. As the information increases and response time decreases there is a need to re-examine the nature of the information and its relationship to the display medium of choice being used". Stone's study posed the following questions:

- What does 3D (as illustrated in Figure 23 - a more recent battlefield visualisation concept called DRAGON, developed by the US Naval Research Laboratory (Durbin *et al.*, 1998)) offer over and above conventional 2D data display (e.g. as currently used for tactical plan position displays - track/threat/Identification, Friend or Foe (IFF)/response data, etc.)?



LOCAL TIME: Tue Nov 4 15:07:06 1997
ZULU TIME: Tue Nov 4 20:07:06 1997

LAST INTERSECTED UNIT
LAV25 (enemy)
Unit Number = 10351
POSITION
LL: 34:30:07N 116:08:14W
UTM: 3818151N 579205E
Elev: 3650 ft (3650 ft)
UNIT STATUS
HUNTER/WARRIOR – NRL
JMCIS FEED
Line 1.
Line 2.
Line 3.
Line 4.
Line 5.
Line 6.

GROUND COORDINATES
LL: 34:30:10N 116:08:12W
UTM: 3818231N 579244E
Elevation:
Elev: 3228 ft (0 ft)

Figure 23: *DRAGON* Battlefield Visualisation Test Bed
Source: Durbin *et al.*, 1998

The report concluded, from the limited evidence at the time, that **3D offered a more intuitive interface from the point of view of situational awareness and navigational strategies**.

- Is there any (preferably experimental or field) evidence to support a "move" by the military from 2D to 3D data display?

Whilst there were a number of programmes under way at the time the report was compiled that demonstrated a military focus to the use of 3D displays, very few (if any – at the time of writing - 1996) had been tested in a field setting or even within a simulated military environment. There was evidence that relates the use of 3D to improved recall and information uptake, although much of this evidence had been generated using quite simplistic experimental comparisons between textual and basic graphical data representations. There were some exceptions. Ware & Franck (1994a;b), for example, conducted experimental studies that demonstrated an **improvement in the understanding of an abstract information net by a factor of 1.6 when 3D was employed, and up to a factor of 3 when head coupling was employed**. They suggest that **motion parallax cues are of greater importance than stereopsis in revealing structural information**. Koike (1993) showed how the introduction of visual 3D frameworks could **dramatically reduce users' cognitive loading when forming mental correlations between multiple window environments** (previously displayed in 2D).

Also of relevance to the present study, Wickens *et al.* (1994) showed that **3D displays using monocular cues produced shorter response times than 2D displays**. However, Hollands *et al.* (1995) provided experimental evidence to suggest that **2D displays generally elicited better trend and difference estimation performance than 3D displays**. The study also found that information displayed in 3D can, as far as the user's performance is concerned, be supported or confounded by the integration of other, essentially monocular or dynamic cues. For example, Ware & Franck (1994a) discovered that **motion cues in information display are more significant than stereo cues, irrespective of the type of motion**. With regard to other supplementary cues Wickens *et* *al*. (1994) and Merwin *et al*. (1995) demonstrated that **colour coding did nothing to help users' understanding of 3D displays**. Sound also appears to play a reasonably important role in enhancing situational awareness, reducing loading on visually dominated sensory systems (eg. Venolia, 1993), with 3D audio improving target location and identification (Perrott *et al*., 1995). 3D audio appears to act in a similar fashion to abrupt visual onsets, but with an attention-capturing capability at much longer distances (Strybel *et al*., 1995). In a general study of audio feedback in conjunction with virtual displays, McKinley *et al*. (1995) found, using both subjective and objective measures, a positive effect of sound on 3D localisation.

- Are there any accepted or emerging human factors *principles* for effective 3D data design?

Apart from one specific public domain summary publication by Wickens *et al*. (1994), with some reference to 3D in the likes of Boff & Lincoln (1988) and Farrell & Booth (1994), there had, at that time, been no serious attempt to develop a methodology, nor collate a series of usable guidelines for the implementation of 3D visualisation systems. The Stone (1996) study went on to propose a possible methodology, applying techniques used in the then-active scientific visualisation community (e.g. Keller & Keller, 1993) to develop a concept 3D display for the Centre for Human Science's Air Defence Scenario for UK STOW '97 (Figure 24).

Since 1996, of course, there have been rapid and extensive developments in the hardware and software tools supporting interactive 3D (i3D) and Virtual Environment (VE) developments for C^2 and tactical information visualisation. However, one of the main criticisms raised in the Stone (1996) study still remains. There is not one single publication to which interface designers can turn for clear guidance on the design and development of 3D displays for a broad range of applications, such as those covered herein, be that relating to interactive 3D content or to the choice of appropriate display hardware. The recent documents published as part of the HFI DTC programme (Stone 2008; 2011), providing guidelines for i3D relating to games-based training systems design, represent but one part of this complex and rapidly growing arena.

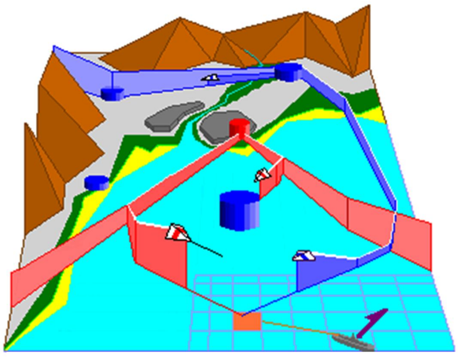


Figure 24: 3D Tactical Display Concept
based on STOW '97
Source: Author's Archives

More recent developments relating to the exploitation of i3D in C^2 applications include the research by St. John *et al*. (2001b). Their study examined spatial judgement and shape understanding with 2D and 3D displays in an experiment that required participants to locate objects within a terrain map (rendered using the 3ds max 3D modelling and animation toolkit) to judge the visibility between objects and from other locations on the terrain. This task is analogous to the task of locating communication antenna or surface-to-air missile site location, ensuring clear lines of sight between the units, yet remaining hidden from the view of enemy patrols. The researchers' results showed that the **3D view produced slower solution times than the 2D view**. St. John *et al*. suggest that "the accurate representations of distances and elevations inherent to the 2D view were more helpful to the precise understanding of the terrain than the more realistic 3D view that inevitably invited line-of-sight ambiguity costs". In contrast, when participants were asked to choose a promising route through the terrain representation, the **3D view enabled participants to form a general understanding of the terrain more quickly than the 2D view**. The authors suggest that these findings support a display design paradigm called "Orient and Operate" in which **3D views would be best for conveying general shape understanding and then 2D views would be best used for solving tactical problems involving precise 3D judgments**. To quote from St. John *et al.* again:

"... **two dimensional views ... are useful for judging relative positions because the normal viewing angles (e.g., top-down, side, front) minimize distortion** ... Ambiguity is confined to a single dimension such as altitude in a top-down view. This confinement of ambiguity to the dimension that is not represented provides better opportunities to deal with the ambiguity. For instance, a user can easily switch among a set of 2D views to obtain exact information about each dimension of interest. In contrast, each dimension of a 3D view is confounded with ambiguity spread across all three dimensions. This ambiguity and distortion make relative-position judgments of any precision difficult"

Continuing on the theme of terrain understanding, Hollands *et al.* (2011) investigated the advantage of visual momentum in the form of smooth transition between 2D and 3D displays of geographical terrain (a scenario that is relevant to future multi-screen C^2 applications). A task was designed where participants were required to determine the relative heights of two points on a natural terrain image and whether one point could be seen from the other. In making these judgements the scene changed from one display format to the other (i.e. between 2D and 3D). A number of transition formats were compared: continuous (a dynamic rotation from one display format to the other), discrete - immediate, discrete - delayed, and discrete - delayed with a preview. The results showed that performance after continuous transition was superior than after discrete transition. The authors concluded that **dynamic transitions from 2D to 3D are recommended when observers examine multiple views of terrain over time**.

Peinsipp-Byma *et al.* (2009) compared monoscopic 2D with four types of stereoscopic displays (anaglyph, polarised, field sequential and autostereo) in a task where participant had to detect objects (e.g. antennas, poles, vehicles, airports, military objects and high buildings) in aerial pictures. No statistical analysis was carried out on the data, but mean values suggest that **performance using the stereo systems was better than the monoscopic system in terms of higher detection rate and lower false detections, although mean working time was lowest for the monoscopic condition**. **Subjective ratings of workload indicated that the participants felt greater physical strain in the stereoscopic conditions, particularly the autostereoscopic condition, where the observer had to maintain a defined position to generate the 3D image**. Ratings of visual strain in the stereoscopic conditions were also taken, more so the autostereoscopic condition which also rated highest for frustration. Between the stereoscopic conditions, polarised and field sequential techniques performed best. These techniques were also rated as most preferred by the users. However, as stated above, no statistical analysis was carried out on the performance data. Indeed, the large standard deviations shown with the mean data, suggest that many of the results may not reach statistical significance at the 0.05 level, using parametric statistical techniques.

6.4.1 Submarine "Traffic"

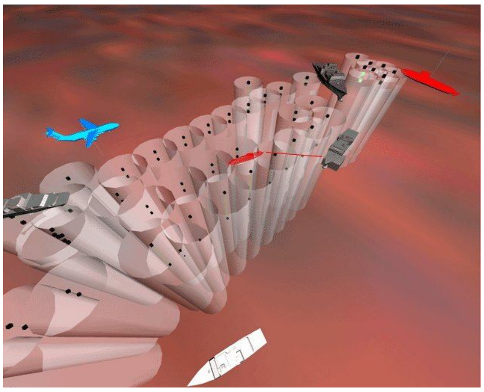


Figure 25: An Early 3D Concept Display for Surface and Submarine Threat Detection (Source: Author's Archive)

In a study related to some of the ATC scenarios addressing identification and collision incident detection/avoidance described above, Durkee *et al*. (2010) evaluated the use of 2D and autostereoscopic 3D displays (which they describe as "low-fidelity" and "high-fidelity") for submarine navigation (an early concept for which is shown in Figure 25). The high-fidelity autostereoscopic display was, in fact, a relatively high-resolution, eye-tracking-enabled device. The researchers discovered that **the 2D display produced highest overall performance, in terms of collision detection and new object identification**. **Response times were also quicker, confidence ratings were higher and fatigue lower in the 2D display condition**. However, workload was also rated higher in the 2D condition and route selection to avoid collision faster in 3D. Durkee *et al*. concluded that there was too much uncertainty and risk to expect human performance advantages for autostereoscopic 3D.

7 Conclusions

The conclusions and recommendations presented here have been collated both from the literature survey undertaken in support of the present HFI DTC study and from practical experience of evaluating and deploying 3D/stereoscopic systems in the laboratory and within third-party facilities.

Tam *et al*. (2011) state that "... the great interest for 3D-TV stems from the recognition that, when compared to standard two-dimensional (2D) television, this technology significantly enhances the entertainment value of television programs". They go on to claim that "surveys indicate that people would rather view S3D images than their two-dimensional counterparts, provided that the stereoscopic images are free from annoying artefacts and are comfortable to view".

History has shown how human interface technologies, with many recent examples having been developed by the entertainment industry, be they early market prototypes or well-established (but unproven from a cost-benefit perspective), quite regularly find their way into mainstream, real-world applications with very little (if any) early attention being paid to the capabilities, limitations and learning requirements of the end user populations. Virtual/Augmented Reality and "Serious Games" are two excellent and relatively recent examples, and the early adoption problems witnessed in these arenas over the past two decades seem to be regularly repeated on a 5-year rolling basis. Similar evidence is, at the time of writing, forming for related interactive technologies, such as portable games consoles, computer tablets, haptic and olfactory "displays", brain-computer interfaces, and so on.

Often, such early, indeed *premature* adoption may be symptomatic of the need to spend surplus end-of-year budgets to prevent losses in the subsequent financial accounting period. Or it may be a case of "if the US has it, so should we". This is an issue that is not only restricted to industrial, commercial or governmental adopters, but to academia as well – the spread of so-called "VR Centres of Excellence" since the mid-1990s helped to preserve the myth that interactive technologies were, perhaps, more mature than was actually the case. Yet the research outputs from these centres did nothing to help real-world adopters avoid making wasteful procurements of unreliable and unusable technologies.

Despite the unique developmental history surrounding the development of 3D displays and viewing technologies, such technology also now falls into this "premature adoption risk" category, courtesy of very recent developments in the commercial exploitation of stereoscopic viewing for cinematic, home and gaming productions. Of equal concern is the fact that stereoscopic and 3D viewing technologies have been under investigation by non-entertainment-based organisation for nearly five decades, yet their widespread adoption has still to be witnessed. As pointed out by one of the authors over 15 years ago, "Unfortunately, the true importance of, and benefits to be gained from the use of 3D in the display of information to human users has never been proven beyond doubt. In the academic and technical literature published since the late 1970s, there have been regular "surges" of interest in 3D and stereo, occurring almost "sinusoidally" on a 2-yearly basis. Initial studies of 3D and true stereoscopic displays (i.e. those which exploit the binocular vision, or stereoptic/stereopsis characteristics of the human viewer) concentrated on video systems, relaying information from teleoperated vehicles stationed at remote and hazardous worksites, (e.g. within nuclear hot cells or subsea). A range of prototypes were constructed, but a good number of these failed to gain full operator acceptance and, therefore, never attained operational status" (Stone, 1996).

Very little seems to have changed today. As has been shown in this report, there are certainly a number of studies that suggest that the use of 3D displays may be associated with enhanced end user performance (or perceived improvements) in very specific task elements, be they represented in a computer-generated fashion (as for ATC and C^2), or generated from remote hazardous or safety-critical environments (as with EOD telerobots and surgery). However, there are also many conflicting papers and the "disconnect" between laboratory studies and real-world domains and experiences seems as evident today as it was in the 1970s and 1980s.

3D displays form one of an ever-increasing range of technology categories that have been described as "non-traditional" human interfaces (Kortum, 2008; see also Stone, 2011). This term was coined to cover a variety of human-computer interface technologies, in particular those that attempt to "satisfy" most, if not *all* of the sensory and motor qualities humans are born with.

> **"Many of these interfaces will evoke a strong "wow" factor ... since they are very rare, and commercial applications are not generally available. Others ... may not seem as exciting, but they are incredibly important because they are widely deployed, and generally very poorly designed".**

(Kortum, 2008; page 1).

8 Recommendations

It is highly recommended that the “temptation” to procure more “exotic” forms of 3D display, such as volumetric or holographic projection technologies, should be avoided for defence applications. However, this category of display should be the focus of annual technology watches, in order to monitor progress and evolution of specific systems into acceptable technology readiness levels (TRLs), suitable for task-specific investigations. Stereoscopic imagery presentation using head-mounted or head-coupled displays forms a special case for early Human Factors treatment, particularly if the imagery being viewed by a single user is presented to other viewers – stereoscopically or not. However, given the current state of development with COTS HMDs, for most tasks their adoption should be discouraged and alternative display techniques sought (see also Stone, 2011).

Quoting from Meesters *et al*. (2004):

- “Stereoscopic TV should be able to provide good quality stereo pictures to multiple viewers who are free to move throughout the room”,
- “Any stereoscopic system should also be able to display monoscopic images without problems, and with an image quality that is at least comparable, but preferably superior to current TV”, and, perhaps most importantly of all,

Of direct relevance to the outcome of the present study, Meesters *et al.* (2004) also conclude:

- “Prototype systems need to be tested outside the controlled laboratory space, for it is the long-term, real-world use of 3-D TV that will prove its impact”.

Therefore, and regardless of the application domain, ANY opportunity to deploy 3D or stereoscopic display systems MUST be preceded by a Human Factors study of the tasks required of the end users and the context in which such technology might be deployed. The results of the study should be presented concisely, in order to help the end user community decide:

(a) whether or not the use of 3D or stereoscopic technologies and/or content are actually justified (from a task enhancement perspective), or are other forms of data/image representation more appropriate?

(b) if there *is* justification from a Human Factors standpoint, then what type of 3D display technology best serves the end users’ needs for the tasks they are expected to perform?

(c) what are the key technology deployment environment issues (e.g. from open spaces to dedicated command centres, particularly)?

(d) what are the key interactive issues (not only with the main display and the 3D data being displayed, but with other human-system interfaces that are likely to be in close proximity as well)?

(e) what are the key health and safety issues (including acceptable viewing durations, the screening of potential end users to exclude those unable to perceive 3D, the need for regular eye testing of actual end users, etc.)?

To support such studies, it is recommended that further research should be undertaken to build on the findings of this report, the literature survey undertaken to date and the contents of Stone (1998, 2008 and 2011) to develop a Human Factors methodology to support:

- The analysis of tasks best suited to implementation in 3D synthetic or stereoscopic form,
- The design of appropriate 3D content and fidelity – be that symbolic, representative of the real world, or a combination of the two,
- The selection of appropriate hardware for 3D stereoscopic data display,
- The analysis of end user contexts and environments into which a 3D or stereoscopic display is intended to be introduced,
- The evaluation of 3D and stereoscopic displays, including the development of an appropriate set of subjective and objective metrics.

In parallel with the development of this methodology, a minimum of two real-world case studies need to be undertaken (with experimental evaluations), based on tasks identified in current UK defence stakeholder projects, to help illustrate and justify the outcomes of the methodology. Where the case studies focus on stereoscopic display technology, then the results of the methodology should be compared across different types of displays, in conjunction with task elements displayed on other non-stereoscopic devices and in different ambient environments.

Further investigations also need to be conducted relating to appropriate interactive devices and 3D display technologies, for both individual and multi-user interaction. This is not only an interactive hardware ergonomics issue, but also concerns how best to represent the virtual interactive elements within 3D graphical displays, based on the tasks required (i.e. navigation, object contact, select and interrogation, object placement, data “drill-down”) and how interaction may be supported or confounded when presented using specific types of 3D or stereoscopic displays.

Finally, a study should also be undertaken to gain some idea as to the likely percentage of military personnel who suffer from stereoscopic defects and the severity of said defects. Such a study should be conducted in a way so as not be perceived as a form of selection test (unless, of course, the individuals under scrutiny are destined for a role that demands very high levels of stereoacuity). In particular, participants from groups within the Armed Forces that may be expected to consider the adoption and deployment of stereoscopic technology in the future should be tested, including those involved in explosive ordnance disposal (manual and telerobotic), unmanned vehicle operators (land, sea and air), C^2 activities (for example operations and control room personnel), medics, helicopter support personnel (e.g. rear-door winch operators or voice marshals), close-range weapons duties, and so on.

The development of a computer-based stereoscopic 3D vision capability testing system, appropriate to the Armed Forces, may also be worthy of study, design and early evaluation. The literature review supporting in this study has certainly highlighted a range of tasks that could be developed further and committed to 3D software programs. The head-coupled perspective concept developed for the iPod and iPad, mentioned earlier in this report, also shows that such a series of tests could be made readily available for portable computing platforms.

References

Alexander, A.L., & Wickens, C.D. (2005). “3D Navigation and Integrated Hazard Display in Advanced Avionics: Performance, Situation Awareness, and Workload”; Technical Report AHFD-05-10/NASA-05-2. Prepared for NASA Langley Research Center Hampton, VA.

Balakrishnan, R., Fitzmaurice, G.W., & Kurtenbach, G. (2001). “User Interfaces for Volumetric Displays”; *Computer (Next Generation Interfaces* Special Edition*)*; March 2001; pp. 37-45.

Basdogan, C., Lum, M., Salcedo, J., Chow, E., Kupiec, S.A., & Kostrewski, A. (2002). “Autostereoscopic and Haptic Visualization for Space Exploration and Mission Design”; In *Proceedings of HAPTICS '02* - the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems; Orlando, Florida; March 24-25, 2002; IEEE.

Berry, H.W. (1979). “An Underwater Three-Dimensional (3D) Television System”; In *Proceedings of the 1979 Offshore Technology Conference*; 30 April to 07 May; Houston, Texas.

Blake, R., & Fox, R. (1973). “The psychophysical inquiry into binocular summation”; *Perception & Psychophysics*; 14(1): pp. 161–85.

Boff, K.R., & Lincoln, J.E. (1988). “Engineering Data Compendium: Human Perception and Performance”; H.G. Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio. Integrated Perceptual Information for Designers Program; 3 volumes plus User's Guide.

Bourgois, M., Cooper, M., Duong, V., Hjalmarsson, J., Lange, M., & Ynnerman, A. (2005). “Interactive and Immersive 3D Visualization for ATC”; In *Proceedings of USA/Europe R&D Seminar ATM 2005*; Baltimore.

Brown, M.A., & Slater, M. (1997). “Some Experiences with Three-Dimensional Displays Design: An Air Traffic Control Visualisation”; In *Proceedings of the IEEE International Workshop on Robot and Human Communication*; pp.296-301.

Braunstein, M.L. (1976). “Depth Perception Through Motion”; New York Academic.

Campbell, F.W., & Green, D.G. (1965). “Optical and Retinal Factors Affecting Visual Resolution”; *J. Physiol (Lond.)*; 181; pp.576-593.

Chen, J.Y.C., Hass, E.C., & Barnes, M.J. (2007). “Human Performance Issues and User Interface Design for Teleoperated Robots”; *IEEE Transactions on Systems, Man, and Cybernetics* – Part C: Applications and Reviews; 37(6); pp.1231-1245.

Chen, J.Y.C., Haas, E.C., Pillalamarri, K., & Jacobson, C.N. (2006). "Human-Robot Interface: Issues in Operation Performance, Interface Design, and Technologies"; Report ARL-TR-3834. U.S. Army Research Laboratory, Aberdeen Proving Ground, MD 21005-5425.

Chen, J.Y., Oden, R.V.N., Drexler, J.M., & Merritt, J.O. (2010a). "Evaluation of Stereoscopic Display for Indirect-Vision Driving and Robot Teleoperation"; *The WSTIAC Quarterly*; 10(2); pp.3-9.

Chen, J.Y., Oden, R.V.N., Kenny, C., & Merritt, J.O. (2010b). "Effectiveness of Stereoscopic Displays for Indirect-Vision Driving and Robot Teleoperation"; Report ARL-TR-5259. U.S. Army Research Laboratory, Aberdeen Proving Ground, MD.

Crescenzio, F.D., Miranda, G., Persiani, F., & Bombardi, T. (2009). "A First Implementation of an Advanced 3D Interface to Control and Supervise UAV (Uninhabited Aerial Vehicles) Missions"; *Presence*; 18(3); pp.171-184.

Draper, J.V., Handel, S., Hood, Kring, C.T. (1991). "Three Experiments with Stereoscopic Television: When it Works and Why"; In *Proceedings of IEEE International Conferences on Systems, Man and Cybernetics, 2*; pp.1047-1052.

Drascic, D., & Grodski, J.J. (1993). "Using Stereoscopic Video for Defence Teleoperation"; In *Proceedings of the SPIE Conference on Stereoscopic Displays and Applications IV*; San Jose; 1993.

Dumbreck, A.A., Scheiwiller, P.M., Murphy, S.P., Kitmer, C.A., Bentz, M.A. (1991). "3-D TV Systems for Remote Operations: Development and Applications"; In *Proceedings of American Nuclear Society 4th Topical Meeting on Robotics and Remote Systems*, Albuquerque; February, 1991.

Durbin, J., Swan II, J.E., Colbert, B., Crowe, J., King, R., King, T., Scannell, C., Wartell, Z., & Welsh, T. (1998). "Battlefield Visualization on the Responsive Workbench"; In *Proceedings of the IEEE Visualization '98*; IEEE Computer Society Press; pp. 463-466.

Durkee, K., Paley, M., Kudzma, K., & Reiners, D. (2010). "Evaluation of 2D and 3D Displays to Validate the Integration of Human Performance Analysis Methods into Submarine Acquisitions"; In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*; pp.2196-2200.

Falk, V., Mintz, D., Grunenfelder, J., Fann, J.I., & Burdon, T.A. (2001). "Influence of Three-Dimensional Vision on Surgical Telemanipulator Performance"; *Surgical Endoscopy*; 15; pp. 1282-1288.

Farrell, R.J., & Booth, J.M. (1984), "Handbook for Imagery Interpretation Equipment"; Boeing Aerospace Company (Seattle, Washington 98124), Report No. D 180-19063-1.

Farvardin, M., & Afarid, M. (2007). "Evaluation of Stereo Tests for Screening of Amblyopia"; *Iranian Red Crescent Medical Journal*; 9(2); pp. 80-85.

Fricke, T.R., & Siderov, J. (1997). "Stereopsis, Stereotests and their Relation to Vision Screening and Clinical Practice"; Clin. Exp. Optom*.;* 80; pp. 165-172.

García-Robledo, P., García-Borrás, P., Barrio, J., Ferre, M., & Aracil, R. (2009). "Teleoperation Platform for Multimodal Experimentation: Common Communications Applied to Teleoperation"; In *Proceedings of the 2009 IEEE International Conference on Mechatronics*; Malaga, Spain; April 2009.

Grossman, T., & Balakrishnan, R. (2006). "An Evaluation of Depth Perception on Volumetric Displays"; In *Proceedings of the Working Conference on Advanced Visual Interfaces. AVI 06*; May 23–26, 2006; Venezia, Italy; ACM Press; pp. 193-200.

Grossman, T., & Balakrishnan, R. (2008). "Collaborative Interaction with Volumetric Displays"; In *Proceedings of CHI 2008 – "Collaborative User Interfaces*; April 5-10, 2008 · Florence, Italy.

Harvey, D. (1987). "VCASS: A Second Look at the Super Cockpit"; *Rotor & Wing International*; 21; pp.32-33 and 63.

Hatch, S.W., & Richman, J.E. (1994). "Stereopsis Testing Without Polarized Glasses: a Comparison Study on Five New Stereoacuity Tests"; *J. Am. Optom. Assoc.*; 65(9): pp.637-41.

Hofmeister, J., Frank, T.G., Cuschieri, A., & Wade, N.J. (2001). "Perceptual Aspects of Two-Dimensional and Stereoscopic Display Techniques in Endoscopic Surgery: Review and Current Problems"; *Seminars in Laparoscopic Surgery*; 8(1); pp.12-24.

Hofstetter, H.W., & Bertsch, J.D. (1976). "Does Stereopsis Change With Age?"; *Am. J. Optom. Physiol. Optics*; **53**: pp.664-7.

Hollands, J.G., Pavlovic, N.J., Enomoto, Y., & Jiang, H. (2008). "Smooth Rotation of 2-D and 3- Representations of Terrain: An Investigation into the Utility of Visual Momentum"; *Human Factors*; 50(1); pp.62-76.

Hollands, J.B., Pierce, B., & Magee, L.E. (1995), "Displaying Quantitative Information in Two and Three Dimensions"; In *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*; Volume 2; pp.1425-1429.

Hubona, G.S., Wheeler, P.N., Shirah. G.W., & Brandt, M. (1999). "The Relative Contributions of Stereo, Lighting and Background Scenes in Promoting 3D Depth Visualization"; In *ACM Transactions on Computer-Human Interaction (TOCHI)*; Volume 6 Issue 3; September 1999; pp. 214–242.

Julesz, B. (1971). "Foundations of Cyclopean Perception"; University of Chicago Press.

Kanduri, A.K., Thomas, G., Cabrol, N., Grin, E., & Anderson, R.C. (2005). “The (In)Accuracy of Novice Rover Operators’ Perception of Obstacle Height From Monoscopic Images”; *IEEE Transactions on Systems, Man and Cybernetics* – Part A: Systems and Humans; 35(4); pp.505-512.

Keller, P., & Keller, M. (1993). “Visual Cues - Practical Data Visualisation”; IEEE Computer Society Press; 2nd Edition.

Koike, H. (1993). “The Role of Another Spatial Dimension in Software Visualisation”; *ACM Transactions on Information Systems*; 11(3); pp.266-286.

Kolasinski, E.M. (1995). “Simulator Sickness in Virtual Environments”; US Army Research Institute Technical Report No. 1027 (Project No. 20262785A791).

Kortum, P. (Ed., 2008). “HCI Beyond the GUI: Design for Haptic, Speech, Olfactory and Other Nontraditional Interfaces”; Morgan Kaufmann Series in Interactive Technologies. Elsevier Inc.

Lange, M., Hjalmarsson, J., Cooper, M., Ynnerman, A., & Duong, V. (2003). “3D Visualization and 3D Voice Interaction in Air Traffic Control”; In *Proceedings of the Annual SIGRAD Conference* - Special Theme: Real Time Simulations; pp.17-22.

Lee, S., & Kim, G.J. (2008). “Effects of Haptic Feedback, Stereoscopy, and Image Resolution on Performance and Presence in Remote Navigation”; *International Journal of Human-Computer Studies*; 66(10); pp.701-717.

Lim, K.Y., & Quek, R.S.M. (2002). “Enhancing Operator Performance of Remote Container Landing: An Assessment of a 3D Stereoscopic Control & Display System”; In *Proceedings of HF 2002*; November 25-27, Melbourne, Australia.

May, J. (2000). “Perceptual Principles and Computer Graphics”; *Computer Graphics Forum*; 19(4); pp. 271-279.

McKinley, R.L., D'Angelo, W.R., Haas, M.W., Perrott, D.R., Nelson, W.T., Hettinger, W.T., & Brickman, B.J. (1995). “An Initial Study of the Effects of 3-Dimensional Auditory Cueing on Visual Target Detection”; In *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*; Volume 1; pp.119-123.

Merwin, D.H., Vincow, M.A., & Wickens, C.D. (1994). “Visual Analysis of Scientific Data: Comparison of 3D Topographic, Colour and Gray Scale Displays in a Feature Detection Task”; In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*; 38 (4); pp**. 240-244.**

Montes-Mico, R. (2001). “Prevalence of General Dysfunctions in Binocular Vision”; *Annals of Ophthalmology*; 33(3); pp. 205-208.

Mueller-Richter, U.D.A., Limberger, A., Weber, P., Ruprecht, K.W., Spitzer, W., & Schilling, M. (2004). "Possibilities and Limitations of Current Stereo-Endoscopy"; *Surgical Endoscopy*; 18; pp.942-947.

Munz, Y., Moorthy, K., Dosis, A., Hernandez, J.D., Bann,S., Bello, F., Martin, S., Darzi., & Rockall, D.T. (2004). "The Benefits of Stereoscopic Vision in Robotic-Assisted Performance on Bench Models"; *Surgical Endoscopy*; 18; pp.611-616.

Ohlsson, J., Villareal, G., Abrahamsson, M., Cavazos H, Sjöström A., & Sjöstrand, J. (2001). "Screening Merits of the Lang II, Frisby, Randot, Titmus, and TNO Stereo Tests"; *J.AAPOS* (Journal of the American Association for Pediatric Ophthalmology and Strabismus); 5; pp.316-22.

Peinsipp-Byma, E., Rehfeld, N., & Eck, R. (2009). "Evaluation of Stereoscopic 3D Displays for Image Analysis Tasks"; In Woods, A.J., *et al.*, (Eds.), *Stereoscopic Displays and Applications XX*. Proceedings of SPIE-IS&T Electronic Imaging; SPIE; Vol. 7237.

Paulos, E., & Canny, J. (1997). "Ubiquitous Tele-Embodiment: Applications and Implications"; *International Journal of Human–Computer Studies*; 46(6); pp.861–877.

Park, S.H., & Wolstad, J.C. (2000). "Multiple Two-Dimensional Displays as an Alternative to Three-Dimensional Displays in Telerobotic Tasks"; *Human Factors*; 42(4); pp.592-603.

Perrott, D.R., Cisneros, J., McKinley, R.L., & D'Angelo, W.R. (1995), "Aurally Aided Detection and Identification of Visual Targets"; In *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*; Volume 1; pp.104-108.

Roscoe, S.N. (1968). "Airborne Displays for Flight and Navigation"; *Human Factors*; 10(4); pp.321-332.

Rozzi, S., Amaldi, P., Wong, W., & Field, B. (2007). "Operational Potential for 3D Displays in Air Traffic Control"; In *Proceedings of the ECCE 2007 Conference*; 28-31 August 2007; London, UK; pp.170-183.

Scribner, D.R., & Gombash, J.W. (1998). "The Effect of Stereoscopic and Wider Field of View Conditions on Teleoperator Performance"; Report ARL-TR-1598. US Army Research Laboratory, Aberdeen Proving Ground, MD 21005-5425.

Sheridan, T.B. (1987). "Telerobotics"; Plenary Presentation for 10th IFAC World Congress on Automatic Control; Munich; July, 1987.

Sheridan, T.B. (1992a). "Telerobotics, Automation and Human Supervisory Control"; MIT Press.

Sheridan, T.B. (1992b). "Musings on Telepresence and Virtual Presence"; *Presence: Teleoperators and Virtual Environments*; 1(1); pp.120–126.

Smallman, H.S., St. John, M., Oonk, H.M., & Cowen, M.B. (2001). "Information Availability in 2D and 3D Displays"; *IEEE Computer Graphics and Applications*. September/October; pp.51-57.

St John, M., Cowen, M.B., Smallman, H.S., & Oonk, H.M. (2001a). "The Use of 2D and 3D Displays for Shape-Understanding Versus Relative-Position Tasks"; *Human Factors*; 43(1); pp.79-98.

St. John, M., Smallman, H.S., Bank, T.E., & Cowen, M.B. (2001b). "Tactical Routing Using Two-Dimensional and Three-Dimensional Views of Terrain"; Technical Report 1849. United States Navy. SSC San Diego, San Diego, CA.

Stone, R.J. (1983). "Human Factors and the Design of Remotely Operated Submersible Systems (ROVs)"; *International Underwater Systems Design*; 5(1).

Stone, R.J. (1996). "3D Representations for Tactical Displays"; Final Contract Report delivered to Gregory Harland Ltd under DERA Centre for Human Sciences Contract reference CHS0/5002S.

Stone, R.J. (2008). "Human Factors Guidelines for Interactive 3D and Games-Based Training Systems Design: Edition 1"; Human Factors Integration Defence Technology Centre Publication; August, 2008.

Stone, R.J. (2011). "Human Factors Guidelines for Interactive 3D and Games-Based Training Systems Design: Edition 2"; Human Factors Integration Defence Technology Centre Publication; September, 2011.

Stone, R.J., & Mason, C.A. (1986). "AERE Contract H2C 547530 ES: Project Definition Study: Human Factors Evaluation of the MA-23M Servomanipulator"; BAe Sowerby Research Centre Report No. JS 10561; March, 1986.

Strybel, T.Z., Boucher, J.M., Fujawa, G.E., & Volp, C.S. (1995). "Auditory Spatial Cueing in Visual Search Tasks: Effects of Amplitude, Contrast and Duration"; In *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*; Volume 1; pp. 109-113.

Sutherland, I. (1965). "The Ultimate Display"; In *Proceedings of the* International Federation for Information Processing *(IFIP) '65*; pp.506-508.

Thomas, G., Goldberg, J.H., Cannon, D.J., & Hillis, S.L. (2002). "Surface Textures Improve the Robustness of Stereoscopic Depth Cues"; *Human Factors*; 44(1); pp.157-170.

Tsirlin, I., Wilcox, L.M., & Allison, R.S. (2011). "The Effect of Crosstalk on the Perceived Depth from Disparity and Monocular Occlusions"; *IEEE Transactions on Broadcasting*; 57(2); pp.445-453.

Van Bergen, P., Kunert, W., Bessell, J., & Buess, G.F. (1998). "Comparative Study of Two-Dimensional and Three-Dimensional Vision Systems for Minimally Invasive Surgery"; *Surgical Endoscopy*; 12; pp.948-954.

Van Beurden, M.H.P.H., Van Hoey, G., Hatzakis, H., & Ijsselsteijn, W.A. (2009). "Stereoscopic Displays in Medical Domains: A Review of Perception and Performance Effects"; *SPIE*; Vol. 7240.

Van Orden, K.F., & Broyles, J.W. (2000). "Visuospatial Task Performance as a Function of Two- and Three-Dimensional Displays Presentation Techniques"; *Displays*; 12; pp.17-24.

Vasilyev, N.V., Novotny, P.M., Martinez, J.F., Loyola, H., Saigo, I.S., Howe, R.D., & del Nido, P.J. (2008). "Stereoscopic Vision Display Technology in Real-Time Three-Dimensional Echocardiography-Guided Intracardiac Beating-Heart Surgery"; *Journal of Thoracic and Cardiovascular Surgery*; 135(6); pp.1334-1341.

Venolia, D. (1993), "Facile 3D Direct Manipulation"; In *INTERCHI '93; Proceedings of the Conference on Human Factors in Computing Systems*; pp.31-36.

Vertut, J., & Coiffet, P. (1986). "Teleoperation and Robotics Evolution and Development"; Kogan Page.

Ware, C., & Franck, G. (1994a). "Evaluating Stereo and Motion Cues for Visualising Information Nets in Three Dimensions"; University of Brunswick Technical Document No. TR94-082.

Ware, C., & Franck, G. (1994b). "Viewing a Graph in a Virtual Reality Display is Three Times as Good as a 2D Diagram"; In *Proceedings of the 1994 IEEE Conference on Visual Languages*; pp.182-183.

Westwood, J.D. (1981). "The Application of Television to Remotely Operated Underwater Vehicles"; *J. Phys. E: Sci. Instrum.*; 14; pp.137-144.

Wickens, C.D. (1995). "Display Integration of Air Traffic Control Information: 3D Displays and Proximity Compatibility"; Final Technical Report AL-95-2/FAA-95-2. Prepared for Federal Aviation Administration System Operations & Engineering Branch Washington, DC.

Wickens, C., Campbell, C., Liang, C., & Merwin, D. (1995). "Weather Displays for Air Traffic Control: The Effect of 3D Perspective"; Aviation Research Laboratory Technical Report: ARL-95-1/FAA-95-1, 1995.

Wickens, C.D., Todd, S., & Seidler, K. (1989). "Three-Dimensional Displays: Perception, Implementations and Applications"; CSERIAC Tech. Rep 89-001, Wright Patterson Air Force Base, Ohio.

Wickens, C.D., Merwin, D.H., & Lin, E.L. (1994). "Implications of Graphics Enhancements for the Visualisation of Scientific Data"; *Human Factors*; 36(1); pp.44-61.

Wickens, C.D., Laclair, M., & Sarno, K. (1995). "Graph-Task Dependencies in Three-Dimensional Data: Influence of Three-Dimensionality and Color"; In *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*; Volume 2; pp.1420-1424.

Wittlief, K., (2007). "Stereoscopic 3D Film and Animation – Getting it Right"; *Computer Graphics*; ACM SIGGRAPH Publication; 41(3). Available at:http://www.siggraph.org/publications/newsletter/volume/stereoscopic-3d-film-and-animationgetting-it-right (last visited, 03/01/2012)

Woods, A., Docherty, T., & Koch, R. (1994a). "Experiences of Using Stereoscopic Video with an Underwater Remotely Operated Vehicle"; In *Proceedings of Underwater Intervention 1994*; San Diego, California; pp.34-38.

Woods, A., Docherty, T., & Koch, R. (1994b). "Field Trials of Stereoscopic Video with an Underwater Remotely Operated Vehicle"; In Fisher, S.S., Merritt, J.O, & Bolas, M.T. (Eds., 1994), *Proceedings of Stereoscopic Displays and Virtual Reality Systems*; SPIE Vol. 2177; pp.203-210.

Woods, A. (1997). **"**The Application of Stereoscopic Video to Underwater Remotely Operated Vehicles"; *Australian Petroleum Production & Exploration Association Journal*; pp.797-800.

Yorchak, J.P. (1986). "Teleoperator Human Factors Study"; MCR 86-558; Martin Marietta, Denver, CO; May, 1986.

- End of Document -

Subject to Protective Marking, National Caveats and announcement or availability restrictions, the information provided in this form will be published outside UK government and should, where possible, be of a commercially non sensitive nature.

To aid completion notes can be found at the end of this form.

Set Protective Marking of this form (See notes 3 & 4)

1. Report Number incl. Version Number:

2a. Report UK Protective Marking:

2b. Report National Caveats:

2c. Report Descriptors:

3. Title:

4a. Title UK Protective Marking:

4b. Title National Caveats:

5a.	**Primary Author SURNAME:**	**Primary Author INITIALS:**
5b.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5c.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5d.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5e.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5f.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5g.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5h.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5i.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**
5j.	**Secondary Author SURNAME:**	**Secondary Author INITIALS:**

6. Name and Address of Publisher	**7. Name and Address of MOD Sponsor**
8. MOD Contract Number:	**9. DSTL Project Number:**
10. Entity Number:	**12: Other Report Numbers:**
11. Work Package Number:	
13. Date of Publication (dd/mm/yyyy):	**14a. Contract Start Date (dd/mm/yyyy):**
15. Number of Pages:	**14b. Contract End Date (dd/mm/yyyy):**

16a. IPR: Please indicate the IPR conditions under which this report and abstract are supplied (See note 5):

16b. Patents: Have any patents or registered design applications been filed on any of the results described in this report? (See note 6) Application Number:

17. Abstract (Max length 2500 characters incl. Spaces; This abstract is intended for publication outside UK government and should be of a commercially non sensitive nature. For reports that are produced outside of the terms of Defcon 705, a non-commercial summary that is suitable for publication outside UK government must be provided):

18a. Abstract UK Protective Marking:

18b. Abstract National Caveats:

19: Keywords (see Note 7):

20a. Report Announcement: Indicate the various stakeholder / customer groups to whom the report can be announced. (Report announcement refers to the inclusion of titles and abstracts in alerting bulletins.) (See note 8)

IF NOR, please state the post of the release authority:

20b. Report Availability: Indicate the stakeholder / customer groups to whom the report can be released without further 'need-to-know' checks. (Report availability refers to the release of the full text of reports.) (See note 9)

IF NOR, please state the post of the release authority:

21. Authorisation (Complete as applicable)

	Name	**Signature**	**Date (dd/mm/yyyy)**
Department Manager			
Project Manager			
Technical Reviewer			

Notes

1. A copy of this form is to be completed by the primary author of this report. When completed, it is to be bound as the last numbered pages of the finished report to which it refers and forms the last numbered pages of the report. To insert the RDP into the report document place the cursor and the end of your report and select the 'Insert' menu, then select 'File'. Navigate to the RDP form and click 'Insert'. Once inserted the margins for the RDP section must be changed to the following: Top 1.02cm; Bottom 0.49cm; Inside 1.69cm; Outside 0.89cm; Gutter 0cm; Mirror margins. To use the form the document must be protected.

2. This form is provided in a locked format to enable data entry and should be used as such. Dropdown menus are provided for fields 2a, 4a, 16a, 16b, 18a, 20a, 20b and the header/footer (see point 3 below). It is the responsibility of the primary author named in this form to ensure that the data entered into this form is correct. If this form is unlocked functionality of the menus are lost and may lead to invalid or incorrect data being recorded which could result in inappropriate information release.

3. As the form is locked the headers/footers are not directly accessible. To set the protective marking in these fields select the desired protective marking from the dropdown menu marked 'PROTECTIVE MARKING NOT SET' then click File Print Preview. This will change the view to print preview and update these fields visibly then close the print preview. The headers/footers should now be updated to indicate the correct protective marking.

4. The marking 'PROTECT' should only be used with a recognised descriptor and not used in isolation.

5. This field should be used to record the contract IPR conditions under which the report was produced. Where DEFCON 705 applies, a Full Rights Version of the report should be supplied for ALL technical deliverables. An additional Limited Rights Version of the report may also be supplied if the Full Rights Version does not include all of the deliverable technical information. The copyright markings applied to the report must be in accordance with the requirements of the relevant IPR condition.

6. Where an application for a patent or a registered design has been made, please ensure that the D/IPR are informed in accordance with the relevant contract.

7. Authors may provide terms or short phrases which identify concisely the technical concepts, platforms, systems etc. covered by the report. Wherever possible, terms should be selected from the UK Defence Thesaurus which is available via the Defence Intranet.

8. Report Announcement Codes: UNL – UNLIMITED announcement; MOD – MOD and MOD Agencies; OGD – Other Government Departments and MOD; UKD – MOD and UK Defence Contractors; OSG – MOD and Overseas Defence Depts; NOR – Not to be Announced.

9. Report Availability Codes; UNL – UNLIMITED Distribution; MOD – MOD and MOD Agencies; OGD – Other government departments and MOD; UKD – MOD and UK Defence Contractors; OSG – MOD and Overseas Defence Depts; NOR – No release without approval of Release Authority.

10. For advice on using this form, please consult the Defence Research Reports Specification which can be obtained by emailing KISreports@dstl.gov.uk. If further information is required please email KISreports with "RDP advice" in the email subject field.