



Perceptual Error in Medical Practice

Dr Paul Greig

*MBChb (Hons) Bsc (Hons) FRCA
MAcadME PGCME*

**Pembroke College
University of Oxford**

Trinity Term, 2016

Thesis submitted for the degree of Doctor of Philosophy in Clinical Neurosciences

Abstract

Perceptual Error in Medical Practice

Submitted for the degree of Doctor of Philosophy in Clinical Neurosciences

Paul Greig MBChB (Hons) BSc (Hons) MAcadME PGCME FRCA

Pembroke College, University of Oxford

Trinity Term, 2016

Introduction

Medical errors are major hazards, and lapses in non-technical skills such as situational awareness contribute to most incidents. Risks are concentrated in acute care, and in crisis situations clinicians can apparently ignore vital information. Poor workplace ergonomics contributes to risk.

Existing work into perceptual errors offers insights, but these phenomena have been little researched in medicine. This thesis considers medical non-technical skills and how they are taught, and explores vulnerability to inattention and change blindness.

Methods

Medical human factors and the psychology of perceptual error were reviewed, and a mixed-methods assessment of postgraduate medical curricula completed.

Experiments assessed clinicians' interaction with clinical monitoring devices using eye-tracking, and studies were conducted exposing clinicians to various perceptual error stimuli using non-clinical and clinical videos, and simulation. A survey was also conducted to assess clinicians' insight into the phenomena of perceptual error.

Results

Non-technical skills feature poorly in medical curricula, and equipment is poorly standardised in critical care areas. Unfamiliar devices slow response times and increase error rate.

Clinical training confers no generalisable advantage in perceptual reliability. Even expert clinicians miss important events. Two out of every three life-support instructors for example missed a critical failure in the patient's oxygen supply when watching a recorded emergency simulation.

The insight and understanding healthcare staff have of perceptual errors is poor, leading to significant overestimates of perceptual reliability that could have consequences for clinical practice.

Conclusions

Perceptual errors represent a latent risk factor contributing to loss of situational awareness. High rates of perceptual error were observed in the video-based experiment. Although lower rates were observed in simulation, important events were still missed by participants that could have serious consequences.

The incidence of perceptual error appears sensitive to the method used to test for it, and this has important implications for the design of future experiments testing for these phenomena.

Mitigating perceptual error is likely to be challenging, but relatively simple adjustments to team practices in emergency situations may be fruitful.

Contents

<i>Abstract</i>	i
<i>Contents</i>	iii
<i>Abbreviations</i>	xi
1 Introduction.....	1
1.1 Chapter overview.....	1
1.2 Medical errors in perspective.....	1
1.2.1 The High-Reliability Organisation.....	5
1.3 Modelling accidents: the systems-based approach.....	6
1.4 Human Factors & Non-Technical Skills in Medical Practice.....	9
1.4.1 Human Factors & Non-Technical Skills defined.....	9
1.4.2 Situational Awareness.....	13
1.4.3 Situational awareness in the context of this thesis.....	15
1.5 Aims and structure of this thesis.....	16
1.6 References.....	19
2 Perceptual Limitation.....	24
2.1 Chapter overview.....	24
2.2 Human errors and the effects of cognitive demand.....	24
2.3 Attention, perception and the attentional set.....	30

2.4	Inattention blindness	34
2.5	Change Blindness	39
2.5.1	Perceptual errors in other sensory modalities.....	41
2.6	Perceptual Errors in Real-World Situations	42
2.7	References	46
3	Human Factors in Medical Training	52
3.1	Chapter overview.....	52
3.2	Introduction	52
3.3	Methods.....	55
3.3.1	Design.....	55
3.3.2	Subjects	55
3.3.3	Outcome Measures.....	55
3.3.4	Study Procedures	55
3.3.5	Analysis.....	58
3.4	Results.....	59
3.4.1	Human factors.....	62
3.4.2	Non-technical skills.....	63
3.4.3	Situational awareness	64
3.4.4	Human factors training	65
3.5	Discussion	68
3.5.1	Summary	71
3.6	References	73

4	From human factors to ergonomics: the impact of equipment design.....	75
4.1	Chapter overview.....	75
4.2	Introduction.....	76
4.3	Methods.....	77
4.3.1	Design.....	77
4.3.2	Subjects.....	77
4.3.3	Outcome Measures.....	77
4.3.4	Study Procedures.....	78
4.3.5	Analysis.....	80
4.4	Results.....	80
4.4.1	Familiarity Scores and Task Performance.....	80
4.4.2	Learning Effects.....	88
4.5	Discussion.....	89
4.5.1	Visual Search.....	90
4.5.2	Familiarity and search performance.....	98
4.5.3	Summary.....	103
4.6	References.....	105
5	Inattentional Blindness and Non-Clinical Stimuli.....	108
5.1	Chapter overview.....	108
5.2	Introduction.....	108
5.3	Methods.....	110
5.3.1	Design.....	110

5.3.2 Subjects 110

5.3.3 Outcome Measures 110

5.3.4 Study Procedures 110

5.3.5 Analysis..... 111

5.4 Results..... 112

5.5 Discussion 114

5.5.1 Object features and event detection 117

5.5.2 Predicting inattentional blindness risk..... 120

5.5.3 Summary 123

5.6 References 125

6 Inattentional Blindness and Clinical Stimuli..... 129

6.1 Chapter overview..... 129

6.2 Introduction 130

6.3 Methods..... 132

6.3.1 Design..... 132

6.3.2 Subjects 132

6.3.3 Outcome Measures..... 132

6.3.4 Study Procedures 132

6.3.5 Analysis..... 134

6.4 Results..... 135

6.4.1 Existing knowledge of the experimental paradigm..... 138

6.5 Discussion 139

6.5.1	Inattentional phenomena and other professional groups.....	144
6.5.2	The inverse-U relationship between load and performance	145
6.5.3	Expectation, memory, and change detection	147
6.5.4	Non-visual inattentional phenomena	148
6.5.5	Summary	151
6.6	References	152
7	Insight into Perceptual Errors amongst Healthcare Personnel.....	157
7.1	Chapter overview.....	157
7.2	Introduction	158
7.3	Methods.....	160
7.3.1	Design.....	160
7.3.2	Subjects	160
7.3.3	Outcome Measures.....	161
7.3.4	Study Procedures	161
7.3.5	Analysis.....	162
7.4	Results.....	163
7.4.1	Participant Demographics.....	163
7.4.2	Predictions of noticing versus actual noticing rates	167
7.5	Discussion	168
7.5.1	Summary	173
7.6	References	174
8	Workload and Perceptual Error during Simulated Cardiac Arrest.....	176

8.1	Chapter overview.....	176
8.2	Introduction	177
8.3	Methods.....	178
8.3.1	Study Design.....	178
8.3.2	Subjects	179
8.3.3	Outcome Measures.....	179
8.3.4	Study Procedures	179
8.3.5	Analysis.....	183
8.4	Results.....	183
8.4.1	Workload Ratings.....	183
8.4.2	Event Timing.....	184
8.4.3	Noticing Rates	185
8.4.4	Eye-tracking results.....	188
8.5	Discussion	189
8.5.1	Summary	197
8.6	References	198
9	Conclusions	201
9.1	Summary of findings	201
9.2	Methodological limitations of the programme of study	207
9.3	Clinical and policy implications of these data.....	211
9.4	Future opportunities for research	215
9.5	Final Comments	218

9.6	References	219
10	Appendices.....	220
10.1	Conflict of interests	220
10.2	Materials & Methods.....	221
10.2.1	Literature search strategy for chapter 1.....	221
10.2.2	Literature search strategy for chapter 2.....	221
10.2.3	Materials used for capturing eye-tracking data	222
10.2.4	Materials used for statistical analyses.....	223
10.2.5	Materials used in generating replica displays described in chapter 3.....	223
10.2.6	Materials used to develop the video sequence described in chapter 6.....	225
10.2.7	Materials used in the online survey described in Chapter 6	229
10.2.8	Materials used to collect online data for the experiment in Chapter 7	241
10.2.9	Materials used to employ the NASA Task Load Index described in chapter 8	260
10.2.10	Materials used to collect data for the experiment described in Chapter 8	262
10.3	Oxford University Hospitals Anaesthetic Monitor Survey.....	267
10.3.1	Introduction	267
10.3.2	Methods.....	267
10.3.3	Results.....	267
10.3.4	Discussion	269
10.3.5	References	270
10.4	Funding statement	271
10.5	Acknowledgements	272

10.7	Publications and Presentations Arising from this Thesis.....	274
10.8	Manuscripts of published work.....	275

Abbreviations

AAGBI	Association of Anaesthetists of Great Britain and Ireland
ABP	Arterial (invasive) Blood Pressure
ACAT	Acute Care Assessment Tool
AED	Automated External Defibrillator
ALMAT	Anaesthetic List Management Assessment Tool
AR	Anaesthetic Room
ARCP	Annual Review of Competency Progression
CbD	Case-based discussion
CEM	College of Emergency Medicine
CEX	Clinical Evaluation Exercise
CRF	Case Record Form
CVP	Central Venous Pressure
DOPS	Direct Observation of Procedural Skills
ECG	Electrocardiogram
EM	Emergency Medicine
EtAA	End-tidal Anaesthetic Agent concentration
EtCO₂	End-tidal Carbon Dioxide concentration
EtN₂O	End-tidal Nitrous Oxide Concentration
EtO₂	End-tidal Oxygen concentration
EWTD	European Working Time Directive
FiAA	Fraction-inspired Anaesthetic Agent concentration
FiCO₂	Fraction-inspired Carbon Dioxide concentration
FiN₂O	Fraction-inspired Nitrous Oxide Concentration
FiO₂	Fraction-inspired Oxygen concentration
GMC	General Medical Council
HR	Heart Rate
HRO	High Reliability Organisation
HSE	Health and Safety Executive
ICM	Intensive Care Medicine
ICU	Intensive Care Unit
LMA	Laryngeal Mask Airway
MiniCEX	Mini-Clinical Evaluation Exercise
MSF	Multi-source Feedback
NAT	Normal Accident Theory
NHS	National Health Service
NHSLA	NHS Litigation Authority
NIBP	Non-invasive Blood Pressure
NPSA	National Patient Safety Agency
NTS	Non-technical Skills
OT	Operating Theatre

OUH	Oxford University Hospitals NHS Foundation Trust
PbD	Problem-based discussion
RCoA	Royal College of Anaesthetists
RoI	Region of Interest
RR	Respiratory Rate
SpO₂	Oxygen saturations
Tv_{exp}	Tidal Volume (expired)
WHO	World Health Organisation
WPBA	Workplace Based Assessment

1 Introduction

'In a country where millions are spent every week on the national lottery, the concept of risk is obviously alien'

Alberti (1)

1.1 Chapter overview

Medical error is a significant, but often under-recognised, problem in modern healthcare. This chapter will present an overview of the scale and impact of medical error, setting the context for the experiments described later in this thesis.

The search strategy used to inform this chapter is presented in Appendix 10.2.1.

1.2 Medical errors in perspective

The deadliest operation ever conducted is thought to have been performed by Scottish surgeon Robert Liston (1794-1847). It was described as *'the only operation in history with a 300% mortality [rate]'* (2). The surgeon, operating in haste, managed to kill not only his patient, but also both an assistant and a bystander who was observing the procedure.

Medical science has progressed considerably in the years since. However, despite modern advances, errors continue to cause significant morbidity and mortality. If considered as a single disease medical error would rank as the sixth leading cause of death in the United States (3), contributing to 98,000 deaths every year (4, 5). Although this number is equivalent to a jumbo-jet crashing every 36 hours there has been little public outcry, and the medical profession has itself been slow to respond. This is in contrast to the high profile afforded other medical issues such as healthcare-associated infection, which carries comparable risk of mortality (6). Furthermore, there are signs that despite increasing awareness, error rates have improved little in recent years (7).

The situation in the United Kingdom resembles that of the USA. Approximately 670,000 people have contact with the National Health Service (NHS) every day (8), and up to 35,000 of them may be harmed by an adverse event (9).

Most of the risk of error is concentrated in acute-care environments, and hospital episodes account for more than half of NHS contacts (10). It is from hospitals that the overwhelming majority of the 1.3 million adverse events reported annually to the National Patient Safety Agency (NPSA) originate (11). The types of incidents recorded by the NPSA in 2010-2011 are summarised in Figure 1-1. Under-reporting of critical incidents is widespread, and the true number of incidents could be as high as 6.5 million (8).

Although estimates vary from study to study, it appears that approximately one patient in ten is affected by some form of medical error (12), and incidents in critical care environments occur twice as frequently as those originating in general wards (13, 14). Most medical errors are minor and do not result in serious or long-term harm to the patient, but up to 1 in 300 patients admitted to hospital suffers an error which contributes to their death (4, 5).

While the human cost of medical error is obvious and considerable, there is also a significant financial burden associated with these incidents. Prolongation of hospital stay as a consequence of adverse events is estimated to be approximately 8 days per incident (9, 15), costing approximately £1 billion annually (8). The extra resources required to treat patients who are the victim of error represents around 1% of total NHS spending, and even greater sums of money must be set aside to deal with the legal consequences that follow.

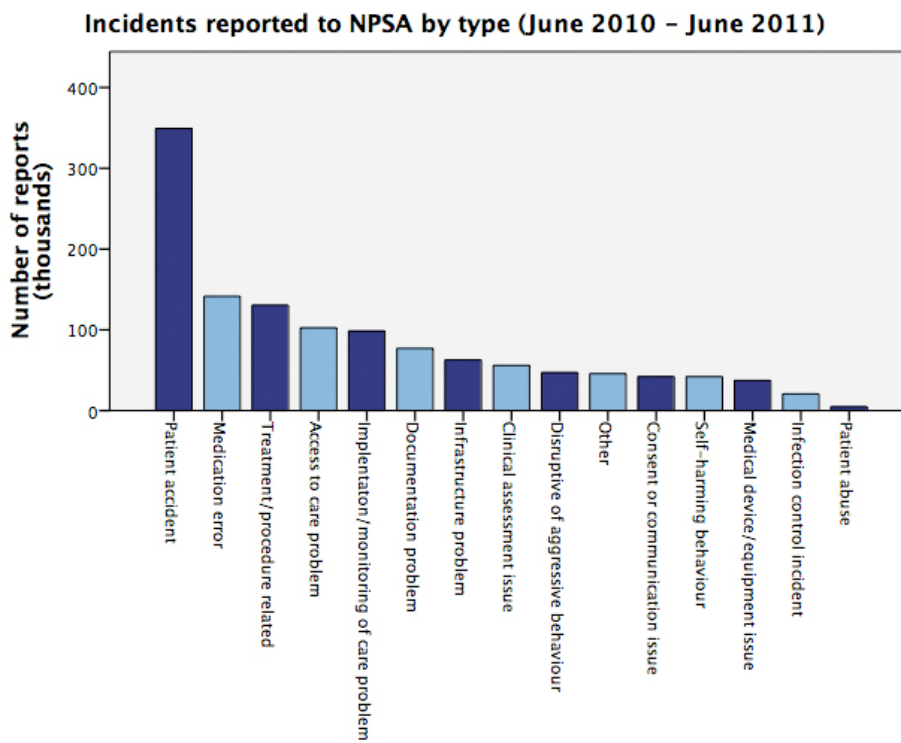
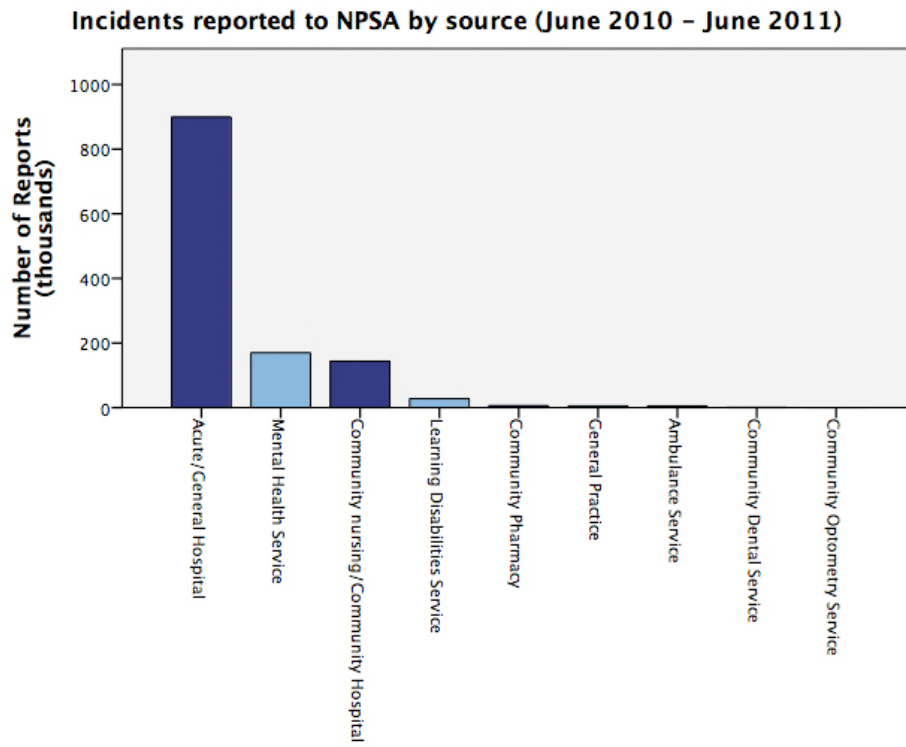


Figure 1-1: Categorisation of adverse events received by the NPSA by incident source and type (June 2010 – June 2011)

The NHS Litigation Authority (NHSLA) indemnifies all NHS Trusts against legal proceedings. In 2011-12 the NHSLA received 9,143 claims relating to clinical work, with 22,512 active claims on file as of 31st March 2012 (16). The cost of settling litigation in 2011-12 amounted to £1,330 million (17); of this sum, £214.5 million was spent on legal costs alone. Total NHSLA spending in 2011/2012 surpassed that of the entire Department for Energy and Climate Change (18). Considered against the total NHS budget (£104.3 billion in 2012-13 (19)) two pennies in every pound spent on healthcare is used to meet costs associated with medical error.

The clinical negligence scheme operated by the NHSLA has spent in excess of £5.3 billion over the last 10 years, and operates with outstanding liabilities of nearly £19 billion (17). Since claims reported to NHSLA take on average 15 months to resolve patients, their families and the staff involved are placed under considerable pressure for protracted periods of time.

Surgery contributes to much of the healthcare workload, and accounts for approximately 60% of adverse events in healthcare (20, 21). The World Health Organisation (WHO) estimates that 234 million operations are carried out annually worldwide (22). If estimates of error rates are globally representative then the number of people affected by surgical error annually is equivalent to the total population of Australia (23). In the UK, the NHS conducted nearly 10.5 million surgical procedures in 2011-12 (24), and by volume surgery represents the most common source of complaints received by the NHSLA.

Given the high prevalence of medical error, and the significant associated costs, one must inevitably ask why are errors so prevalent? Studies consistently find human error to be the leading cause of adverse events, accounting for up to two-thirds of critical incidents in non-healthcare settings (25), and the situation in medicine appears similar; closed-claims analyses have found human errors contribute to 80% of medical negligence cases (26).

One must hope that these errors do not occur because clinicians are wilfully negligent, so what then are the reasons that human error is so prevalent? Many factors in healthcare contribute to risk, and clinicians from critical-care specialties including anaesthetics, intensive care, and emergency medicine are at particularly high risk of error, reflecting the highly dynamic and complex environments in which they work. Staff depend on recalling large numbers of facts, often in time-critical or stressful environments. Fatigue can be an important contributor, and recent changes to out-of-hours service provision have increased the number of handovers, increasing risk of communication difficulties and disrupting continuity of care. These are all factors that are considered 'risky' in industrial settings.

1.2.1 The High-Reliability Organisation

Safety-critical industries such as aviation, railway transportation, and nuclear power have long recognised that such hazards exist. These industries are comparable with healthcare insofar as they are complex working environments with multiple potential risks, and when adverse events occur the consequences are likely to be significant. Steps have been taken over the last decade to adapt lessons learned from these industries into healthcare, many of which arise from the concept of the 'High Reliability Organisation' (HRO).

The HRO model arose from work in the late 1980s after it was recognised that certain industrial or military processes are inherently risky, and that these organisations must work consciously and continuously to minimise the potential for error (27). Risk-reduction is brought about by a combination of training, safe working practices, workplace culture, active monitoring, and robust incident reporting systems.

The aviation industry is often held up as the prototypical HRO, with an estimated mortality rate of under one per million opportunities (13). Psychology and ergonomics are used extensively in the development of aviation technology and working practices, and have been for several decades after it became clear that human error contributed to a high percentage of aircraft accidents (28). This led

to the development and delivery of structured human factors training to address communication and team working issues on the flight deck (29). Airline human factors training programmes typically focus on three domains: error avoidance through good communication and situational awareness, error trapping through safety-conscious behaviours, and mitigation by training in emergency drills.

Much work is on-going to develop the NHS in the UK into an HRO after high-profile failures, such as the death of Elaine Bromiley (30), have highlighted the urgent need to advance the understanding of human error in healthcare.

1.3 Modelling accidents: the systems-based approach

Errors occur where a sequence of actions fails to achieve a desired outcome, under circumstances that cannot be attributed to chance alone. They can be classified as failures either of planning (where a sequence of events is completed as intended but the strategy is itself faulty), or of execution (where a fault occurs after carrying out an appropriately drafted plan). Errors can also be considered as slips, lapses, or mistakes (31).

Sociology professor Charles Perrow described 'Normal Accident Theory' (NAT) in 1984 after investigations into the Three Mile Island nuclear power plant accident (32). NAT views errors as an inevitable feature of complex systems, and highlights the potential for cascading failures as coincidental malfunctions in different parts of a system accumulate, with unforeseen or unforeseeable effects.

NAT describes complex systems in terms of the interactivity of the variables within the systems and their coupling. A highly interactive system is one in which multiple variables are interdependent, and therefore deviations can propagate throughout the system in ways that may be opaque to the user. Tight coupling exists where actions have immediate and direct consequences. It makes a system less tolerant of change, since there is little scope for users to recover from errors before serious

consequences ensue. Inflexible processes that require actions to be taken in strict order, with only one route to a successful outcome, are at particularly high risk.

Critical care environments are an example of the complex, dynamic workplace that NAT would describe as high-risk (33). Since physiological parameters exhibit complex interdependency, not all of which are necessarily fully understood or predictable, patients can be considered as highly interactive systems. Most emergency situations demonstrate tight coupling, where decisive and specific corrective actions must be taken promptly to prevent undue harm occurring to the patient.

The Bromily case alluded to above illustrates some of the difficulties faced by clinical teams. The critical mistake – a fixation error which resulted in the team failing to follow a recognised failed intubation drill – was largely the responsibility of the anaesthetist, but this was not the only factor that contributed to the poor outcome. There were several other failures, notably in communication, overly hierarchical team structures, misallocation of resources, and failure to prioritise.

Recognising that critical incidents seldom arise from any single failure, psychologist James Reason devised the 'Swiss cheese' model of organisational accidents. According to this model the final, active error made by a system operator has usually been enabled by a series of preceding systemic weaknesses (Figure 1-2) (34).

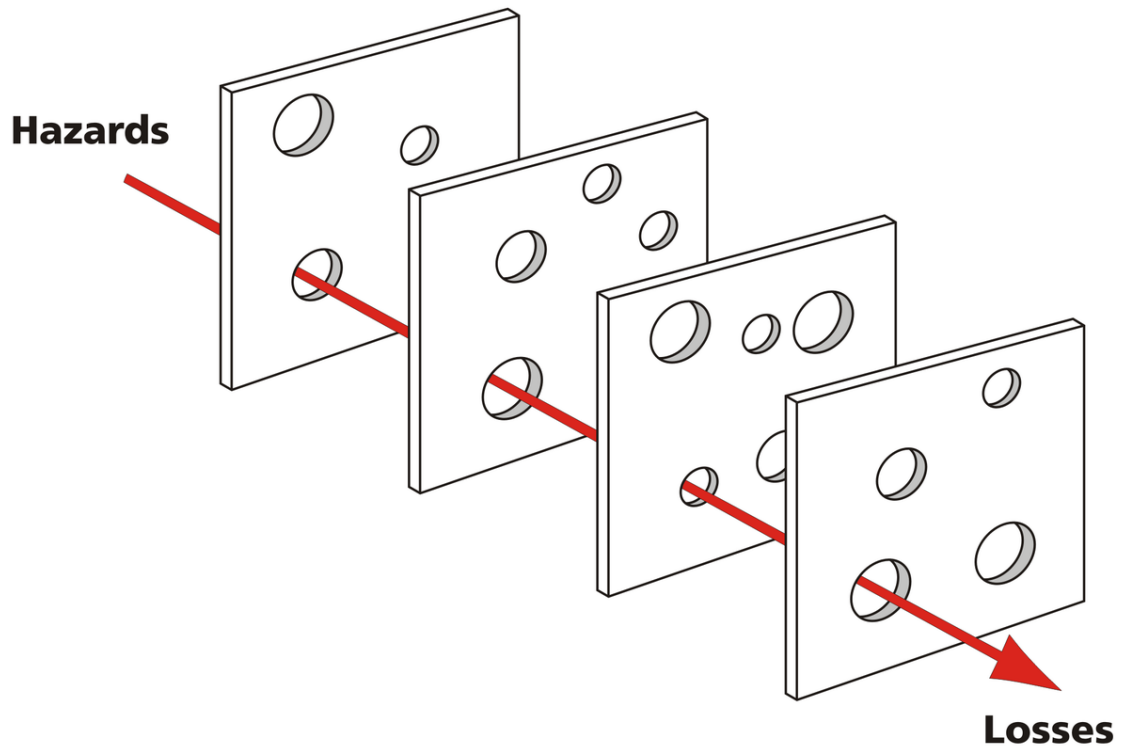


Figure 1-2: Reason's Swiss-cheese model of organisational accidents; reproduction of an original diagram by David Mack distributed under creative commons licence (CC BY-SA3.0)

Reason visualises systems as a series of steps. At each step an error can either be trapped or pass through to the next. These layers of defence are analogous to slices of cheese, while the holes represent weaknesses in the processes at each stage. Under normal circumstances any error that passes a layer will be captured by the next, but if certain errors coincide the holes will align, leading to a critical incident. Accidents therefore occur after a cascade of failures that, whilst individually inconsequential, can combine to overwhelm defences engineered into the system. The holes in each layer are described as latent faults.

These latent factors represent individual failings, training gaps, poor process design, or awkwardly designed equipment that each contribute to risk. While several such defects may be recognised in a process, it is possible for unknown latent faults to persist for protracted periods of time and may only come to light after an incident occurs. It is here that the value of high quality critical-incident (and near-miss) reporting systems becomes particularly relevant, as recurring themes in reports may highlight unsuspected latent factors.

The last line of defence in most processes is the end-user: the human operating the system. It is here that the final 'active' fault occurs to cause the accident. These active faults are generally the most proximate to the incident, so are the most obvious causative factors. Often, particularly in healthcare, investigations focus on the active fault and apportion much of the blame to the individual (33). Whilst individual chastisement may satisfy an impulse for retribution, it does nothing to address those on-going latent faults which, if allowed to persist, will contribute to future incidents. It is a feature of HROs that investigations adopt a systems-based approach and actively search for those intrinsic vulnerabilities that place staff in positions where risks can be taken (33, 35).

Several latent risks in the hospital environment have been identified (33). These include complex non-standardised equipment, inadequate staffing or training, distractions, and non-standard communications. A feature common to each of these is that they all place increased cognitive demands on the individual staff members. As shall be described later, workload may be a significant and under-recognised contributor to error and risk.

1.4 Human Factors & Non-Technical Skills in Medical Practice

1.4.1 Human Factors & Non-Technical Skills defined

Technical skills, the ability to perform practical tasks such as patient examination or cannulation, are obvious and vital contributors to delivering safe and effective healthcare. Less well recognised are the human factors which complement them.

The study of human factors originated in human-machine interactions, a branch of science now better known as ergonomics. The term later broadened into human interactions more generally (25, 36), and 'human factors' now draws upon experimental and cognitive psychology, sociology, and ergonomics, to address the complex interplay between people and teams, the devices they use, and the processes they work within. It is in this broader sense that the term is used in this thesis. This

definition is exemplified by the UK Health and Safety Executive (HSE), who describe human factors as *'environmental, organisational and job factors, and human and individual characteristics which influence behaviour at work...'* (37).

Beneath this umbrella term there exist a series of domains examining the individual facets of human factors described above. Non-technical skills (NTS) represent the behaviours and traits exhibited by people that impact on individual and team functioning, and were first described in aviation (38). Specific NTS training is part of the Joint Aviation Authorities requirement for commercial aviators, and aircrew are subject to regular review of both their technical and non-technical proficiency (39).

Specific attributes associated with NTS occur in two broad domains: mental skills (judgement, planning and situational awareness) and social skills (communication and team leadership) (40). In the context of this thesis, NTS are viewed through the prism of the Anaesthetic Non-Technical Skills (ANTS) system devised by the University of Aberdeen (41). This rating system further subdivides NTS into four sub-domains, entitled situational awareness, team-working, task-management, and decision-making. The ANTS system goes on to describe specific, observable behaviours that can be sought to determine a practitioner's achievement of each the elements that comprise the system. Many other NTS taxonomies exist, but this is one of the best established, and most widely used systems.

Although an understanding of human factors is a fundamental feature of an HRO, systematic study of their role in healthcare has only relatively recently become commonplace. The contribution that NTS make to patient safety was established after the publication in 2000 of the US Institute of Medicine report *'To err is human'* (4). Steps are currently being taken to develop NTS rating systems, and/or incorporate NTS training and assessment, for a variety of specialities. These include emergency medicine (42-44), anaesthetics (38), critical care (45), resuscitation (46-48), theatre nursing (49, 50), pre-hospital care (51), obstetrics (52), and general surgery (53, 54). NTS assessment has also been considered as a potential discriminator in recruitment to UK specialist training (55).

Training and maintaining NTS is important. In healthcare, data exist that demonstrate that when NTS are optimised in high-risk critical-care environments, error rates fall (14, 56). In the operating theatre NTS have been described as being '*as important as technical proficiency*' (40). These skills are not practiced in isolation, and it has been noted that technical performance correlates with non-technical skill (57-59). Unfortunately the ability to self-assess NTS is limited (60, 61), thus training and formal measurement by an expert faculty is required. This has led to an increasing role for medical simulation, with structured human-factors training now embedded into some under- and post-graduate medical training curricula.

Anaesthesia has been one of the earliest and most enthusiastic adopters of human factors training in hospital settings (62). This may have occurred for both philosophical and pragmatic reasons. It has been suggested that this pro-safety focus may result from the understanding that there is seldom a direct therapeutic benefit from undergoing anaesthesia (63). The role of the anaesthetist is therefore to ensure the safe conduct of a procedure under circumstances tolerable to the patient. Parallels have also been drawn between the highly technical, team-orientated practice of anaesthesia and aviation where a wealth of human factors and NTS literature already exists (64).

Anaesthetic interest in NTS may also be more pragmatic. One of the barriers to widespread adoption of human factors training in medicine is that although NTS are '*essential to learn, [they are] difficult to measure*' (65). The conduct of an anaesthetic has, unlike many medical interventions, a definable structure, and is reasonably tightly constrained. It may therefore lend itself more easily to industry-style process-analysis (66). This may simplify the development of taxonomies, to enable proper research into the role of NTS, and smooth the adaptation of lessons gleaned from the existing body of industrial human factors research.

Modern medical practice generally, and anaesthetics particularly, depends upon effective team working. Anaesthetists are often significant contributors to the management of critically ill medical

and surgical patients (67). They seldom work in isolation (68) and often must fulfil leadership roles, particularly during crisis situations (69).

Team leaders must be able to discharge their personal responsibilities effectively, whilst also allocating the physical and human resources available to them efficiently. This can be particularly challenging, because medical teams are usually transient in nature (70). It is not uncommon for emergency team members to work with each other for the first time over a critically unwell patient. Managing complex situations under these conditions taxes NTS, and leadership style can have a significant impact on safe working practices (71).

High levels of workload are associated with greater risk of critical incidents (72), an observation perhaps explicable by the increased cognitive demands placed on staff. Riem and colleagues model these effects using '*Cognitive Load Theory*' (57) which describes mental performance under pressure. The model acknowledges that cognitive resources, like any other, are limited in supply and must be carefully managed.

The limiting step in many cognitive processes is working memory. Riem's model suggests that mental shortcuts can be developed to circumvent the limitations of short-term memory by developing protocols and procedures which are stored in long-term memory. Although less immediately accessible long-term memory possesses much higher capacity. These shortcuts can be used to guide reactions during incidents, and these intuitive responses have been likened to a '*sixth sense*' (73).

Non-technical skills suffer under high cognitive demand. Reductions in workload therefore free cognitive resources, allowing more to be diverted from technical to non-technical skills such as communication and situational awareness. Improvements in non-technical skills have been associated with improved technical ability (74) and care-delivery during critical incidents (75). Andersen *et al* highlighted the need to '*anticipate and plan*', '*avoid fixation errors*' and '*allocate*

attention wisely' as important to effective delivery of critical care. The common thread linking these elements is that they are all required to develop good quality situational awareness.

1.4.2 Situational Awareness

Situational awareness is one of the most important factors in exercising judgement in the hospital environment (76). It represents the ability to maintain the overview necessary to function in complex situations (77), and the concept arose originally with the military. It first described the ability of combatants to maintain a mental model of the tactical environment in which they were deployed (78), and has since been applied to a variety of civilian activities in which information must be gathered and used to make decisions. In industrial settings, loss of situational awareness accounts for a high proportion of human-error related accidents (79, 80), and its loss has been described as *'the most frequent cause of errors in real time activities'* (81). Situational awareness is particularly crucial in critical-care (12) and emergency medicine (82), and is a marker of excellence in anaesthesia (69, 83). The concept is, however, also relevant to practitioners in less acute areas (84).

How individuals develop and maintain situational awareness has been the subject of research. People both construct models based on the information available and can also actively seek data to inform progress (85). Achieving good situational awareness (defined as having a complete and accurate understanding of events) depends on the integrated function of disparate cognitive functions. Elements of divided and focussed attention, short- and long-term memory, perception, and spatial awareness are all required (79).

Achieving good situational awareness can be viewed as a three-step process (85). These steps comprise firstly the basic perception of relevant information, followed by synthesis of these data into a mental model, and finally the projection of the constructed model into the future. Attainment of each level is based upon the one preceding, so that the quality of a mental model constructed about scenario is limited fundamentally by the completeness of data acquisition.

The process of developing situational awareness must begin with scanning the environment to gather salient data. This is both a conscious and sub-conscious process. The information acquired by the sense organs then passes through a series of pre-conscious filters which modulate sensory data after acquisition, but before the individual becomes aware of it. Much of the sensory data potentially available in the environment is discarded prior to reaching consciousness (a process over which the individual has little or no conscious control). Attaining this basic level of situational awareness, termed level one awareness by Endsley, lays the foundation for creating a more sophisticated understanding of events. The experimental chapters to follow will largely consider where faults can arise during these early stages in the development of situational awareness, which then limit clinicians' ability to respond to a scenario correctly.

Those data that reach consciousness will be integrated to form a coherent mental model of current events. The individual must then be able to use the constructed model to make predictions and decisions (86, 87). This allows the clinician to function pro-actively, one marker of expertise (85). If there is a failure to achieve adequate level one situational awareness then any mental model will be incomplete or inaccurate, and any decisions made using it may be faulty.

Maintenance of situational awareness, and the recognition of factors that contribute to its loss, are fundamental components in the medical NTS repertoire (88), and in most medical NTS taxonomies situational awareness is one of the headline categories (39, 42). It encompasses elements such as information-gathering, comprehension and anticipation (89), and while improvements in most NTS domains have been demonstrated by training, it is less clear that situational awareness is as amenable to teaching interventions (74, 79, 81).

Although anaesthesia is recognised as increasingly safe (69) it continues to be the case that human error is the leading cause of unintended morbidity and mortality (90), outweighing errors caused by equipment failure by nearly six to one (91). It is estimated that at least one in twenty anaesthetic incidents result from a lapse in attention or carelessness (91), and loss of situational awareness

appears to contribute to approximately 30% of reported critical incidents (92, 93). Trainee doctors may be particularly at risk of failures in situational awareness (94), a finding that is likely to be related to experience.

Jones and Endsley estimate 76% of situational awareness errors occur by failing to achieve level one awareness, in many cases despite all relevant information being nominally available (85). Diagnostic decision-making is particularly vulnerable to error (26), and a recent primary-care study investigating cancer misdiagnosis demonstrated that basic situational awareness failings were common (95).

1.4.3 Situational awareness in the context of this thesis

Situational awareness is an active process that is fundamentally dependant on perception, but psychologists can clearly demonstrate that perception is more than a simple matter of 'seeing' or 'hearing'. Far more data processing occurs at a subconscious level than most individuals realise, but if human cognitive resources are limited, what are the implications for perception when these resources are taxed? These questions have not been systematically addressed in medicine, and this thesis therefore intends to examine those situations in which particularly level one situational awareness may fail. In this I seek to build bridges between existing psychological experiments, and real-life clinical practice.

Situational awareness can be examined retrospectively or prospectively. Retrospective methods include analysis of medical records or critical incident data, but would be problematic in the context of this work. Record quality would be a major limitation, because insight into these psychological phenomena is generally poor (96), and it is known that critical incident reporting represents only the 'tip of the iceberg' regarding medical mishaps and near-misses. The psychology of situational awareness is seldom taught and little understood by frontline staff, so for these reasons it is highly unlikely that clinicians will have referred to the effects of interest in their reporting. Any analysis would therefore, at best, be based upon inference. Little insight could be gained into how and why

these effects occurred. Focus groups, interviews, or questionnaires would likewise be limited by the lack of insight into the psychological effects of interest.

Prospective methods offer the opportunity for richer data collection. Data could be obtained using direct observation of clinical practice, or indirectly using video recordings of relevant scenarios. This would be ethically complex, due to issues around patient confidentiality, and practically very difficult. While ten percent of patients might experience a critical incident over the course of their admission, there is no guarantee that any incident would be relevant to this study. Furthermore, the chance of catching any specific incident during any period of study would be relatively low, meaning excessive time periods would need to be observed to capture sufficient numbers of relevant events. The large number of confounding variables that would need to be accounted for in a functioning hospital environment would make any collected data very difficult to analyse.

Simulation was therefore chosen to offer the optimal balance between applicability to real-life practice and control over the experimental environment. Low-fidelity methods, such as using video-recordings, or high-fidelity methods, using a simulation manikin, mean that the experimental stimuli can be tightly standardised between participants, and can be designed to answer specific questions about the underlying psychological processes. By eliminating the potential to capture data not relevant to the experiment (particularly confidential data about patients) the ethical aspects of the research are greatly simplified, and costs are kept manageable. Finally, the methods can also be designed to parallel the original psychological experiments, making it straightforward to map findings from the experiments described here on to existing data from outside of medicine.

1.5 Aims and structure of this thesis

This thesis will investigate potential psychological explanations for why clinicians, even those with significant experience, can appear to ignore crucial information that should change the management of their patient. A greater understanding of how these apparent lapses occur is important to medicine, and these data will be most timely. Traditionally investigations into adverse events in

medicine have focussed on finding fault, often assuming that the individuals involved are at fault or indeed guilty of negligence (97). What if, instead, there were latent psychological risk factors the profession was as-yet unaware of? The following aims are therefore intended:

1. To understand the role perceptual errors play in clinicians' development of situational awareness
2. To understand how workload and experience interact to determine vulnerability to perceptual error in a variety of clinical situations
3. To define how different branches of medicine approach the teaching and acquisition of non-technical skills, with particular reference to situational awareness

The cognitive processes underpinning perception are little understood by most medical practitioners, and yet clinicians depend on these functions and suffer their limitations every day. The next chapter will therefore describe what has already been established about psychological phenomena of 'perceptual errors', and comment on how such effects limit the achievement of level one situational awareness in a non-medical context. This will lend context to the subsequent experimental chapters.

These investigations represent a new phenomenon in a clinical context, and therefore a case must be built incrementally. Although 'situational awareness' is well-understood in certain medical fields such as anaesthetics, it is by no means clear that all clinicians understand what is meant by the term. Chapter 3 will therefore consider how clinicians are taught non-technical skills, using a detailed analysis of the training curricula for hospital-based medical specialties. Understanding will then be gained about how specialties priorities and use non-technical skills in their workplace. This knowledge will be essential when disseminating material from this body of work to specialists outside of critical care areas where non-technical skills training is considered routine.

Subsequent chapters present a series of experiments which explore how and when clinicians can experience the effects of two specific perceptual failures, termed ‘inattention blindness’ and ‘change blindness’. This begins with an exploration of how devices that clinicians use can affect their development of situational awareness, presented in Chapter 4.

Next, in Chapter 5, a replication of a commonly cited psychological experiment is presented, albeit using a clinically-trained population. This experiment will test assumptions around whether clinical training confers any generalised perceptual advantage to those who receive it. Later, new resources will be developed to test different factors that affect and contribute to situational awareness in different clinical contexts. Chapter 6 uses established methods to test for perceptual errors in a healthcare population, but using for the first time published anywhere a clinically-relevant dynamic scenario.

Chapter 7 will present data about clinicians’ understanding of their vulnerability and insight into perceptual errors, while Chapter 8 follows up the prior video-based experiments with an exploration of perceptual error amongst clinicians in a high-fidelity simulated environment directly applicable to real-life practice.

Chapter 9 then draws these strands together to demonstrate how situational awareness can fail, not because clinicians are poorly trained or negligent, but rather because clinicians are in the end human, and subject to the psychological limitations this entails.

1.6 References

1. Alberti KGMM. Medical errors: a common problem. *British Medical Journal*. 2001;322(7285):501-2.
2. Gordon R. *Great Medical Disasters*. London: House of Stratus; 2001.
3. Hoyert D, Xu J. Deaths: Preliminary Data for 2011. *National Vital Statistics Reports*; vol 61 no 6. Hyattsville, MD: National Center for Health Statistics; 2012.
4. Kohn L, Corrigan J, Donaldson M. *To err is human: building a safer health system*. Washington, DC: National Academy Press; 2000.
5. Donaldson L. *Good doctors, safer patients*. London: The Stationery Office; 2006.
6. Kleven RM, Edwards JR, Richards CL, Horan TC, Gaynes RP, Pollock DA, et al. Estimating Health Care-Associated Infections and Deaths in U.S. Hospitals, 2002. *Public Health Reports*. 2007;122:161-6.
7. Sevdalis N, Hull L, Birnbach DJ. Improving patient safety in the operating theatre and perioperative care: obstacles, interventions, and priorities for accelerating progress. *Br J Anaesth*. 2012;109(suppl 1):i3-i16.
8. House of Commons Health Committee. *Patient Safety: 6th report of session 2008-2009*. London: The Stationery Office; 2009.
9. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *British Medical Journal*. 2001;322(7285):517-9.
10. Dew C. *Hospital Episode Statistics London: The Health and Social Care Information Centre; 2012* [Available from: <http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=1132>].
11. NRLS Quarterly Data Workbook (up to September 2011) [Website]. London: National Health Service Commissioning Board Special Health Authority; 2012 [updated 20/03/2012. Available from: <http://www.nrls.npsa.nhs.uk/resources/healthcare-setting/acute-hospital/?entryid45=133438>].
12. Vincent C. *Patient Safety*. 2nd ed. Oxford: Blackwell; 2010.
13. Bion JF, Abrusci T, Hibbert P. Human factors in the management of the critically ill patient. *Br J Anaesth*. 2010;105(1):26-33.
14. Reader TW, Flin R, Cuthbertson BH. Communication skills and error in the intensive care unit. *Curr Opin Crit Care*. 2007;13(6):732-6.
15. Sari AB-A, Sheldon TA, Cracknell A, Turnbull A, Dobson Y, Grant C, et al. Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital. *Qual Saf Health Care*. 2007;16(6):434-9.
16. NHS Litigation Authority. *The NHS Litigation Authority factsheet 3: claims information*. London: The Stationery Office; 2012.
17. NHS Litigation Authority. *The NHS Litigation Authority factsheet 2: financial information*. London: The Stationery Office; 2012.
18. Hoban M. *Budget 2011 (HC836)*. London: The Stationery Office; 2011.
19. Hoban M. *Budget 2012 (HC1583)*. London: The Stationery Office; 2012.
20. Flin R, Burns C, Mearns K, Yule S, Robertson EM. Measuring safety climate in health care. *Qual Saf Health Care*. 2006;15(2):109-15.
21. Yule S, Flin R, Maran N, Rowley D, Youngson G, Paterson-Brown S. Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World J Surg*. 2008;32(4):548-56.
22. World Alliance for Patient Safety. *Safe surgery saves lives*. Geneva: World Health Organisation Press, Safety WafP; 2008.
23. Department of Economic and Social Affairs. *World population prospects: The 2010 revision; Volume 1 Comprehensive Tables*. New York: United Nations, Department of Economic and Social Affairs P, Division; 2011.

24. Dew C. Hospital episode statistics 2011-2012: main procedures and interventions London: Health and Social Care Information Centre,; 2012 [Available from: <http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=215>].
25. Helmreich R, Davies J. Human factors in the operating room: interpersonal determinants of safety, efficiency and morale. *Baillieres Clinical Anaesthesia*. 1996;10(2):277-95.
26. Singh H, Thomas EJ, Petersen E, Studdert D. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Archives of Internal Medicine*. 2007;167(19):2030-6.
27. Sutcliffe KM. High reliability organizations (HROs). *Best Practice & Research Clinical Anaesthesiology*. 2011;25(2):133-44.
28. Fletcher G, Flin R, McGeorge P. The identification and measurement of anaesthetists' non-technical skills. Aberdeen: University of Aberdeen; 2003.
29. Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Anaesthetists' Non-Technical Skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth*. 2003;90(5):580-8.
30. Harmer M. Independent Review on the care given to Mrs Elaine Bromiley on 29 March 2005 London: Clinical Human Factors Group; 2005 [Available from: <http://www.chfg.org/articles-films-guides/elaine-bromiley-report>].
31. Reason J. Safety in the operating theatre - Part 2: human error and organisational failure. *Qual Saf Health Care*. 2005;14(1):56-60.
32. Perrow C. *Normal Accidents: Living with High-Risk Technologies*. New York: Basic Books; 1984.
33. van Beuzekom M, Boer F, Akerboom S, Hudson P. Patient safety: latent risk factors. *Br J Anaesth*. 2010;105(1):52-9.
34. Carthey J, Clarke J. *Implementing human factors in healthcare*. London: National Patient Safety Agency; 2010. Available from: www.patientsafetyfirst.nhs.uk.
35. Reason J. Human error: models and management. *British Medical Journal*. 2000;320(7237):768-70.
36. Scanlon MC, Karsh BT. Value of human factors to medication and patient safety in the intensive care unit. *Crit Care Med*. 2010;38(6 Suppl):S90-6.
37. Health and Safety Executive. *Reducing Error and Influencing Behaviour*. 2nd ed. London: HSE Books; 1999.
38. Flin R, Patey R, Glavin R, Maran N. Anaesthetists' non-technical skills. *Br J Anaesth*. 2010;105(1):38-44.
39. Flin R, Patey R. Non-technical skills for anaesthetists: developing and applying ANTS. *Best Practice & Research Clinical Anaesthesiology*. 2011;25(2):215-27.
40. Fletcher G, McGeorge P, Flin R, Glavin R, Maran N. The role of non-technical skills in anaesthesia: a review of current literature. *Br J Anaesth*. 2002;88(3):418-29.
41. Flin R, Maran N, Glavin R, Patey R. *Framework for Observing and Rating Anaesthetists' Non-Technical Skills*. Aberdeen: University of Aberdeen; 2003.
42. Flin R, Maran N. Identifying and training non-technical skills for teams in acute medicine. *Qual Saf Health Care*. 2004;13 Suppl 1:i80-4.
43. Steinemann S, Berg B, DiTullio A, Skinner A, Terada K, Anzelon K, et al. Assessing teamwork in the trauma bay: introduction of a modified "NOTECHS" scale for trauma. *Am J Surg*. 2012;203(1):69-75.
44. Flowerdew L, Brown R, Vincent C, Woloshynowych M. Development and validation of a tool to assess emergency physicians' nontechnical skills. *Ann Emerg Med*. 2012;59(5):376-85.e4.
45. Reader T, Flin R, Lauche K, Cuthbertson BH. Non-technical skills in the intensive care unit. *Br J Anaesth*. 2006;96(5):551-9.
46. Walker S, Brett S, McKay A, Lambden S, Vincent C, Sevdalis N. Observational Skill-based Clinical Assessment tool for Resuscitation (OSCAR): development and validation. *Resuscitation*. 2011;82(7):835-44.

47. McKay A, Walker ST, Brett SJ, Vincent C, Sevdalis N. Team performance in resuscitation teams: comparison and critique of two recently developed scoring tools. *Resuscitation*. 2012;83(12):1478-83.
48. Andersen PO, Jensen MK, Lippert A, Østergaard D. Identifying non-technical skills and barriers for improvement of teamwork in cardiac arrest teams. *Resuscitation*. 2010;81(6):695-702.
49. Mitchell L, Flin R, Yule S, Mitchell J, Coutts K, Youngson G. Evaluation of the Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS) system. *Int J Nurs Stud*. 2011.
50. Rutherford JS, Flin R, Mitchell L. Non-technical skills of anaesthetic assistants in the perioperative period: a literature review. *Br J Anaesth*. 2012;109(1):27-31.
51. Shields A, Flin R. Paramedics' non-technical skills: a literature review. *Emerg Med J*. 2012.
52. Bahl R, Murphy DJ, Strachan B. Non-technical skills for obstetricians conducting forceps and vacuum deliveries: qualitative analysis by interviews and video recordings. *Eur J Obstet Gynecol Reprod Biol*. 2010;150(2):147-51.
53. Crossley J, Marriott J, Purdie H, Beard JD. Prospective observational study to evaluate NOTSS (Non-Technical Skills for Surgeons) for assessing trainees' non-technical performance in the operating theatre. *Br J Surg*. 2011;98(7):1010-20.
54. Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. *Qual Saf Health Care*. 2009;18(2):104-8.
55. Gale TC, Roberts MJ, Sice PJ, Langton JA, Patterson FC, Carr AS, et al. Predictive validity of a selection centre testing non-technical skills for recruitment to training in anaesthesia. *Br J Anaesth*. 2010;105(5):603-9.
56. Mishra A, Catchpole K, Dale T, McCulloch P. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. *Surg Endosc*. 2008;22(1):68-73.
57. Riem N, Boet S, Bould MD, Tavares W, Naik VN. Do technical skills correlate with non-technical skills in crisis resource management: a simulation study. *Br J Anaesth*. 2012;109(5):723-8.
58. Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N. The impact of nontechnical skills on technical performance in surgery: a systematic review. *J Am Coll Surg*. 2012;214(2):214-30.
59. Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and error in the operating room: analysis of skills and roles. *Ann Surg*. 2008;247(4):699-706.
60. Arora S, Miskovic D, Hull L, Moorthy K, Aggarwal R, Johannsson H, et al. Self vs expert assessment of technical and non-technical skills in high fidelity simulation. *Am J Surg*. 2011;202(4):500-6.
61. Moorthy K, Munz Y, Adams S, Pandey V, Darzi A, Imperial College--St. Mary's Hospital Simulation G. Self-assessment of performance among surgical trainees during simulated procedures in a simulated operating theater. *Am J Surg*. 2006;192(1):114-8.
62. Gaba DM. Anaesthesiology as a model for patient safety in health care. *British Medical Journal*. 2000;320(7237):785-8.
63. Toff NJ. Human factors in anaesthesia: lessons from aviation. *Br J Anaesth*. 2010;105(1):21-5.
64. Gronlund SD, Dougherty MRP, Durso FT, Canning JM, Mills SH. *Planning in Air Traffic Control (DOT/FAA/AM-01/16)*. Washington, D.C.: Federal Aviation Administration Office of Aerospace Medicine; 2001.
65. Zausig YA, Grube C, Boeker-Blum T, Busch CJ, Bayer Y, Sinner B, et al. Inefficacy of simulator-based training on anaesthesiologists' non-technical skills. *Acta Anaesthesiologica Scandinavica*. 2009;53(5):611-9.
66. Phipps D, Meakin GH, Beatty PCW, Nsoedo C, Parker D. Human factors in anaesthetic practice: insights from a task analysis. *Br J Anaesth*. 2008;100(3):333-43.
67. Handy JM, Morris C. The evolving roles of the anaesthetist during emergency medical care. *Anaesthesia*. 2013;68:1-2.
68. Smith AF, Mishra K. Interaction between anaesthetists, their patients, and the anaesthesia team. *Br J Anaesth*. 2010;105(1):60-8.

69. Larsson J, Holmström IK. How excellent anaesthetists perform in the operating theatre: a qualitative study on non-technical skills. *Br J Anaesth*. 2013;110(1):115-21.
70. van Beuzekom M, Boer F, Akerboom S, Dahan A. Perception of patient safety differs by clinical area and discipline. *Br J Anaesth*. 2013;110(1):107-14.
71. Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson G. Teaching surgeons about non-technical skills. *Surgeon*. 2007;5(2):86-9.
72. Weinger MB. Human Factors Research in Anesthesia Patient Safety: Techniques to Elucidate Factors Affecting Clinical Task Performance and Decision Making. *J Am Med Inform Assoc*. 2002;9(90061):58S-63.
73. Smith AF, Arfanis K. I. 'Sixth sense' for patient safety. *Br J Anaesth*. 2013;110(2):167-9.
74. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual Saf Health Care*. 2009;18(2):109-15.
75. Andersen PO, Jensen MK, Lippert A, Ostergaard D. Identifying non-technical skills and barriers for improvement of teamwork in cardiac arrest teams. *Resuscitation*. 2010;81(6):695-702.
76. Gosbee J. Handoffs and communication: the underappreciated roles of situational awareness and inattention blindness. *Clinical Obstetrics & Gynecology*. 2010;53(3):545-58.
77. Busby S, Witucki-Brown J. Theory development for situational awareness in multi-casualty incidents. *Journal of Emergency Nursing*. 2011;37(5):444-52.
78. McManus J, Huebner K, Scheulen J. The science of surge: detection and situational awareness. *Acad Emerg Med*. 2006;13(11):1179-82.
79. O'Brien KS, O'Hare D. Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*. 2007;50(7):1064-91.
80. Singh H, Petersen LA, Thomas EJ. Understanding diagnostic errors in medicine: a lesson from aviation. *Qual Saf Health Care*. 2006;15(3):159-64.
81. Walker GH, Stanton NA, Kazi TA, Salmon PM, Jenkins DP. Does advanced driver training improve situational awareness? *Appl Ergon*. 2009;40(4):678-87.
82. Flowerdew L, Brown R, Vincent C, Woloshynowych M. Identifying nontechnical skills associated with safety in the emergency department: a scoping review of the literature. *Ann Emerg Med*. 2012;59(5):386-94.
83. Smith AF, Glavin R, Greaves JD. Defining excellence in anaesthesia: the role of personal qualities and practice environment. *Br J Anaesth*. 2011;106(1):38-43.
84. Johnston PW, Fioratou E, Flin R. Non-technical skills in histopathology: definition and discussion. *Histopathology*. 2011;59(3):359-67.
85. Endsley M, Garland D. Theoretical underpinnings of situation awareness: a critical review. *Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
86. Fedack K, Fedack JM. Fixed wing transport airway management utilizing a situational awareness paradigm. *Air Medical Journal*. 2008;27(1):30-6.
87. Wright M, Endsley M. Building shared situation awareness in healthcare settings. In: Nemeth C, editor. *Improving healthcare team communication: building on lessons from aviation and aerospace*. Aldershot, UK: Ashgate; 2008. p. 97-114.
88. Fioratou E, Flin R, Glavin R, Patey R. Beyond monitoring: distributed situation awareness in anaesthesia. *Br J Anaesth*. 2010;105(1):83-90.
89. Patey R, Flin R, Fletcher G, Maran N, Glavin R. Developing a taxonomy of anaesthetists' nontechnical skills. In: Henriksen K, Battles JB, Marks ES, editors. *Advances in patient safety: from research to implementation*. 4. 2011/01/21 ed. Rockville, Maryland: Agency for Healthcare Research and Quality; 2005.
90. Flin R, Fletcher G, McGeorge P, Sutherland A, Patey R. Anaesthetists' attitudes to teamwork and safety. *Anaesthesia*. 2003;58(3):233-42.

91. Cooper JB, Newbower RS, Long CD, McPeck B. Preventable anesthesia mishaps: a study of human factors. *Anesthesiology*. 1978;49(6):399-406.
92. Staender S, Davies J, Helmreich B, Sexton B, Kaufmann M. The anaesthesia critical incident reporting system: an experience based database. *Int J Med Inf*. 1997;47(1-2):87-90.
93. Marcus R. Human factors in pediatric anesthesia incidents. *Paediatr Anaesth*. 2006;16(3):242-50.
94. Nyssen AS, De Keyser V. Improving training in problem solving skills: analysis of anaesthetists' performance in simulated problem situations. *Le Travail Humain*. 1998;61(4):387-401.
95. Singh H, Giardina TD, Petersen LA, Smith MW, Paul LW, Dismukes K, et al. Exploring situational awareness in diagnostic errors in primary care. *BMJ quality & safety*. 2012;21(1):30-8.
96. Levin DT, Momen N, Drivdahl S. Change blindness blindness: the metacognitive error of overestimating change-detection ability. *Visual Cognition*. 2000;7(3):397-412.
97. Reason J. Beyond the organisational accident: the need for "error wisdom" on the frontline. *Quality and Safety in Health Care*. 2004;13(suppl 2):ii28-ii33.

2 Perceptual Limitation

'We cannot change the human condition, but we can change the conditions under which humans work'

Reason (1)

2.1 Chapter overview

The psychology of attention is not generally taught at medical school. It does not feature at all in, for example, the curriculum of the University of Oxford Medical School (2). This chapter will therefore present a brief overview of the psychology of error and attention, as it pertains to this thesis.

Data will be presented on psychological phenomena of special relevance to the achievement of level 1 situational awareness or, more correctly, in its failure. These data will therefore establish the current understanding of perceptual errors, and create a framework for the subsequent experimental chapters.

The search strategy used to inform this chapter is presented in Appendix 10.2.2.

2.2 Human errors and the effects of cognitive demand

Errors represent unintentional acts resulting in failure to accomplish a task (3). Even highly-trained people are fallible, particularly when placed in complex environments or working under stress. Fallibility arises because the brain has finite processing capacity. As sensory or cognitive demands increase there are fewer mental resources available (4), therefore marshalling cognitive capacity may be one of the most important skills in healthcare (5).

Many people have somewhat naïve beliefs about the fidelity of perception (6), and it is often assumed that sense organs function somewhat like recording devices which passively

acquire information from the individual's environment. Developing awareness of one's surroundings is actually a complex, multi-stage process, and the subjective experience that arises is based upon information that has been variously tested, filtered, and modified prior to reaching consciousness (7). There is interaction between elements in the external world and the internal processes of the brain that determine how an object is represented in the mind of an observer.

The primate brain is highly adapted for visual perception. The visual system underpins basic functions such as object avoidance, and also more complex tasks such as facial or textual recognition (8). The exact mechanisms by which perception occurs are complex and incompletely understood. Early researchers such as Gibson characterised the process as 'bottom-up' (9), a process whereby perception is driven by intrinsic object-features such as conspicuousness or motion. Gibson theorised that to understand perception one had to consider how the brain interpreted light on the retina. He termed the visual information available from the environment the '*ambient optic array*' and that this was composed of various features that contributed to the interpretation of scene.

Later researchers considered an algorithmic approach to these bottom-up processes (10), investigating how the brain actually derives meaning from visual information. This led to greater appreciation of the 'top-down' processes that influence perception. These are processes that originate endogenously, for example accessing memory to aid in recognition of familiar objects. This was described as a constructivist approach, whereby the use of information known to the individual contributes meaning to new data as they are received. This process has been likened to hypothesis generation (8), as the brain creates a series of interpretations for each scene and searches for the best fit. Sometimes the best-fitting interpretation may be incorrect, leading to misinterpretation or misrecognition of objects, and these processes may become saturated at times of high perceptual load.

Memory is also important to successful task completion, and is subject to limitations. Short-term memory stores information of immediate relevance, and encoding data into it rarely requires conscious effort on the part of the individual. This contrasts with long-term memory which is of much higher capacity, but requires repeated exposure or deliberate practice to store data. Different systems exist within short-term memory for the storage of images or sounds, but storage capacity is small, limited to around seven independently associated items, which can be held for periods of up to eight seconds (11). Memory capacity is important to situational awareness, and indeed, one in ten situational awareness errors occurs after information that was successfully registered is subsequently forgotten (12).

Certain situations can arise which make the capacity of short-term memory appear much higher than these data suggest. Experienced chess players, for example, can retain the position of 15 or more pieces and reconstruct a game in progress after viewing play for only a few seconds (13). Superficially this might lead one to conclude that chess mastery confers significant memory advantages, but if the pieces are instead scattered randomly across the board this apparent advantage disappears. This is a powerful example of the use of 'perceptual chunking'; objects in memory that share common properties can be grouped and encoded as single items. In this instance then, the experienced chess player can recognise patterns and meanings in the game in progress and use these relationships to group pieces. Randomly placed pieces share no such commonalities and must be stored individually, resulting in far fewer items being recalled successfully.

Short-term memory can be likened to a labelling system in which category titles are stored that point to locations in long-term memory for data retrieval (13). Improved pattern-recognition and more efficient perceptual chunking is one feature of expertise (14), which enables individuals to work beyond the apparent limits of cognition. For any given task

novices report a higher perceived workload than do experts (15), thus this ability to better chunk items is a feature of emerging expertise. The effect of experience may also be mediated by memory-guided attentional cueing (16). The experienced practitioner can use pattern recognition and memories of previous exposure to similar situations to improve their visual search strategy. This is analogous to findings from professional sports, whereby expert players can more quickly discern meaning from games in-play than novices (13, 17, 18).

After collecting and storing information these data can then be used to make decisions. Decision-making is a dynamic process that can be approached using a variety of strategies, reviewed by Glavin in a 2011 paper (19). Glavin describes four main methods by which an individual can reach a decision; creatively, through object-appraisal, based on rules, or primed by recognition.

Creative solutions are the most cognitively demanding to reach, but are required when a novel solution is required for a newly-encountered problem. Some decisions are, however, constrained by external factors such as availability of equipment, time, or personnel. In these cases, the individual must evaluate a range of choices and choose from within the generated options. This process of 'option appraisal' is therefore somewhat less cognitively demanding than seeking a solution within an unconstrained decision-space.

Rules-based decision making is still less demanding. This occurs when actions can be guided by protocolised knowledge (19). Here the decisions have in a sense been made in advance, and applying the rules requires fewer mental resources than having to seek a creative solution. This, of course, depends upon there being a understanding that learned rules apply.

The least-cognitively demanding decision is 'recognition-primed'. Here a previously encountered problem is recognised, and a known solution applied with minimal thought,

but the likelihood of making this recognition depends on previous experience. Decision-making of all types is degraded when task complexity, time pressure, or stakes are high.

One factor that impacts importantly on decision-making in modern, technologically sophisticated workplaces, is the availability of equipment and tools. Human interaction with devices has been widely studied, and in industrial settings much effort has been expended to make the use of technology seamless. This is not necessarily the case in healthcare, where nearly 20 years ago Reason described theatres as an '*ergonomic nightmare*' (20) and arguably this is a situation that continues today in most critical-care environments (21). Many clinicians fail to recognise the risks that are posed by their poor working environments (22).

One example from a relatively newly-built theatre complex in a major teaching hospital illustrates this point. When built, the piped medical gas supply outlets were placed upon the back wall of each anaesthetic room. This constrained the positioning of equipment, and meant that the anaesthetic machine must be placed behind the anaesthetist, making it impossible to reach the controls or observe monitors without turning away from the patient (Figure 2-1).

Whilst the individual devices installed in this room have been carefully designed, and each functions correctly when operated, the way in which the environment has been organised adds to hazards and increases the physical and mental demands made of staff working here. When one considers that more than half of the steps required to anaesthetise a patient are equipment-related (23), the scope for error is significant.

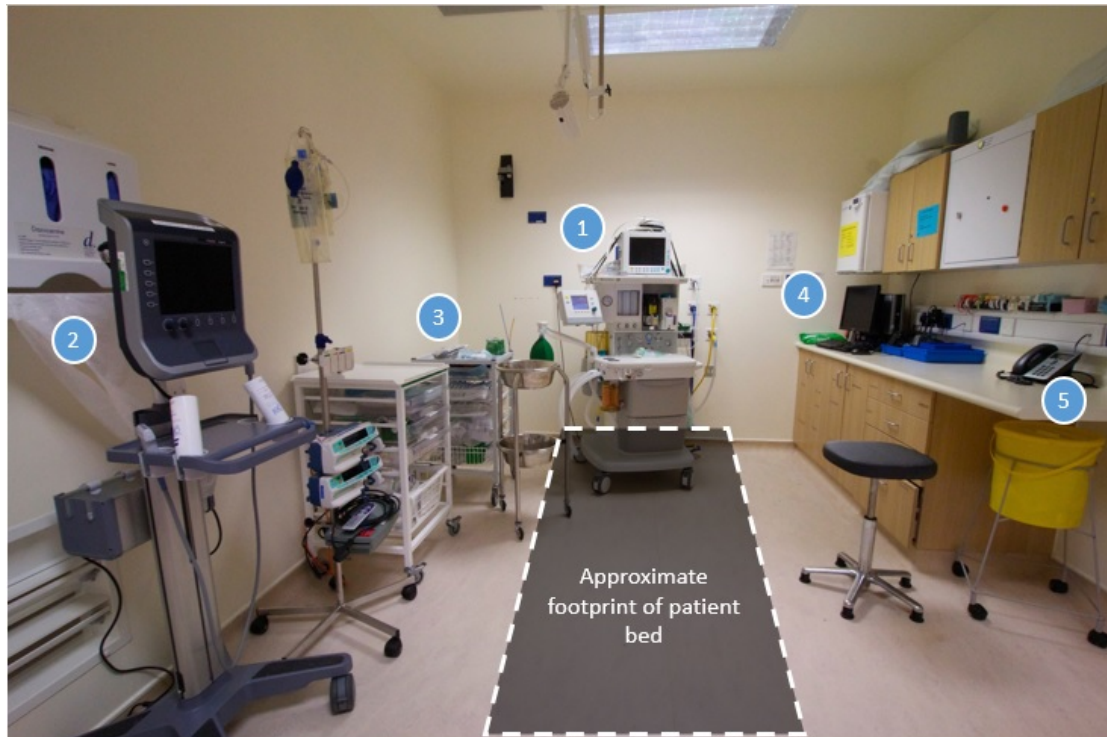


Figure 2-1: Image of an anaesthetic room highlighting a series of ergonomic problems; (1) the anaesthetist must position themselves between the patient and anaesthetic machine, making it difficult to see both the patient and the monitoring equipment; (2) access to personal-protective equipment is obstructed by other devices; (3) the anaesthetic assistant is positioned on the wrong side of the bed to be able to reach the supply cupboards; (4) there is no leg-space available when sitting at the computer; and (5) access to the sharps bin is obstructed by the worktop [Photo credit: Paul Greig]

Equipment is also poorly standardised, even within individual theatre complexes. This compounds risk, because staff have fewer opportunities to develop familiarity with the operation of each device. Decisions involving the operation of unfamiliar equipment will likely be more cognitively demanding, and standardisation is associated with improved performance in clinical practice (24).

The patient is clearly at the heart of safety interventions, but it is worth considering that there are benefits for staff as well. Improvements in working hours and staffing ratios reduce fatigue and physical demands, while care bundles, guidelines, and checklists all reduce the mental demands of clinical practice by moving towards more rules-based decision making. More than simply providing a pleasant workplace, better working practices coupled with environmental and task optimisation are important to reducing distractions and allowing staff to focus on their primary duties.

2.3 Attention, perception and the attentional set

Although vigilance is recognised as a vital component of safe anaesthesia (25) few clinicians understand the cognitive factors which contribute. In medical literature vigilance is generally portrayed as a bottom-up process; sufficient information will inevitably be acquired if the individual attends to their environment, and data-sources such as device alarms are made sufficiently conspicuous.

The reality is much more complex. Alluded to in the previous chapter, the death of Elaine Bromiley (26) is a tragic, but informative, demonstration of how these perceptual and cognitive failures can combine to result in a very serious adverse outcome. In this case, a team of appropriate trained staff in a well-equipped facility failed to recognise the progression of a failed-intubation situation. This was despite national monitoring standards being met, and therefore all the necessary information to make a diagnosis being available to the team. There is no doubt that the team involved were appropriately motivated to resolve the situation, so there can be no question that they actively ignored the information. Why then, did they fail to integrate critical information such as pulse oximetry into their mental models? Why should their situational awareness have failed so badly? The reality of attention, perception, and therefore situational awareness, is that these processes are subject to limitations in cognitive processing (27), and therefore prone to failure in certain situations.

The terms 'perception' and 'attention' are closely related, but are not synonymous (28). Attention should be considered as the devotion of sensory and cognitive resources toward a stimulus, whereas perception represents the subjective, conscious awareness of that stimulus. This distinction is important, because the subject of whether perception can occur in the absence of attention has been the subject of much research activity.

Although our subjective experience of visual perception is highly detailed, there is much evidence that perception is neither automatic nor complete (29, 30), and there are data demonstrating that objects can be looked at (i.e. be subjected to attention) but not seen (perceived). Natural scenes are complex and continually varying, and theoretically there is a limitless amount of data that could be encoded about each view. Human cognition is clearly a limited resource, and would be overwhelmed if all this information were to receive detailed processing. To prevent overload, mechanisms must exist to selectively filter out extraneous non-task-relevant details.

These mechanisms are described by perceptual load theory (31, 32), which posits that in order to minimise processing load the brain gathers only sufficient data to gain a broad overview of any given scene (18, 33). Detail is then filled in as required, likened by some to using the environment as an *'external memory bank'* (34). As mental workload increases, attentional focus must narrow (12).

Psychologist Ronald Rensink notes that there is a balance between the details stored about an object versus the level of meaning assigned to it, stating:

'No visual representations are both coherent and detailed... all that is really needed is a representation of only those objects—and those particular properties—involved in the task at hand' (18)

He further suggests that the detail we subjectively perceive is an illusion created by the brain, which instead generates a *'virtual representation'* of each object. This minimalist framework comprises little detail but can readily be completed as circumstances require. Rensink notes that this method of representation is *'highly sensitive to the demands of the task and the expectations of the observer'*. This expectation-dependency may make individuals prone to perceptual failures caused by defects in attention.

Even the most highly conspicuous objects cannot be perceived without some level of attention being directed towards them (35). Attentional focus is to an extent under voluntary control, so individuals can intentionally focus on a particular item based on motivation and expectation (36). There is evidence of cultural differences in how people view scenes and the meaning they extract from them (37), and inter-individual differences including age (38), affect (39), and experience of the specific task (40, 41) also contribute to intrinsic variability in the processing of information. It is, however, unlikely there are fundamental differences in underlying perceptual or cognitive ability between those who experience perceptual failures and those who do not (42).

Items can also capture attention reflexively. The likelihood of capture is based on situational features such as uniqueness or conspicuousness. Items may be inherently conspicuous by virtue of unusual shape or colour, or may carry particular personal relevance for the individual. These are termed sensory- and cognitive-conspicuousness respectively (43). Emotional factors, particularly negative emotional content, may contribute to cognitive conspicuousness (44).

These features then influence how the brain selects for relevancy. An 'attentional set' operates pre-consciously to detect items of particular relevance to the task at hand (45), and to filter out items that are irrelevant to the goals of the individual (processing of which would otherwise waste cognitive capacity). These irrelevant objects can then be automatically ignored. This means, for example, when instructed to track white squares moving across a display screen few or no details may be stored about non-relevant black objects. Actively ignoring irrelevant or distracting stimuli has also been described as '*selective looking*' (46), but difficulty arises because this process is not fully under conscious control so if important task-relevant information becomes available from an object not part of the attentional set it is likely to be missed. This phenomenon is termed a 'perceptual

error', because the information associated with the object is discarded *before* the individual becomes subjectively aware of the item.

An attentional set can therefore be considered an automatic means of reducing cognitive workload, and its function has been described as '*tune[ing] attention for features that then receive visual priority*' (47). Most *et al* showed that when experimental participants performed a task involving coloured objects they were more likely to notice unexpected stimuli with similar colour or shape to the items they were tracking. It was reasoned this resulted from the new stimulus falling within the attentional set. It was also noted that counting accuracy fell in participants confronted with increasing numbers of distractors even when they were not consciously aware of them. The authors suggested this was evidence of additional cognitive resources being consumed by pre-conscious processing of the additional stimuli.

It is known that objects of high personal significance, and therefore high cognitive conspicuousness, are less prone to being ignored. The 'cocktail party effect' is one such example (48), whereby an individual can filter out background chatter in a noisy environment but still retain the ability to detect their name from amongst the unattended conversations. Similar effects operate in the visual domain (49), and changes in a scene that affect people or animals are often noted earlier than inanimate objects (50). It may be that while the contents of the attentional set are generated on a task-by-task basis, certain objects do automatically receive attentional prioritisation.

The limitations enforced by availability of cognitive resources, and the systems employed by the brain to operate within these limits, are therefore a double-edged sword. They offer the subjective experience of a richly detailed environment, but they also carry the potential for errors. In most daily activities, these errors may carry few consequences, but in high risk workplaces these slips may have serious consequences.

If these phenomena are intrinsic features of cognition, and therefore affect all members of the workforce, it is incumbent on high-reliability organisations (HROs) to design robust systems that can accommodate these sources of error safely. The first stage in mitigating the effects of any risk is to increase systemic awareness of the contributing factors (5). While inattention (in the lay sense of missing important details) is frequently cited as a contributing factor in human errors in clinical practice, there has not yet been a systematic investigation into the role of perceptual limitations on ability to work in a healthcare setting.

Certain recurring failures of perception have been extensively demonstrated by psychologists over the last 40 years. Two of the best-investigated are termed 'inattention blindness' and 'change blindness'. Both have significant implications for clinicians.

2.4 Inattention blindness

Although many find it difficult to accept that perception can be highly unreliable, perceptual errors are in fact a part of day-to-day experience. Consider only that an anatomical blind spot exists on the retina at the optic disc which produces results in a blind spot, *'equivalent to the size of a lemon held at arm's length'* (34). Despite this sizeable gap in visual field a blind spot is not a part of our daily experience. Consciousness has adapted to function seamlessly within these limits, and only when specific demonstrations are designed to highlight the effect do people become aware of its presence.

Inattention blindness is an error of perception associated with failure to perceive objects or events that intuitively should be easily visible. It is associated with perceptual failure occurring whilst performing tasks that require concentration, and is demonstrable using a variety of computer-, laboratory- or naturalistic models. Investigations into the effect began with laboratory studies, using tightly controlled stimuli to elicit mechanisms that result in the phenomena. More recent work has extended these findings to situations closely

applicable to real-world tasks. Investigators have examined the phenomenon in the context of sports (51), driving (52), and in criminal investigations (53).

Neisser first described inattention blindness in 1975 using a pair of superimposed video sequences. Observers were instructed to attend to events in one sequence whilst ignoring activity in the other (54), and many failed to notice odd or otherwise conspicuous events in the unattended stream. Mack and Rock later re-examined the phenomenon in the late 1990s, requiring volunteers to discriminate between the lengths of arms on a cross briefly displayed on a computer screen. During critical trials an additional shape was introduced to the display, and participants were later asked about their awareness of this unexpected stimulus. Again most were completely unaware of the additional shape, even though they could easily locate it when not engaged in the length-discrimination task (55).

The phenomenon was further explored under dynamic conditions in a famous experiment by Simons and Chabris (56). This experiment demonstrated inattention blindness could be induced during observation of a real-world scenario. Observers watched a video comprising two teams passing basketballs. One team was dressed in white, the other in black. Observers counted either the total number of passes, or kept simultaneous counts of two different types of throw.

Part way through the video sequence a collaborator dressed in a gorilla costume walked slowly through the scene, pausing in the centre and then exiting the stage. In the original experiment 46% of naïve observers overall failed to notice the gorilla. Even greater proportions missed the event when maintaining two separate pass-counts (increasing cognitive demands) or when another scene was superimposed over the basketball video (increasing perceptual load). This has since become known as the 'Invisible Gorilla' and the experimental video has been viewed on Internet video site YouTube in excess of seven million times (57). The sequence has been highly successful in publicising the effect of

inattentional blindness amongst non-specialist audiences. The effect observed in this experiment has been termed 'central inattentional blindness'¹, reflecting that the effect arises from a lack of cognitive capacity (58).

Workload has a major effect on the likelihood of noticing the critical stimulus. Under conditions of high perceptual load (more items to attend to) or high cognitive load (increasing task difficulty) inattentional blindness is more likely as fewer resources can be devoted to less salient features (59). Fewer objects from outside the attentional set can then be highlighted by pre-conscious processing and passed to conscious awareness. This explains why observers counting the white team's passes were less likely to notice the gorilla (which was coloured black) than observers tracking the black team. The attentional set of the observer in this case was tuned to white moving objects, meaning that information about black objects was simply discarded pre-consciously. This effect is a consistent finding across a range of experiments using a variety of different experimental paradigms (28, 47, 60, 61).

Although inattentional blindness can occur in the absence of any distracting stimuli, and can even blind people to the only object displayed on screen (62), the phenomenon is more prevalent when there are higher numbers of distractors (46). Sub-threshold distractors that are perceived, but which fail to reach consciousness, may still 'tie up' cognitive resources and impair performance even when observers are subjectively unaware of them (45, 49).

The highest rates of inattentional blindness are seen in experimentally-naïve participants exposed to an unexpected stimulus for the first time. One might expect that having been exposed to a particular inattentional blindness stimulus the critical event would then be reliably noticed on repeated viewings. Some data suggest this may not be the case.

¹ Other forms of inattentional blindness have also been described; *spatial* inattentional blindness for example arises from attentional misdirection at the point the critical event occurs. This is a technique frequently employed by stage magicians.

In 2010 Simons exposed subjects who had prior knowledge of the 'Invisible Gorilla' experiments to a modified version of the video sequence. This second video, entitled 'Monkey Business', contained new critical events in addition to the presence of the gorilla (63). Subjects in this experiment could reasonably be expected to anticipate the nature of the experiment and would therefore be watching in a heightened state of vigilance. All participants did indeed notice the gorilla, and therefore specific foreknowledge appeared to protect them from one specific perceptual error. Perhaps surprisingly though the subjects were not more likely to notice the sequence's other events, and may have even been less likely to detect them. They therefore enjoyed no generalizable improvement in observational reliability. Simons notes these findings may reflect 'search satisfaction', but it could equally be postulated that additional attentional resources expended on searching for the (semi-expected) gorilla consumed processing power that might otherwise have been devoted to other (truly unexpected) occurrences. Conscious effort to maximise vigilance cannot therefore guarantee reliable perception for all stimuli.

Task-difficulty is therefore a critical predictor of vulnerability (64). Various studies have observed that task-familiarity reduces risk (51, 65), probably because expertise confers advantages in cognitive efficiency (13, 17). Unfortunately these benefits are not transferrable outside the field of expertise (42).

Predicting which individuals are susceptible to inattention blindness is challenging. The difficulties in retesting subjects mean that few data are available to identify individual traits associated with vulnerability. Some evidence does imply differences between 'noticers' and 'non-noticers', for example visual search strategy or distractibility may vary (66), and it has often been suggested that people with lower working memory capacity are at increased risk (60, 65, 67, 68) although this finding is not universal (69).

Many of these findings are relatable to clinical practice. Induction of- and emergence from-general anaesthesia are the periods of greatest workload for the anaesthetist (15, 70), and this is where many critical events do occur. Interestingly, and despite the lower task demands, a relatively high number of incidents are still observed during the maintenance phase of anaesthesia (23, 71). This seems somewhat paradoxical, however operating under conditions of minimal workload have also been associated with perceptual failures. Low cognitive load conditions have been associated with increased distractibility in laboratory tests (72), and boredom can lead to inattention (73). Disengagement with the task (a particular risk where the activity is repetitive, protracted, or unvarying) is associated with reduced attention and easier distractibility (74).

Listening to music is a common practice in the operating theatre, and might be assumed to consume cognitive resources that would increase the risk of inattention blindness. This has been criticised a safety hazard that can impair communication (75), but there are data to suggest that an appropriate level of background music could confer benefits (59). In this paper Beanland suggested improved visual attention may be mediated either by changes in affect; by means of forcing observers to distribute attention more widely; by a loosening of the attentional set; or by suppression of internal distractors. It may therefore be justifiable to permit a low-level of background music in the workplace, although clearly if it impairs the sharing of information or conflicts with auditory prompts from monitoring equipment this must be remedied.

Failures in vigilance or monitoring contribute to up to 60% of medical errors (76). Clinical teams perform in high-stress conditions with high mental and physical workloads so it is plausible to expect that inattention blindness contributes to a high proportion of such incidents. While inattention blindness is potentially the most commonplace error in the dynamic clinical environment, given the importance of monitoring to detect trends other

perceptual errors such as change blindness probably also contribute significantly to loss of situational awareness in the medical workplace

2.5 Change Blindness

Change blindness occurs when observers fail to notice changes in a visual scene, particularly (but not exclusively) when the change coincides with some manner of interruption in the view (77). Change blindness can be classified according to the visual artefacts that cause it (18, 78, 79):

- *Shift-contingent*: blindness occurs to changes occurring during visual saccades
- *Blink-contingent*: blindness occurs to changes occurring during an eye blink
- *Occlusion-contingent*: blindness occurs after obscuring the site of the change with another object
- *Cut-contingent*: blindness occurs to changes occurring across movie scenes
- *Splat-contingent*: blindness occurs during the simultaneous display of distracting objects which do not necessary obscure the changing object
- *Time-contingent*: blindness occurs to gradual change occurring over a period of several seconds, even when the change is in plain sight

Change blindness arises through a combination of perceptual and memory artefacts. Detection of change depends both on visual awareness and working memory, both of which are subject to limitations in processing ability. It appears that certain features of attention, executive function, working memory, and visual processing correlate with change detection ability in certain tasks (80).

The sparseness of visual representations described earlier in this chapter means that most cases of change blindness probably arise through failure in comparison of pre- and post-change scenes (30, 81, 82). Objects are not continuously perceived in their entirety, and

instead are sampled dynamically according to task relevance (83). This explains why individuals can remain blind to changes in objects that are the subject of attention (84), and in the absence of any additional distracting stimuli (85). Expectations modulate change-detection because observers direct more focus on areas of a scene they expect to change, although as with inattention blindness this may be at the cost of missing events in areas they monitor less vigilantly. Cueing observers to expect an event does not therefore automatically protect them from change blindness (30, 33).

Change blindness is not only a laboratory phenomenon; this effect was highlighted in the experiment conducted by Simons and Levin in which a conversation partner was changed during the interaction (86). This has subsequently become known as the 'Door Experiment', as the switch was conducted while two collaborators interrupted the conversation by carrying a door between the experimenter and subject. The change was overlooked by nearly 50% of participants overall, and certain features of the target affected the likelihood of the participant noticing the change. Simons stated that during social interactions individuals apply high-level labels to the representation of the person with whom they are conversing. If the change does not violate this categorisation (for example by changing race, gender, or social group) the change is unlikely to be noted. Behavioural continuity may also be relevant to change detection (29).

Few people are aware how limited perceptual abilities are. As with inattention blindness people greatly overestimate their ability to detect changes between scenes in movies, photographs and even live encounters (87). Individuals have such poor insight into their perceptual limitations that Levin coined the term '*change-blindness blindness*' (88) – 70-90% of people believe they will detect changes in movie sequences that in fact none notice when tested. There are also data to suggest that a discontinuity exists between *change-detection* and *change-identification*. People are often better at determining when and where change

has occurred, but are often less able to specifically describe the change (89). It is a reasonably consistent feature of many change- and inattentional-blindness experiments that spontaneous reports of critical stimuli are somewhat lower than responses to prompting questions, so it does appear that some memory traces persist for events that people are initially unaware of.

2.5.1 Perceptual errors in other sensory modalities

Inattentional- and change-detection errors are most closely associated with the visual system, but similar phenomena have been noted in other sensory modalities: analogues of both visual change-blindness and inattentional blindness have been detected using auditory stimuli (labelled change deafness (90) and inattentional deafness (91) respectively). Inattention phenomena have also been observed using tactile stimuli (55).

Some of the earliest work to consider selective attention, the foundations of inattentional and change blindness, was actually conducted using non-visual stimuli (80). Dichotomous listening tasks were studied in the 1950s and 1960s in which participants were presented with different audio-tracks simultaneously in each ear. Participants in these experiments were instructed to follow one stream, whilst ignoring the other. Such studies highlighted that individuals can filter out the unattended voice and fail to notice changes such as speaker or language, although people do remain sensitive to certain cognitively conspicuous prompts (such as their own name) in the unattended stream.

Like visual processing, auditory processing is similarly impaired by workload. For example, a deterioration in response time to verbal commands, and a fall in word recognition accuracy, is observed when stimuli are uttered by a changing voice (92). It is hypothesised that the additional attentional resources required to attend to the new vocal characteristics are responsible, and this is consistent with the perceptual load theory described by Lavie.

It is interesting to note that increasing workload in one sensory modality also adversely impacts on cognitive processing elsewhere. Studies of auditory processing demonstrate that when stimuli do not reach consciousness they are still subjected to significant pre-attentive processing, even to the point that they can influence a person's behaviour (49, 93). There are data that increasing auditory workload can trigger visual inattentive phenomena (94, 95). The reciprocal has also been demonstrated; that inattentive deafness could be induced under conditions of high visual load (4). Sinnett termed the process underlying this '*shared attentional capacity*', and this is important to consider when relating experimental data to the multi-sensory environment encountered in the real-world.

The finding that perceptual resources are not modality-specific but instead must be spread across all incoming stimuli is highly relevant to the hospital environment, where relevant data can be derived from spatially disparate sources and presented using a variety of sensory modalities including visual, auditory, touch, and on occasion smell.

2.6 Perceptual Errors in Real-World Situations

Perceptual errors such as change-blindness or inattentive blindness are well-described, robust phenomena that are relatively easy to demonstrate in controlled settings, but what of their impact in real-world situations? Although not always referenced specifically, and perhaps not always defined as such by the authors, the effects of perceptual failures can often be detected in descriptions of critical incidents.

Aviation accident investigations are often a rich source of human factors case studies. Reading a report of the crash of Eastern Airlines Flight 401 on December 29, 1972 reveals a set of circumstances in which a flight crew apparently remained unaware that the aircraft was descending, despite a functional altimeter and auditory alarms. The crew had become fixated on a minor fault in an indicator lamp, and apparently ignored the gradual loss of altitude until only moments before the aircraft struck the ground. The accident investigation

recorded the cause of the crash as *'the failure of the flight-crew to monitor the flight instruments during the final four minutes of flight, and to detect an unexpected descent soon enough to prevent impact'* (96). It is not a radical leap to suggest that inattentive blindness- and deafness were likely factors in this accident.

In more everyday circumstances, driving is one task that is certainly affected by perceptual error (52, 97). Visual inattention is frequently implicated in road-traffic accidents, and is disproportionately cited in collisions involving inexperienced drivers (98). The risk to drivers is also sensitive to mental workload. Conversing with another individual, either in person (41) or using mobile telephony (94, 99, 100), adversely affects performance (80), and even experienced drivers can be affected.

Other routine tasks can be adversely affected by increasing mental demand. Hyman *et al* examined the reaction of pedestrians to an unexpected event whilst they crossed a University square in which a collaborator in a clown costume was unicycling. A significant proportion of people were unaware of the clown, and mobile-phone users were the least likely to report seeing the unicyclist. Interestingly, participants who were walking with a partner were up to three times more likely to report the experimental stimulus (101). Studies on military teams have also demonstrated benefits of working in teams on situational awareness, provided that members can communicate freely (102). These data raise the interesting possibility that teams are to a degree protected against the perceptual failures that may impair the function of its individual members, a finding that could be usefully employed within the medical environment.

While perceptual error has received attention in financial trading (103), aviation (104), and military exercises (105, 106), less attention has been paid in healthcare. Only a handful of studies have considered perceptual error in clinical practice, and none have considered it systematically. Only a single case report to date has implicated inattentive blindness in a

real critical incident after a femoral line guidewire was lost in a patient, and then missed on a series of radiographs in the days following (107). It is likely that this paucity of data reflect more a failure to recognise the phenomena, than any immunity against them amongst the medical profession.

Clinicians generally assume that simply increasing the conspicuousness of objects will be enough to make important information visible (108, 109), however this is not automatically true. As Simons himself points out, the infamous gorilla is certainly conspicuous, and study participants often express shock when confronted with stills of the sequence (56). Another example can be found in the provision of workplace emergency equipment. Fire extinguishers are deliberately designed to be highly conspicuous, and the average office worker should never be more than 30m away from one (110). It has, however, been demonstrated that the vast majority of employees are unable to describe the location of fire extinguishers in their own workplaces (111). It is likely that while conspicuousness assists in perception, if the item is not related to the individual's immediate needs it will not be part of their attentional set and therefore is liable to be overlooked.

Manipulation of the attentional set to include all possible items is neither plausible nor practical. The only pragmatic measures to reduce inattention blindness risk would be those that offer consistent reductions in mental workload. Closer integration of theatre equipment, standardisation of devices and layout, and simulator-based training are all steps that may be beneficial. Increased uptake of checklists and guidelines in crisis situations, already standard practice in other HROs, may also offer cognitive advantages by decreasing the mental workload necessary to manage such events.

Change-blindness-blindness, and the overconfidence many people have in their perceptual abilities, has also been associated with real-world consequences. Levin highlights a 1998 study that found 90% of criminal convictions later overturned on DNA evidence were based

on erroneous eye-witness statements, and commented on a false association between the confidence with which testimony was delivered and the perceived accuracy of the statement (29). Other perceptual errors, including inattentional blindness, may have also contributed to wrongful convictions (53).

Perceptual errors have demonstrable real-world implications, but it is not yet clear how these phenomena impact on clinical practice. The following experiments intend to lend insight into how these affect healthcare staff, and to offer suggestions for how the effects might be mitigated.

2.7 References

1. Reason J. Human error: models and management. *British Medical Journal*. 2000;320(7237):768-70.
2. Curriculum Development Committee. *Integrated Curriculum in Medicine and Surgery 2014-2015*. Oxford: University of Oxford Medical School; 2015.
3. Reason J. *Human Error*. Cambridge, UK: Cambridge University Press; 1990.
4. Macdonald JSP, Lavie N. Visual perceptual load induces inattentive deafness. *Attention, Perception, and Psychophysics*. 2011;73(6):1780-9.
5. Fletcher G, McGeorge P, Flin R, Glavin R, Maran N. The role of non-technical skills in anaesthesia: a review of current literature. *Br J Anaesth*. 2002;88(3):418-29.
6. Winer GA, Cottrell JE, Gregg V, Fournier JS, Bica LA. Fundamentally misunderstanding visual perception: Adults' belief in visual emissions. *American Psychologist*. 2002;57(6-7):417-24.
7. Nobre AC, Kastner S, editors. *The Oxford Handbook of Attention*. Oxford: Oxford University Press; 2014.
8. Braisby N, Gellatly A. *Cognitive Psychology*. 2nd ed. Oxford, UK: Oxford University Press; 2012.
9. Gibson J. *The senses considered as a single system*. Boston, MA: Houghton Mifflin; 1966.
10. Marr D. *Vision: a computational investigation into the human representation and processing of visual information*. New York: WH Freeman & Co; 1982.
11. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*. 1956;63(2):81-97.
12. Endsley M, Garland D. *Theoretical underpinnings of situation awareness: a critical review*. *Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
13. Chase WG, Simon HA. Perception in chess. *Cognitive Psychology*. 1973;4(1):55-81.
14. Garland DJ, Barry JR. Cognitive advantage in sport: the nature of perceptual structures. *Am J Psychol*. 1991;104(2):211-28.
15. Weinger MB, Herndon OW, Zornow MH, Paulus MP, Gaba DM, Dallen LT. An Objective Methodology for Task Analysis and Workload Assessment in Anesthesia Providers. *Anesthesiology*. 1994;80(1):77-92.
16. Summerfield JJ, Lepsien J, Gitelman DR, Mesulam MM, Nobre AC. Orienting attention based on long-term memory experience. *Neuron*. 2006;49(6):905-16.
17. Furley P, Memmert D, Heller C. The dark side of visual awareness in sport: Inattentive blindness in a real-world basketball task. *Attention, Perception, and Psychophysics*. 2010;72(5):1327-37.
18. Rensink RA. Change detection. *Annual Review of Psychology: Annual Reviews Inc.*; 2002. p. 245-77.
19. Glavin RJ. Human performance limitations (communication, stress, prospective memory and fatigue). *Best Practice & Research Clinical Anaesthesiology*. 2011;25(2):193-206.
20. Reason J. *Human Error*. British Medical Association Conference on Healthcare Errors; London2000.
21. Wong SW, Smith R, Crowe P. Optimizing the operating theatre environment. *ANZ J Surg*. 2010;80(12):917-24.
22. Wears RL, Perry SJ. Human factors and ergonomics in the emergency department. *Ann Emerg Med*. 2002;40(2):206-12.
23. Phipps D, Meakin GH, Beatty PCW, Nsoedo C, Parker D. Human factors in anaesthetic practice: insights from a task analysis. *Br J Anaesth*. 2008;100(3):333-43.

24. Arriaga AF, Bader AM, Wong JM, Lipsitz SR, Berry WR, Ziewacz JE, et al. Simulation-Based Trial of Surgical-Crisis Checklists. *N Engl J Med.* 2013;368(3):246-53.
25. Vincent C. Patient Safety. 2nd ed. Oxford: Blackwell; 2010.
26. Harmer M. Independent Review on the care given to Mrs Elaine Bromiley on 29 March 2005 London: Clinical Human Factors Group; 2005 [Available from: <http://www.chfg.org/articles-films-guides/elaine-bromiley-report>].
27. Nobre A, Correa A, Coull J. The hazards of time. *Current Opinion in Neurobiology.* 2007;17(4):465-70.
28. Most SB, Scholl BJ, Clifford ER, Simons DJ. What you see is what you set: sustained inattentive blindness and the capture of awareness. *Psychological Review.* 2005;112(1):217-42.
29. Levin DT, Drivdahl SB, Momen N, Beck MR. False predictions about the detectability of visual changes: The role of beliefs about attention, memory, and the continuity of attended objects in causing change blindness. *Consciousness and Cognition.* 2002;11(4):507-27.
30. Levin DT, Simons DJ, Angelone BL, Chabris CF. Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology.* 2002;93(Pt 3):289-302.
31. Lavie N. The role of perceptual load in visual awareness. *Brain Research.* 2006;1080:91-100.
32. Lavie N, Lin Z, Zokaei N, Thoma V. The role of perceptual load in object recognition. *Journal of Experimental Psychology: Human Perception and Performance.* 2009;35(5):1346-58.
33. Simons DJ. Current approaches to change blindness. *Visual Cognition.* 2000;7(1-3):1-15.
34. O'Regan JK. Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology.* 1992;46(3):461-88.
35. Cohen MA, Alvarez GA, Nakayama K. Natural-scene perception requires attention. *Psychological Science.* 2011;22(9):1165-72.
36. Doherty JR, Rao A, Mesulam MM, Nobre AC. Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience.* 2005;25(36):8259-66.
37. Masuda T, Nisbett RE. Culture and change blindness. *Cognitive Science.* 2006;30(2):381-99.
38. Graham ER, Burke DM. Aging Increases Inattentive Blindness to the Gorilla in Our Midst. *Psychology and Aging.* 2011;26(1):162-6.
39. Becker MW, Leininger M. Attentional selection is biased toward mood-congruent stimuli. *Emotion.* 2011;11(5):1248-54.
40. Werner S, Thies B. Is 'change blindness' attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition.* 2000;7(1-3):163-73.
41. White CB, Caird JK. The blind date: the effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accident; analysis and prevention.* 2010;42(6):1822-30.
42. Memmert D, Simons DJ, Grimme T. The relationship between visual attention and expertise in sports. *Psychology of Sport and Exercise.* 2009;10(1):146-51.
43. Watson DS. Inattentive blindness: what you don't see may harm your patient. *AORN journal.* 2010;92(1):110-3.
44. Most SB, Chun MM, Widders DM, Zald DH. Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review.* 2005;12(4):654-61.

45. Simons DJ. Attentional capture and inattentive blindness. *Trends in Cognitive Sciences*. 2000;4(4):147-55.
46. Koivisto M, Revonsuo A. The role of unattended distractors in sustained inattentive blindness. *Psychological Research*. 2008;72(1):39-48.
47. Most SB, Astur RS. Feature-based attentional set as a cause of traffic accidents. *Visual Cognition*. 2007;15(2):125-32.
48. Wood N, Cowan N. The Cocktail Party Phenomenon Revisited: How Frequent Are Attention Shifts to One's Name in an Irrelevant Auditory Channel? *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1995;21(1):255-60.
49. Bressan P, Pizzighello S. The attentional cost of inattentive blindness. *Cognition*. 2008;106(1):370-83.
50. New JJ. A content-specific attenuation of change blindness: Preferential attention to animate beings in natural scenes. *Journal of Vision*. 2003;3(9):643a.
51. Memmert D. The effects of eye movements, age, and expertise on inattentive blindness. *Consciousness and Cognition*. 2006;15(3):620-7.
52. Strayer DL, Drews FA, Johnston WA. Inattention-blindness behind the wheel. *Journal of Vision*. 2003;3(9):157a.
53. Chabris CF, Weinberger A, Fontaine M, Simons DJ. You do not talk about Fight Club if you do not notice Fight Club: Inattentive blindness for a simulated real-world assault. *i-Perception*. 2011;2(2):150-3.
54. Neisser U, Becklen R. Selective looking: Attending to visually specified events. *Cognitive Psychology*. 1975;7(4):480-94.
55. Mack A, Rock I. *Inattentive blindness: an overview*. Cambridge, MA: MIT Press; 1998.
56. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattentive blindness for dynamic events. *Perception*. 1999;28(9):1059-74.
57. Simons DJ. Selective Attention Test: YouTube.com; 2010 [updated 10/3/2010. Available from: <http://www.youtube.com/watch?v=vJG698U2Mvo>].
58. Most SB. What's "inattentive" about inattentive blindness? *Consciousness and Cognition*. 2010;19:1102-4.
59. Beanland V, Allen RA, Pammer K. Attending to music decreases inattentive blindness. *Consciousness and Cognition*. 2011;20(4):1282-92.
60. Beanland V, Pammer K. Minds on the blink: The relationship between inattentive blindness and attentional blink. *Attention, Perception, and Psychophysics*. 2012;74(2):322-30.
61. Most SB, Simons DJ, Scholl BJ, Jimenez R, Clifford E, Chabris CF. How not to be seen: the contribution of similarity and selective ignoring to sustained inattentive blindness. *Psychological Science*. 2001;12(1):9-17.
62. Most SB, Alvarez GA. But it's the only thing there! Sustained inattentive blindness for a solitary stimulus. *Journal of Vision*. 2002;2(7):444a.
63. Simons DJ. Monkeying around with the gorillas in our midst: familiarity with an inattentive-blindness task does not improve the detection of unexpected events. *i-Perception*. 2010;1(1):3-6.
64. Simons DJ, Jensen MS. The effects of individual differences and task difficulty on inattentive blindness. *Psychonomic Bulletin & Review*. 2009;16(2):398-403.
65. Richards A, Hannon EM, Derakshan N. Predicting and manipulating the incidence of inattentive blindness. *Psychological Research*. 2010;74(6):513-23.
66. Richards A, E MH, Vitkovitch M. Distracted by distractors: Eye movements in a dynamic inattentive blindness task. *Consciousness and Cognition*. 2012;21(1):170-6.

67. Seegmiller JK, Watson JM, Strayer DL. Individual Differences in Susceptibility to Inattentive Blindness. *Journal of Experimental Psychology: Learning Memory and Cognition*. 2011;37(3):785-91.
68. Hannon EM, Richards A. Is inattentive blindness related to individual differences in visual working memory capacity or executive control functioning? *Perception*. 2010;39(3):309-19.
69. Bredemeier K, Berenbaum H, Simons DJ. Individual differences in controlled attention and susceptibility to inattentive blindness. *Visual Cognition*. 2011;19(10):1299-303.
70. Weinger MB, Reddy SB, Slagle JM, Editor S. Multiple Measures of Anesthesia Workload During Teaching and Nonteaching Cases. *Anesthesia and Analgesia*. 2004;98(5):1419-25.
71. Cooper JB, Newbower RS, Long CD, McPeck B. Preventable anesthesia mishaps: a study of human factors. *Qual Saf Health Care*. 2002;11(3):277-82.
72. Forster S, Lavie N. High Perceptual Load Makes Everybody Equal: Eliminating Individual Differences in Distractibility With Load. *Psychological Science*. 2007;18(5):377-81.
73. Cheyne JA, Carriere JS, Smilek D. Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition*. 2006;15(3):578-92.
74. Cheyne JA, Solman GJF, Carriere JSA, Smilek D. Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*. 2009;111(1):98-113.
75. Way TJ, Long A, Weihing J, Ritchie R, Jones R, Bush M, et al. Effect of Noise on Auditory Processing in the Operating Room. *J Am Coll Surg*. 2013;216(5):933-8.
76. Singh H, Thomas EJ, Petersen E, Studdert D. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Archives of Internal Medicine*. 2007;167(19):2030-6.
77. Rensink R, O'Regan J, Clark J. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*. 1997;8(5):368-73.
78. Grimes J. On the failure to detect changes in scenes across saccades. In: Akins K, editor. *Perception*. Vancouver Studies in Cognitive Science. 5. New York, USA: Oxford University Press; 1996. p. 89-111.
79. O'Regan J, Rensink R, Clark J. Change-blindness as a result of 'mudsplashes'. *Nature*. 1999;398:34.
80. Jensen MS, Yao R, Street WN, Simons DJ. Change blindness and inattentive blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*. 2011;2(5):529-46.
81. Simons DJ, Chabris CF, Schnur T, Levin DT. Evidence for preserved representations in change blindness. *Consciousness and Cognition*. 2002;11(1):78-97.
82. Simons DJ, Levin DT. Change blindness. *Trends in Cognitive sciences*. 1997;1(7):261-7.
83. Simons DJ, Rensink RA. Change blindness: Past, present, and future. *Trends in Cognitive Sciences*. 2005;9(1):16-20.
84. Simons DJ. In sight, out of mind: when object representations fail. *Psychological Science*. 1996;7(5):301-5.
85. Simons DJ, Franconeri SL, Reimer RL. Change blindness in the absence of a visual disruption. *Perception*. 2000;29:1143-54.
86. Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998;5(4):644-9.
87. Levin DT, Angelone BL. The visual metacognition questionnaire: a measure of intuitions about vision. *American Journal of Psychology*. 2008;121(3):451-72.
88. Levin DT, Momen N, Drivdahl S. Change blindness blindness: the metacognitive error of overestimating change-detection ability. *Visual Cognition*. 2000;7(3):397-412.

89. Hughes HC, Caplovitz GP, Loucks RA, Fendrich R. Attentive and pre-attentive processes in change detection and identification. *PLoS ONE*. 2012;7(8).
90. Vitevitch MS. Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*. 2003;29(2):333-42.
91. Dalton P, Fraenkel N. Gorillas we have missed: Sustained inattentive deafness for dynamic events. *Cognition*. 2012;124(3):367-72.
92. Goldinger SD, Pisoni DB, Logan JS. On the Nature of Talker Variability Effects on Recall of Spoken Word Lists. *Journal of Experimental Psychology: Learning Memory and Cognition*. 1991;17(1):152-62.
93. Mack A, Clarke J. Gist perception requires attention. *Visual Cognition*. 2012;20(3):300-27.
94. Pizzighello S, Bressan P. Auditory attention causes visual inattentive blindness. *Perception*. 2008;37(6):859-66.
95. Sinnett S, Costa A, Soto-Faraco S. Manipulating inattentive blindness within and across sensory modalities. *Quarterly Journal of Experimental Psychology*. 2006;59(8):1425-42.
96. National Transportation Safety Board. Aircraft Accident Report (File 1-0016). Washington D.C.: National Transportation Safety Board; 1973.
97. Lee YC, Lee JD, Boyle LN. Visual attention in driving: The effects of cognitive load and visual disruption. *Hum Factors*. 2007;49(4):721-33.
98. Underwood G. Visual attention and the transition from novice to advanced driver. *Ergonomics*. 2007;50(8):1235-49.
99. Scholl BJ, Noles NS, Pasheva V, Sussman R. Talking on a cellular telephone dramatically increases 'sustained inattentive blindness'. *Journal of Vision*. 2003;3(9):156a.
100. McEvoy SP, Stevenson MR, McCartt AT, Woodward M, Haworth C, Palamara P, et al. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study. *British Medical Journal*. 2005;331.
101. Hyman IE, Boss SM, Wise BM, McKenzie KE, Caggiano JM. Did you see the unicycling clown? Inattentive blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*. 2009;24(5):597-607.
102. Tollner-Burngasser A, Riley MA, Nelson WT. Individual and team susceptibility to change blindness. *Aviation Space and Environmental Medicine*. 2010;81(10):935-43.
103. Andersen JV, Nowak A, Rotundo G, Parrott L, Martinez S. "Price-quakes" shaking the world's stock exchanges. *PLoS ONE*. 2011;6(11).
104. Beck MR, Levin DT, Angelone B. Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition*. 2007;16(1):31-51.
105. Beck MR, Levin DT, Angelone BL. Metacognitive errors in change detection: Lab and life converge. *Consciousness and Cognition*. 2007;16(1):58-62.
106. DiVita J, Obermayer R, Nugent W, Linville JM. Verification of the change blindness phenomenon while managing critical events on a combat information display. *Hum Factors*. 2004;46(2):205-18.
107. Lum TE, Fairbanks RJ, Pennington EC, Zwemer FL. Profiles in patient safety: misplaced femoral line guidewire and multiple failures to detect the foreign body on chest radiography. *Acad Emerg Med*. 2005;12(7):658-62.
108. Gosbee J. Handoffs and communication: the underappreciated roles of situational awareness and inattentive blindness. *Clinical Obstetrics & Gynecology*. 2010;53(3):545-58.
109. Institute for Safe Medication Practice. IMSP Safety Alert: What captures your attention? In: Practice IfSM, editor. Horsham, PA: Institute for Safe Medication Practice; 2009.

110. Anonymous. Fire extinguishing installations and equipment on premises. London: British Standards Institute; 2012.
111. Castel AD, Vendetti M, Holyoak KJ. Fire drill: Inattention blindness and amnesia for the location of fire extinguishers. *Attention Perception & Psychophysics*. 2012;74(7):1391-6.

3 Human Factors in Medical Training

'...a good deal of student learning is not in fact about understanding biology, or political science, or engineering, but about adapting to the requirements of teachers... students will study what they think will be assessed...'

Ramsden (1)

3.1 Chapter overview

The aim of this study is to establish the extent to which non-technical skills and human factors are prioritised and understood by different medical specialties. This will be achieved by an examination of a series of hospital-based postgraduate training curricula.

Training curricula offer a unique and objective insight into each branch of medicine's thinking in this area, because they present a comprehensive list of the items of knowledge and skill that a novice must acquire to be considered qualified in that field.

In the author's experience, both in an educational and clinical capacity, it is by no means clear that all medical professionals understand what is meant by terms such as 'human factors' and 'non-technical skills'. It is crucial therefore to establish these data, in order to appropriately frame the presentation of data from the experiments on situational awareness to professionals from differing professional backgrounds.

3.2 Introduction

The content of assessments has a prime motivating effect on what people learn (2). Non-technical skills (NTS) are a suite of behaviours and skills that facilitate workplace activities, and in anaesthesia are subjected to both formative and summative assessments (3).

Given the prominence of NTS failures in medical errors involving trainees (4), it would be expected that all medical specialties would offer similar teaching and assessment opportunities from an early point in each practitioners' career. It is not at all clear, however, that this is the case in all postgraduate training programmes.

The need to understand patient safety is set out by medical authorities in publications by organisations including the General Medical Council (GMC) and many of the medical Royal Colleges. The GMC publication 'The Trainee Doctor' from 2011 lists patient safety as the first among nine training domains (5), but it does not specify how such training should be structured or delivered. Details regarding the design and provision of postgraduate training are matters devolved to individual specialties through their Colleges' training committees.

These training committees have responsibilities to the profession, to patients, and to the individual trainees to set appropriate standards. Adult learning theories highlight the role of the learner as an active component in their education (6), and characterise the role of the teacher as a guide or facilitator (7). The educator must demonstrate which behaviours to observe, to highlight the interrelations between skills and knowledge (8), and set valid assessments that motivate desirable behaviours in the learner (9). Curriculum developers must therefore specify what a trainee must learn at each stage of training and how the attainment of these objectives should be measured. Adult learners need to appreciate the personal relevance of taught material to make learning effective (8), so trainers must be able to demonstrate to learners why each item is required knowledge.

Modern medical education is designed around competency-based training, with a series of formative and summative assessments based on workplace-based assessments (WPBA) and an annual review of competency progression (ARCP) to review performance. There are also professional examinations conducted by each Royal College that act as gateways between stages of training. This represents a move away from what was historically time-based training, in which

progression through grades was determined by time spent in the programme, towards a system that reflects the skills and abilities acquired at each stage.

All training must be delivered within the working week defined by the European Working Time Directive (EWTD) (10). This has raised concerns regarding curtailment of educational opportunities particularly in highly technical or surgical disciplines, where the reduction in operating hours has been a particular concern (11). The delivery of medical training must therefore be carefully mapped to the curriculum to ensure that every educational encounter has relevance (12). A curriculum document is therefore an essential guide for both educator and learner.

It should not be forgotten that alongside the need for junior doctors to learn, there is also a significant requirement for service-delivery. It is usual in hospital practice for trainee doctors to form the front-line of patient care. The first medical contact many patients have are therefore with clinicians who are amongst the least experienced members of staff. This team, drawn from a variety of speciality backgrounds, may be primarily responsible for care for several hours after admission (13). This may be particularly important in emergency situations where the initial responses are led by doctors in training. The responding team must make a rapid assessment of the patient's condition and reach decisions about whether to continue managing a case themselves or seek the immediate involvement of more experienced colleagues. The non-technical skills of junior staff are not simply an academic question, but also one highly relevant to the patients under their care.

Situational awareness, leadership, and communication abilities are skills practiced daily in all parts of medicine, but there is wide variation in how these skills are understood by different medical specialties. A mixed-methods analysis of medical training curricula was conducted to determine the extent to which different medical specialties set training objectives in NTS, and to seek trends in the prominence with which these skills feature. Qualitative differences in the presentation of such objectives and the associated assessment tools were also sought.

3.3 Methods

3.3.1 Design

This study is an exploratory sequential mixed-methods analysis.

3.3.2 Subjects

Data collection was carried out in mid-2013. The then-current curricula for all adult hospital-based medical, surgical, and critical-care specialties were obtained, along with each curriculum's immediate predecessor, where this was available. Older curricula were included as there were still trainees following these programmes and, although there were no new trainees joining the older systems, the documents would not be fully retired until the last trainee had either gained their CCT or migrated to a current version.

Medical specialty curricula are available from the Joint Royal Colleges of Physicians Training Board (www.jrcptb.org.uk) and surgical curricula from the Intercollegiate Surgical Curriculum Programme (www.iscp.ac.uk). Other specialty curricula were accessed directly from the website of the appropriate College. The full list of curricula and versions analysed are summarised in Table 3-1.

3.3.3 Outcome Measures

The primary outcome of this study was a quantitative analysis of learning objective clustering around pre-specified human factors learning themes. This enabled comparison to be drawn between specialty curricula based upon the frequency and pattern of content loading.

3.3.4 Study Procedures

To ensure adequate capture of relevant learning objectives, a series of search terms relevant to NTS were generated using a modified Delphi process (14). A list of 20 keywords derived from the ANTS (15), NOTSS (16), and NOTECHS (17) non-technical skills systems was produced by the author, with terms grouped under headings of 'task management', 'team working', 'situational awareness', and 'decision making'. In the first round this list was circulated amongst six experts with extensive

experience in human-factors training. These trainers came from anaesthetic, intensive care, or acute medical specialties, and all served as faculty with the Oxford Centre for Simulation, Teaching, and Research (OxSTaR). The participants were invited to add or delete terms as appropriate, and supply their reasoning for each change.

To be marked for deletion, two participants had to flag a term for removal from the list, but terms were added to the list after only a single recommendation. This strategy was chosen to ensure as wide as possible a list was available for searching.

In the first round, none of the original 20 keywords were deleted and 12 new words were suggested. The revised list was circulated amongst the panel, and again no terms from the expanded list were deleted. This time no new terms were added, thus as consensus had been reached, the planned third round was not undertaken.

The contents of the final version were grouped under the headings 'task management', 'teamworking', 'situational awareness', and 'decision making'. These keywords were then transformed into search terms suitable to be used in functions within the analysis software (see section 3.3.5). The final list of keywords and their related search terms is summarised in Table 3-2.

List of analysed curricula by specialty group

Field	Specialty	Version(s) Analysed
Medicine	Acute Medicine	2009 (amended 2012)
	Allergy	2007 [†] , 2010
	Audiology	2007, 2010
	Cardiology	2007, 2010
	Clinical Genetics	2007 [†] , 2010 [†] (amended 2012)
	Clinical Neurophysiology	2007 [†] , 2010 [†] (amended 2012)
	Core Medical Training	2010
	Clinical Pharmacology	2010
	Dermatology	2007, 2010 (amended 2012)
	Endocrinology & Diabetes	2007, 2010 (amended 2012)
	Gastroenterology	2007, 2010
	General Medicine	2007, 2009 (amended 2012)
	Genito-Urinary Medicine	2007, 2010 (amended 2012)
	Geriatrics	2007, 2010 (amended 2012)
	Haematology	2007 [†] , 2010 (amended 2012)
	Immunology	2007, 2010
	Infectious Disease	2007 [†] , 2010 [†] (amended 2012)
	Medical Oncology	2007, 2010
	Metabolic Medicine	2007 [†] , 2010
	Neurology	2007 [†] , 2010 [†]
	Nuclear Medicine	2007 [†] , 2010
	Ophthalmology	2007, 2010
	Paediatric Cardiology	2007, 2010 (amended 2012)
	Palliative Medicine	2007, 2010 (amended 2012)
	Pharmaceutical Medicine	2007, 2010 (amended 2011)
	Rehabilitation Medicine	2007 [†] , 2010 [†]
	Renal Medicine	2007 [†] , 2010 (amended 2012)
	Respiratory Medicine	2007 [†] , 2010
	Rheumatology	2007 [†] , 2010
	Sports and Exercise Medicine	2007, 2010
Stroke Medicine	2007 [†] , 2010	
Critical Care	Anaesthetics	2007, 2010
	Emergency Medicine	2007, 2010
	Intensive Care Medicine	2007, 2010 (including Dual-CCT curricula)
Surgery	Cardiothoracic Surgery	2007, 2010
	Core Surgical Training	2010
	Oto-laryngeal Surgery	2007, 2010
	General Surgery	2007, 2010
	Oro-maxillo-facial Surgery	2007, 2010
	Neurosurgery	2007, 2010
	Obstetrics and Gynaecology	2007 (amended 2011, 2012)
	Paediatric Surgery	2007, 2010
	Plastic Surgery	2007, 2010
	Trauma and Orthopaedics	2007, 2010
	Urology	2007, 2010
	Vascular Surgery	2013

[†]denotes a curriculum with no relevant instances of a keyword

Table 3-1: List of curriculum documents

Keywords and NVIVO search terms		
Category	Stem Keyword	Search Term (including Boolean 'wild card')
Task Management	Prioritisation	Priorit*
	Resource utilisation	Resource*
	Role model	Role [AND] model*
	Standards	Standard*
Team Working	Assertiveness	Assert*
	Authority	Authorit*
	Collaboration	Collaborat*
	Communication	Communicat*
	Leadership	Leadership
	Supportiveness	Support*
	Team-working	Team*
Situational Awareness	Anticipate	Anticipat*
	Attention	Atten*
	Detect	Detect*
	Deterioration	Deteriorat*
	Monitor	Monitor*
	Observation	Observ*
	Recognise	Recogn* [NEAR] signs [OR] symptoms [OR] complications [OR] features
	Situational Awareness	Situation* [AND] Awareness
	Vigilance	Vigilan*
Watch	Watch*	
Decision Making	Balance	Balance
	Choose	Choose
	Decide	Decide
	Decision	Decision
	Option	Option*
	Plan	Plan*
	Re-evaluate	Re-evaluat*
	Select	Select*
Core Terms	Human Factors	Human [AND] factor*
	Non-technical skill	Non-technical [OR] NTS
	Situational Awareness	Situation* [AND] Awareness
	Simulation	Simulat*

Table 3-2: Keywords generated by the Delphi process, and their related search terms

3.3.5 Analysis

Each curriculum was analysed using NVIVO 10 (QSR International, Warrington, UK). To ensure coding validity a sample of curricula was chosen using a random number generator. The sample comprised two curricula from each of the three specialty areas (medicine, surgery, and critical care). Two

investigators, the author and a qualitative researcher with an interest in education (and who did not participate in the Delphi process), independently examined the sample curricula and assessed each instance of a keyword for relevance to non-technical skills. The results were subjected to a test-of-agreement which yielded a high degree of agreement (98%; Kappa = 0.814), and it was therefore felt to be appropriate to single-code the remaining documents.

3.4 Results

Curricula for 31 medical specialties, 12 surgical specialties, and three critical care specialties were reviewed. These represented approximately two million words spread over 88 documents.

NTS terms occurred infrequently across most of the examined texts, with most occurrences in anaesthesia, emergency medicine (EM), or intensive care medicine (ICM). Only these three critical-care specialties specify requirements for formal training in NTS. Methods recommended for training included self-study, tutorials, and simulation. In these curricula NTS are generally described as both knowledge suitable for formal assessment, and also as behaviours requiring direct observation and critique (3).

No comparable training or assessment requirements exist in any of the curricula for medical and surgical specialties. Although well placed to train and assess NTS, simulation is seldom mentioned outside critical-care specialties and almost exclusively recommended only for procedural skills training or resuscitation. NTS-related objectives, where specified, are generally considered knowledge (as opposed to practical skills), and assessed accordingly using methods such as case-based discussion.

Several medical curricula contained no relevant instances of any keywords. While these curricula inferred aspects of NTS, they lacked specific detail on learning objectives and assessment recommendations. For example, communication was considered only in the context of doctor-patient relationships, language skills, or record-keeping. Safety was considered in terms of

adherence to guidelines, evidence-based practice, and continuing professional development. Understanding some components of team-working is expected by many of these curricula, but it was generally unclear what the curriculum-authors consider these components to be, or how they should be taught. These data are summarised in Table 3-3.

Instances and examples of core non-technical skills keywords					Assessment Tool
HF	NTS	SA	Sample text from 2010 curricula		
Anaesthetics	14	28	12	[Demonstrate] "Awareness of human factors... and the importance of non-technical skills in achieving consistently high performance such as: effective communication, team-working, leadership, decision-making and maintenance of high situation awareness" [A27 Critical Incidents/Attitudes and Behaviours] [Describe] "The central role human factors plays in developing a culture of safe practice and how collaboration and team-working enhances safety" [Advanced Training (Domain 1: Clinical Practice)/Knowledge and Skills] [The trainee] "Ensures the primacy of patient safety in all aspects of communication and cooperation and is able to utilise cognitive strategies, human factors and CRM to maximise this" – CCL5 Communication with colleagues and cooperation/Behaviour in Emergency Department context [Demonstrate] "Awareness of human factors concepts and terminology and the importance of non-technical skills in achieving consistently high performance such as: effective communication, team-working, leadership, decision-making and maintenance of high situation awareness" – [C Critical Incidents/Attitudes and Behaviours] "Demonstrate good non-technical skills such as: effective communication, team working, leadership, decision making and maintenance of high situational awareness" [C Critical Incidents/Skills] "Outline human factors theory and understand its impact on safety" [2.7 Prioritisation of patient safety in clinical practice/Knowledge & 11.3 Identifies hazards and promotes safety for patients and staff/Knowledge] [Demonstrate] "Awareness of human factors concepts and terminology and the importance of non-technical skills in achieving consistently high performance such as: effective communication, team-working, leadership, decision-making and maintenance of high situation awareness" [3.10 Critical Incidents/Attitudes and Behaviours]	Cbd, CEX, Examination, Simulation ALMAT, MSF ACAT, Cbd, CEX, MSF ACAT, Cbd, Examinations ACAT, Cbd, DOPS, Simulation Cbd CEX, Cbd, Examinations, Simulation
Emergency Medicine	3	4	4	"Outlines human factors theory and understands its impact on safety" [1.7 Prioritisation of patient safety in clinical practice/Knowledge]	Cbd
Intensive Care Medicine	3	4	2	"Outlines human factors theory and understands its impact on safety" [1.7 Prioritisation of patient safety in clinical practice/Knowledge]	Cbd
Allergy	1	0	0	"Be able to describe... the human factors which lead to drug use errors" [36 Drug Errors/Knowledge]	Pbd
Clinical Pharmacology	1	0	0	"Situation awareness: capacity to monitor and anticipate situations that may change rapidly" [Selection criteria/Personal Skills]	Application, interview, references
Surgical Specialties ¹	0	0	1	"Some surgical specialities (cardiothoracics, neurosurgery, oro-maxillo-facial surgery, plastics, paediatric surgery, trauma & orthopaedics, and urology) each contained this single instance of 'situational awareness'.	
ACAT	Acute care assessment tool			Direct observation of a trainee over an extended period, commenting on management, reasoning, interpretation of signs, team-working, and resource management.	
ALMAT	Anaesthesia List Management Assessment Tool			Direct observation of trainee's conduct over a whole session, including management, record keeping, interpretation of signs, team-working, leadership, and communication.	
Cbd	Case-based discussion			A post-hoc assessment of case management intended to demonstrate clinical reasoning, decision making, and medical knowledge.	
CEX	Clinical evaluation exercise			Direct observation of a trainee over shorter periods. Feedback covers assessment, diagnostic skill, clinical reasoning and decision-making, safety, communication, and professionalism.	
DOPS	Direct Observation of Procedural Skills			An assessment of technical skill performance; in some specialities (notably emergency medicine) reference is also made to NTS components such as SA.	
MSF	Multi-source feedback			Collection of feedback from a trainees peers from various professional backgrounds, covering knowledge, skills, performance, contribution to safety and quality, communication and team-working, and probity.	
Pbd	Problem-based discussion			A post-hoc assessment of knowledge demonstrating reasoning, decision-making and application of knowledge in relation to drug treatment usually at a population level	

Table 3-3: Count of occurrences of NTS keywords by specialty and year

3.4.1 Human factors

With one exception (Emergency Medicine, 2007) human factors were not specifically mentioned in older versions of any specialty curricula. The single mention in the 2007 Emergency Medicine curriculum describes a knowledge skill related to maintaining Good Clinical Care. In the then-current Emergency Medicine curricula human factors features more prominently, associated with high-level performance in the domain '*CC15: Communication with colleagues and cooperation*'. Here human factors and resource management skills are aligned with ensuring '*the primacy of patient safety in all aspects of communication and co-operation*'.

The Emergency Medicine Curriculum also incorporates non-technical skills in its formative assessment programme. They are included in the rating scales for the evaluation of practical skills using tools such as the Case-Based Discussion (CbD) and Mini-Clinical Evaluation Exercise (MiniCEX). The assessment of NTS is therefore situated within a realistic and relevant context. Assessment of NTS is also included in summative examinations for membership of the College of Emergency Medicine (CEM).

The Anaesthetics curriculum is unique in that it devotes an entire chapter to human factors in clinical practice. There is an overt link made between NTS training and patient safety, with trainees required to demonstrate '*clear understanding of why effective [NTS] are required... to ensure clinical sessions are delivered safely*' (Advanced Training: Clinical Practice [Objective AT_D1_01]). Objectives in sub-specialty areas express similar needs. These provide specific guidance on the level of knowledge trainees are expected to possess regarding human factors, noting they must '*demonstrate mastery... and clear understanding of the central role human factors plays in developing a culture of safe practice and how collaboration and team working enhances safety*'. Many of these objectives are assessed using the Anaesthesia List Management Assessment Tool (ALMAT) and review of Multi-source Feedback (MSF), which the College describes thusly:

'The ALMAT... look[s] formally at the way these abilities are put together to undertake safe and effective day-to-day clinical practice. The MSF... [is] particularly useful in terms of gathering information about the trainee's professional demeanour, communication skills and ability to cooperate within a team'

Royal College of Anaesthetists 2010(3)

Critical care specialities expressly stipulate a formal educational requirement for human factors, exemplified in anaesthetic objective TM_IS_09: *'[the trainee should] participate in human factors and patient safety training'*. It is recognised that such training will be delivered using a variety of methods including self-study, formal tutorials, and simulation. Conversely, human factors are not referenced in medical curricula until 2010, where they are referenced once each in the learning objectives of only three programmes:

'Outlines human factors theory and understands its impact on safety' – Allergy 2010

'Be able to describe... the human factors which lead to drug use errors' – Clinical Pharmacology 2010

'Outlines human factors theory and understands its impact on safety' – Immunology 2010

In each of these instances the learning objectives appear in the context of prioritising patient safety and in every case the objective is defined as knowledge, assessed by tools such as CbD, or in some cases the Problem-based Discussion (PbD). There is no mention of a practical skills assessment using a tool such as direct observation of procedural skill (DOPS) or MSF.

The search term 'human factors' did not appear in any surgical curricula.

3.4.2 Non-technical skills

Non-technical skills were explicitly mentioned only in the curricula of critical care specialties. The term appears most frequently in Anaesthetics documents where the need for structured, formal,

non-technical skill training featured prominently. Many objectives in this domain were identically worded and numbered between the three critical-care curricula.

NTS assessment is also uniquely incorporated into formal examinations set by the RCoA:

'The Primary FRCA Examination... assesses the knowledge and understanding [and some of the attitudes and non-technical skills] required to inform and underpin the clinical practice of basic level training.'

Royal College of Anaesthetists(3)

3.4.3 Situational awareness

In common with the other search terms situational awareness occurred most frequently in critical care specialties. It was, however, the only core phrase that occurred in any surgical curricula.

There was a single instance of situational awareness in each of the current curricula for cardiothoracic surgery, oro-maxillo-facial surgery, otolaryngeal surgery, plastics, trauma, urology and neurosurgery. The occurrences in every case were as a *'personal skill'* concerning *'capacity to monitor and anticipate situations that may change rapidly'*. The skill was listed as an *'essential selection criteria for application to surgical posts'*, measured by the application form, interview/selection centre and by references, but without obvious practical or continuing assessment.

In only the 2007 edition of the general surgery curriculum was situational awareness considered an essential knowledge skill, listed under the domains of *'managing clinical care'* and *'managing a team'*, where the trainee must understand *'situational awareness and the causes of loss of situational awareness'*. Surprisingly then, situational awareness was not specifically mentioned in the more recent version of the curriculum at all.

Regarding situational awareness in critical-care specialties, the CEM afford a high priority identifying it as a standard of performance on the CEM-DOPS form. The RCoA similarly mandates the trainee

should *'commit to the principle of maintaining situational awareness at all times'* (objective CC_D1_04), as well as be able to *'demonstrate appropriate situational awareness'* during high risk periods (objective TF_IS_09).

Anaesthetic objectives AT_D2_10 (*'[the trainee should recognise] the importance of maintaining high levels of individual and team situation awareness at all times'*) and TF_AS_36 (*'Demonstrates situational awareness and contributes to the maintenance of situational awareness in others'*) are amongst a very limited set of objectives that recognise the limited ability for individuals to maintain good quality situational awareness alone, and place a value on awareness in the context of team performance.

There were instances where situational awareness was described without using the specific terminology. For example: *'anticipating and managing risks to personal safety'* is specifically mentioned in the 2010 curriculum for Palliative Care 2010. Similarly, Metabolic Medicine 2010 states that a trainee should be able to *'Recognise and respond to the manifestations of a patient's deterioration or lack of improvement... and support other members of the team to act similarly'*. Although not explicitly stated, these objectives cannot be fulfilled without the application of effective situational awareness.

3.4.4 Human factors training

An explicit requirement for human factors training was specified only in critical-care specialties, and there are commonalities in the learning objectives between each document. Learning objectives specify integrated multidisciplinary team training, as illustrated by objective CL_BK_34: *'Discusses the importance of understanding the need... To practice response protocols in resuscitation room or in simulation with other healthcare professionals as appropriate'*.

Implicit references to the development of quality non-technical skills could be found in most curricula, as exemplified in the 2007 General Surgical curriculum: *'displays attitudes commonly accepted as essential to professionalism.'* This statement, taken from a section of the document

dealing with reliability, continuing professional development and self-evaluation, is open to interpretation as there is likely to be broad opinions on what constitutes professional behaviour within and between each specialty. Simulation training was not mentioned at all in any curriculum belonging to a surgical specialty prior to 2012. It is mentioned briefly in the 2012 vascular surgery curriculum in the context of technical skills training, but only in the 2012 otolaryngeal surgery curriculum is there mention made of team-training:

'The role of simulation in [otolaryngeal surgery] is still being considered but... it should be incorporated into the curriculum with both technical skills... and more integrated team skills including full theatre simulation.'

Where simulation training is mentioned at all in medical specialties (cardiology, neurophysiology, dermatology, general/internal medicine, oncology, ophthalmology, neurology, palliative medicine, pharmaceutical medicine, rheumatology, and sports and exercise medicine), it is almost exclusively considered in the context of procedural skills training or resuscitation.

The General (Internal) Medicine curriculum 2010 includes a domain entitled *'Team Working and Patient Safety'* which highlights a need for effective collaboration. This lists amongst knowledge skills the need for the trainee to *'outline the factors adversely affecting a doctor's performance and methods to rectify these'*. Qualities that the trainee should display include the ability to deliver a *'detailed handover between shifts and areas of care'*, and (listed under *'Communication with colleagues and co-operation'*) to have knowledge of *'the features of good team dynamics'* and the *'the principles of effective inter-professional collaboration to optimise patient or population care'*. There is a clear desire to instil in the trainee an appreciation of the value of close and effective team working, but the focus appears primarily upon multi-disciplinary team working and team-leadership, without significant reference to situational awareness or group dynamics.

3.4.4.1 Curricula with no instances of relevant keywords

Five specialties, all medical, published then-current curricula in which there were no relevant instances of any keywords in any version. These data are summarised in Table 3-4.

Curricula with no relevant instances of NTS keywords		
Field	2007 Curriculum	Current Curriculum
Medicine	Allergy	Clinical Genetics
	Clinical Genetics	Clinical Neurophysiology
	Clinical Neurophysiology	Infectious Disease [†]
	Haematology	Neurology
	Infectious Disease [†]	Rehabilitation Medicine
	Metabolic Medicine	
	Neurology	
	Nuclear Medicine	
	Rehabilitation Medicine	
	Renal Medicine [†]	
	Respiratory Medicine [†]	
	Rheumatology	
	Stroke Medicine	

[†]denotes an acute medical specialty, as defined by the Royal College of Physicians (18)

Table 3-4: Curriculum documents that had no relevant instances of NTS keywords

These curricula do have domains dealing with aspects of non-technical skills but they lack the comprehensive approach seen in the curricula of other specialties. The 2010 Clinical Genetics curriculum, for example, discusses communication but primarily addresses communication between doctor and patient or family. Similarly, the 2010 Neurophysiology curriculum section on communication discusses the importance of *'communication in the English language, verbally and in writing'*, and the need for the doctor to *'produce legible case notes'*.

Objectives dealing with professional behaviour in many curricula tend to consider matters such as continuing professional development, stress, conflict resolution or work-life balance. There are sections expressly detailing patient safety in some, but these generally reflect a somewhat narrower focus than found in critical care curricula. The 2010 Clinical Genetics section dealing with safety highlights the *'importance of evidence-based practice in relation to clinical effectiveness'*, and the need for the trainee to *'outline local health and safety protocols'*. Knowledge skills in this section do

require the trainee to understand '*...the components of effective collaboration and team working*', but do not set out what these components are, nor how they should be taught and assessed.

The 2010 Infectious Disease Curriculum contains a domain on '*interaction with other healthcare teams*' which deals with conflict resolution and communicating '*...effectively and respectfully with other relevant professionals by means appropriate to the urgency of the situation*'. The patient safety domain of this curriculum (section 1.7) mentions health and safety considerations, avoiding drug side effects, and that the trainee must '*be aware of [their] own limitations, and operate within them competently*'. Again, the communication domain is largely focussed on handling doctor-patient communication rather than between professionals.

The 2010 Rehabilitation Medicine contains a section addressing Safety and Quality, but this addresses issues of clinical governance, service evaluation, and audit, rather than the performance of staff and teams. Communication skills in rehabilitation medicine are, in common with the other specialities above, focussed on communication with patients. Inter-professional communication is only addressed at the level of care planning in multi-disciplinary team meeting settings.

3.5 Discussion

It is reasonable to contend that if the content of assessment drives learning, a paucity of relevant, specific, and measurable learning objectives will lead to poorer understanding of NTS amongst trainees. There is evidence that NTS training impacts patient safety in a wide range of clinical areas (19-21), but explicit NTS learning objectives feature rarely outside critical-care specialties. Although at first glance human factors seem primarily of interest to practitioners dealing with the most acute situations, the principles are widely applicable and still have value in less urgent settings. Situational awareness, for example, is required in any circumstances in which a decision must be made; from determining the appropriate treatment plan for an individual patient, through managing issues of personal security, to making strategic decisions on behalf of a whole department (22). Team-working and leadership are expected of all medical trainees (23), and prioritisation and the adequate

deployment of resources matter whether managing a single resuscitation team or a whole hospital. Furthermore, industrial data show clearly that human errors contribute to critical incidents in a breadth of circumstances (24). Although error frequency is higher when time and workload pressures increase (25), significant incidents do still occur at lower operating tempos (26).

The principles embodied by human factors may therefore be widely employed, but the scarcity of the specific terms implies that they receive little formal recognition outside of critical care. It seems likely that most curricula are drafted assuming NTS can be acquired implicitly, for example through the observation of suitable role models, rather than recognising a need for formalised training and assessment. This informal learning is often termed the 'hidden curriculum' (27), and the trainers themselves may not even be aware of everything they are teaching through this process. Self-assessment of NTS quality is unreliable (28), and the role models available may not provide best-practice examples for the trainee to follow. The acquisition of learning is therefore likely to be inconsistent, haphazard, and unpredictable. A move to explicit and standardised NTS teaching and assessment in all curricula is desirable.

To assist in this process there is a need for validated assessment tools to enable both trainee and assessor to objectively determine acceptable standards. The availability of a validated assessment tool, in the form of ANTS (15), greatly facilitates the delivery of NTS training in anaesthesia. The tool was developed over several years and has been specifically tailored to anaesthetic practice. Similar systems exist, or are being developed, in surgery (16), emergency medicine (29), and for non-medical theatre practitioners (30), but these have not yet been integrated into the relevant curricula (31, 32). There is an urgent need to develop an NTS system in general medicine, where the formal recognition of NTS within curricula remains particularly poor.

Specifying NTS objectives are an important first step, but this alone will not be sufficient to improve trainees' acquisition of these skills. Validated summative and formative assessments are required to ensure appropriate achievement. NTS are fundamentally a system of behaviours and traits, thus it is

unlikely that instruments designed to assess knowledge are appropriate for measuring them. Assessments based on written materials or retrospective case reviews lack face- and construct-validity for a purpose such as this (33).

To adequately determine NTS performance, assessments must therefore include an observational component ideally within the context of a clinical task. The model adopted by the three critical care specialties sees NTS learning objectives appear alongside relevant core technical skills within the main body of the curriculum. These skills are assessed in realistic contexts, with NTS items appearing on assessment forms intended for the real-time evaluation of procedural skills or list-management (DOPS, ALMAT).

Much of medical practice is social in nature (34), based on human interactions between doctor and patient, or between doctor and professional colleagues. Furthermore, active participation in learning in context, so-called 'situated learning' is likely to facilitate the adoption of professional roles in medicine (35). It is therefore disappointing that many curricula in medical and surgical specialties are structured in a manner that separates these social skills from their clinical application. Most surgical and medical documents present objectives as groups of symptom- or system-based competencies, focussing on the recognition of specified conditions or presentations and their technical management. They are grounded in a symptom-recognition-treatment model that devotes little overt attention to non-technical skills. Communication and team working are, if present at all, frequently relegated to a separate 'professional skills' domain. Non-technical skills are, by their very nature, integrated into all aspects of healthcare delivery, thus it does not seem justifiable to teach or assess them outside of a clinical context.

Many specialty programmes are recognising these unmet needs at a local level (36), however teaching and assessment varies within and between institutions. A fundamental feature of multi-disciplinary teams working in high-reliability organisations is that team members observe each other's practice and offer feedback (37). This requires each member of the team, regardless of

professional background, to have a shared understanding of NTS (3). Although existing NTS scoring systems broadly overlap (elements of situational awareness, decision-making, communication, team-working, task management and leadership are usually present) behavioural descriptors and assessment focus can be inconsistent between scales (38). Variability in training methods means that the body of literature and educational materials generated by one specialty would at best be opaque, and at worst inaccessible, to another. This presents significant barriers to delivering multi-disciplinary team-training.

3.5.1 Summary

The philosophy underpinning most medical and surgical specialties appears to be that NTS are acquired implicitly, for example through the observation of suitable role models. NTS failures are frequently either a causative or contributing factors in many industrial incidents (39), but it is clear that the appreciation of the topic is very limited in healthcare. Aircrew must undergo specific training in the development and maintenance of NTS, but with the very few exceptions found in the critical care specialties there is no equivalent requirement in healthcare.

Of relevance to the NHS generally, this work highlights issues of poor standardisation in terminology and training. It suggests a need to move to explicit, standardised training in NTS across all medical specialties, and consistency in assessment. Many specialty programmes are recognising this need at a local level, however teaching practices and assessment is not standardised. This will impede communication between educators from each branch of practice. Improving the quality of NTS amongst healthcare staff would be expected to improve quality of care and patient safety, and would represent an additional step by the NHS towards becoming a high-reliability organisation.

Within the context of this thesis these data are also important. This work has established that human factors are inconsistently taught and assessed. It seems likely that trainees, at least those outside of critical care specialities, will have little understanding of terminology such as 'situational

awareness'. It does not seem implausible to go further and suggest that their understanding of psychological concepts such as perceptual error will be even more limited.

This means that, should perceptual error be detectable in clinical staff, when materials are developed to inform and educate the workforce any such resources will have to be written with an assumption that human factors will need some introduction.

Self-assessment of non-technical skill is inaccurate, and insight into effects of perceptual error is poor. Clinicians are not, in the author's experience, used to thinking of themselves as vulnerable. It is likely therefore that when presenting data arising from later experiments, audiences will need to be convinced by a comprehensive argument. The personal relevance to the clinician will need to be explored, and the context of the relevant psychological theory explained. Similar issues are likely to arise when presenting such data to policy-makers and stakeholders. These data validate the incremental approach to experimentation that subsequent chapters will describe.

3.6 References

1. Ramsden P. Learning to teach in higher education. London: Routledge; 1992.
2. Epstein RM. Assessment in Medical Education. *N Engl J Med.* 2007;356(4):387-96.
3. Curriculum Development Working Group. Curriculum for a CCT in Anaesthetics. 2nd ed. London: The Royal College of Anaesthetists; 2010.
4. Singh H, Thomas EJ, Petersen E, Studdert D. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Archives of Internal Medicine.* 2007;167(19):2030-6.
5. General Medical Council. The Trainee Doctor: Foundation, Specialty, including GP Training. London: General Medical Council; 2011.
6. Kolb DA, Fry R. Toward an applied theory of experiential learning. In: Cooper C, editor. *Theories of Group Process.* London: John Wiley; 1975.
7. Kaufman DM. Applying educational theory in practice. *BMJ.* 2003;326(7382):213-6.
8. Maudsley G, Strivens J. Promoting professional knowledge, experiential learning and critical thinking for medical students. *Med Educ.* 2000;34(7):535-44.
9. Assessment Working Group. Developing and maintaining an assessment system - a PMETB guide to good practice. Davies H, Joshi H, Holsgrove G, Rowley D, editors. London: Postgraduate Medical Education and Training Board; 2007.
10. European Working Time Directive Working Group. A guide to the implications of the European Working Time Directive for doctors in training. London: EWTD Reference Group; 2009.
11. Postgraduate Medical Education Training Board. PMETB and the European Working Time Directive; September 2005 to September 2009. London: General Medical Council; 2009.
12. Tasker F, Newbery N, Burr B, Goddard AF. Survey of core medical trainees in the United Kingdom 2013 – inconsistencies in training experience and competing with service demands. *Clinical Medicine.* 2014;14(2):149-56.
13. Stein A, MacIntyre A, Ingham-Clark C, Hopson C, Welsh C, Dalton D, et al. NHS Services, seven days a week forum: Summary of Initial Findings. London: NHS England; 2013.
14. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ : British Medical Journal.* 1995;311(7001):376-80.
15. Flin R, Maran N, Glavin R, Patey R. Framework for Observing and Rating Anaesthetists' Non-Technical Skills. Aberdeen: University of Aberdeen; 2003.
16. Crossley J, Marriott J, Purdie H, Beard JD. Prospective observational study to evaluate NOTSS (Non-Technical Skills for Surgeons) for assessing trainees' non-technical performance in the operating theatre. *Br J Surg.* 2011;98(7):1010-20.
17. Misra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. *Quality and Safety in Health Care.* 2009;18(2):104-8.
18. Federation of the Royal Colleges of Physicians. The Physician of Tomorrow: Curriculum for General Internal Medicine (Acute Medicine) Competence Levels 1 and 2. London: Royal College of Physicians; 2006.
19. Starmer AJ, Spector ND, Srivastava R, West DC, Rosenbluth G, Allen AD, et al. Changes in Medical Errors after Implementation of a Handoff Program. *N Engl J Med.* 2014;371(19):1803-12.
20. Neily J, Mills PD, Young-Xu Y, et al. Association between implementation of a medical team training program and surgical mortality. *JAMA.* 2010;304(15):1693-700.
21. Arriaga AF, Bader AM, Wong JM, Lipsitz SR, Berry WR, Ziewacz JE, et al. Simulation-Based Trial of Surgical-Crisis Checklists. *N Engl J Med.* 2013;368(3):246-53.
22. Endsley M, Garland D. Theoretical underpinnings of situation awareness: a critical review. *Situational awareness: analysis and measurement.* Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
23. General Medical Council. Good Medical Practice. London: General Medical Council; 2013.

24. Flin R, Maran N. Identifying and training non-technical skills for teams in acute medicine. *Qual Saf Health Care*. 2004;13 Suppl 1:i80-4.
25. van Beuzekom M, Boer F, Akerboom S, Hudson P. Patient safety: latent risk factors. *Br J Anaesth*. 2010;105(1):52-9.
26. Harmer M. Independent Review on the care given to Mrs Elaine Bromiley on 29 March 2005 London: Clinical Human Factors Group; 2005 [Available from: <http://www.chfg.org/articles-films-guides/elaine-bromiley-report>].
27. Giroux H, Penna A. Social education in the classroom: the dynamics of the hidden curriculum. In: Giroux H, Purpel D, editors. *The Hidden Curriculum and Moral Education*. Berkely, CA: McCutchan Publishing Corporation; 1983.
28. Arora S, Miskovic D, Hull L, Moorthy K, Aggarwal R, Johannsson H, et al. Self vs expert assessment of technical and non-technical skills in high fidelity simulation. *Am J Surg*. 2011;202(4):500-6.
29. Flowerdew L, Gaunt A, Spedding J, Bhargava A, Brown R, Vincent C, et al. A multicentre observational study to evaluate a new tool to assess emergency physicians' non-technical skills. *Emerg Med J*. 2013;30(6):437-43.
30. Rutherford JS, Flin R, Irwin A, McFadyen AK. Evaluation of the prototype Anaesthetic Non-technical Skills for Anaesthetic Practitioners (ANTS-AP) system: a behavioural rating system to assess the non-technical skills used by staff assisting the anaesthetist. *Anaesthesia*. 2015;70(8):907-14.
31. Greig PR, Higham H, Vaux E. Lack of standardisation between specialties for human factors content in postgraduate training: an analysis of specialty curricula in the UK. *BMJ quality & safety*. 2015;24:558-60.
32. Greig PR, Higham H, Vaux E. Authors response to "Human factors training in post graduate curricula" [In press]. *BMJ Quality and Safety*. 2015.
33. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgical Endoscopy And Other Interventional Techniques*. 2003;17(10):1525-9.
34. Lave J, Wenger E. *Experiential learning: legitimate peripheral participation*. Cambridge: Cambridge University Press; 1991.
35. Mann KV. Theoretical perspectives in medical education: past experience and future possibilities. *Med Educ*. 2011;45(1):60-8.
36. Youngson G, Paterson-Brown S, Russ S. Response to: 'Lack of standardisation between specialties for human factors content in postgraduate surgical training: an analysis of specialty curricula in the UK' by Greig et al. *BMJ quality & safety*. 2015.
37. Baker DP, Day R, Salas E. Teamwork as an Essential Component of High-Reliability Organizations. *Health Serv Res*. 2006;41(4 Pt 2):1576-98.
38. Dietz AS, Pronovost PJ, Benson KN, Mendez-Tellez PA, Dwyer C, Wyskiel R, et al. A systematic review of behavioural marker systems in healthcare: what do we know about their attributes, validity and application? *BMJ quality & safety*. 2014.
39. Walker GH, Stanton NA, Kazi TA, Salmon PM, Jenkins DP. Does advanced driver training improve situational awareness? *Appl Ergon*. 2009;40(4):678-87.

4 From human factors to ergonomics: the impact of equipment design

'...the enhancement of operator situational awareness has become a major design goal for those developing operator interfaces... in a wide variety of fields...'

Endsley (1)

4.1 Chapter overview

The previous chapter established that the understanding of human factors amongst medical professionals is not universal. This variance has clinical consequences, but also has less obvious implications for other decisions made by doctors.

Clinicians in specialties such as anaesthetics have input into the acquisition and provision of equipment in their workplace. The author has heard speculation that these decisions are based largely on commercial or financial considerations. The limited understanding that most clinicians are expected to gain about human factors in training means that sound ergonomic principles are not necessarily prime motivators when equipping the workplace.

The psychology of perceptual errors described in Chapter 2 demonstrates that mental workload contributes to risk of error, and it could therefore be suggested that poorly designed or unfamiliar equipment will increase vulnerability.

The aim of the experiment presented here is therefore to determine the impact of a clinician's familiarity with an item of monitoring equipment on their performance in tasks deriving information from that display. It is hypothesised that there is a clear relationship between familiarity and performance, and the experiment has been designed to explore this in a straightforward manner. A

secondary objective is to explore methods of assessing display complexity, and to seek associations between complexity and performance. A tertiary objective is the opportunity for the author to gain familiarity in the use and of eye-tracking equipment and analysis of the data generated by the system.

4.2 Introduction

Situational awareness describes the acquisition of information, using data to formulate mental models, then employing those models to make decisions or predictions (1). Good situational awareness is safety-critical (2), and is of particular relevance to acute-care specialties such as anaesthesia. The clinical monitor is one of the most frequently checked sources of information in theatre (3), and is therefore of crucial importance in developing and maintaining situational awareness.

Monitoring provides early warning of deterioration, ideally at a stage where intervention can prevent significant harm coming to the patient. The Association of Anaesthetists of Great Britain and Ireland (AAGBI) has specified minimum monitoring standards (4), which stipulates the nature and frequency of observations required during an anaesthetic procedure. A high proportion of errors in critical care occur because of human failings related to monitoring (5), and change blindness is associated with failures to detect events on monitor displays (6).

Competence in the use of monitors is a mandatory requirement of basic-level training in anaesthesia (7). Trainees are required to demonstrate that they can correctly apply monitoring to the patient, manage settings and alarms appropriately, interpret, and then record the measured parameters. This also requires the anaesthetist to differentiate real change from artefact.

Given the importance of monitoring devices, one might assume that their provision would be subject to rigorous regulation. By analogy, the design of an aircraft flight deck is strictly controlled, and its crew must be certified to operate each particular type of aircraft (8). Changes in how

information is generated and displayed between aircraft types was one of the contributing factors to the Kegworth air disaster (9), as flight crew fatally misinterpreted the data being presented to them.

No such equivalent exists in anaesthesia. A pre-study survey was conducted in the hospitals of Oxford University Hospitals NHS Foundation Trust that determined no consistency in equipment provision, even with individual theatre suites. This survey is summarised in Appendix 10.3, and the results of this work were used to inform the design of the following study, intended to test the prediction that there is a relationship between users' familiarity with their devices, and their performance when reading data from it.

4.3 Methods

4.3.1 Design

This is a within-subjects repeated-measures observational study.

4.3.2 Subjects

45 participants were recruited into three groups, based on their professional background. Recruits were drawn from the local group of anaesthetists, intensive care nurses, and clinical medical students.

The sample size was calculated to detect an estimated effect size of 0.13, based upon data on memory-guided cuing reported by Summerfield (10). This assumed $\alpha = 0.05$ and $\beta = 0.8$.

4.3.3 Outcome Measures

The primary outcome was determined to be time to complete two data recording tasks. Secondary outcomes were the time taken to locate the first and last item on each task display, the number of visual fixations required to locate targets on each display, the extent of eye movements across each display, and the order in which items were located on each display. The incidence of errors was also tracked, with separate counts of missed and incorrectly recorded values. Participant familiarity was

recorded using a visual analogue scale (VAS), calibrated between '0 – I have never seen this display before' to '100 – this is the only display that I use'.

Separately a record was kept of the number of on-screen objects and total alphanumeric characters placed on each display (counted using the methods described by Wolfe (11)), whilst display complexity was measured using the algorithm developed by Rosenholtz (12). Both feature-congestion and sub-band entropy scores were generated for each display using publically accessible scripts (13) run in MATLAB r2014a (Mathworks UK, Cambridge, UK).

4.3.4 Study Procedures

The study was approved by the University of Oxford Central Research Ethics Committee (MSD-IDREC-C1-2014-030).

The test comprised two sequential tasks. Initially participants were instructed to find and record values for heart rate (HR), non-invasive blood pressure (NIBP), oxygen saturation (SpO_2), end-tidal carbon dioxide concentration (E_tCO_2) and end-tidal anaesthetic agent concentration (E_tAA) from a series of 12 clinical displays. The second task comprised finding only a single specified parameter, respiratory rate (RR), on each display.

The development of these displays is described in Appendix 10.2.5. All 12 monitor recreations (Figure 4-1) were used in both tasks, and were presented in a new random order each time. A custom Visual Basic 2010 application (Visual Studio 2010, Microsoft, Redmond, Washington, USA) was created by the author to present the displays and record responses.

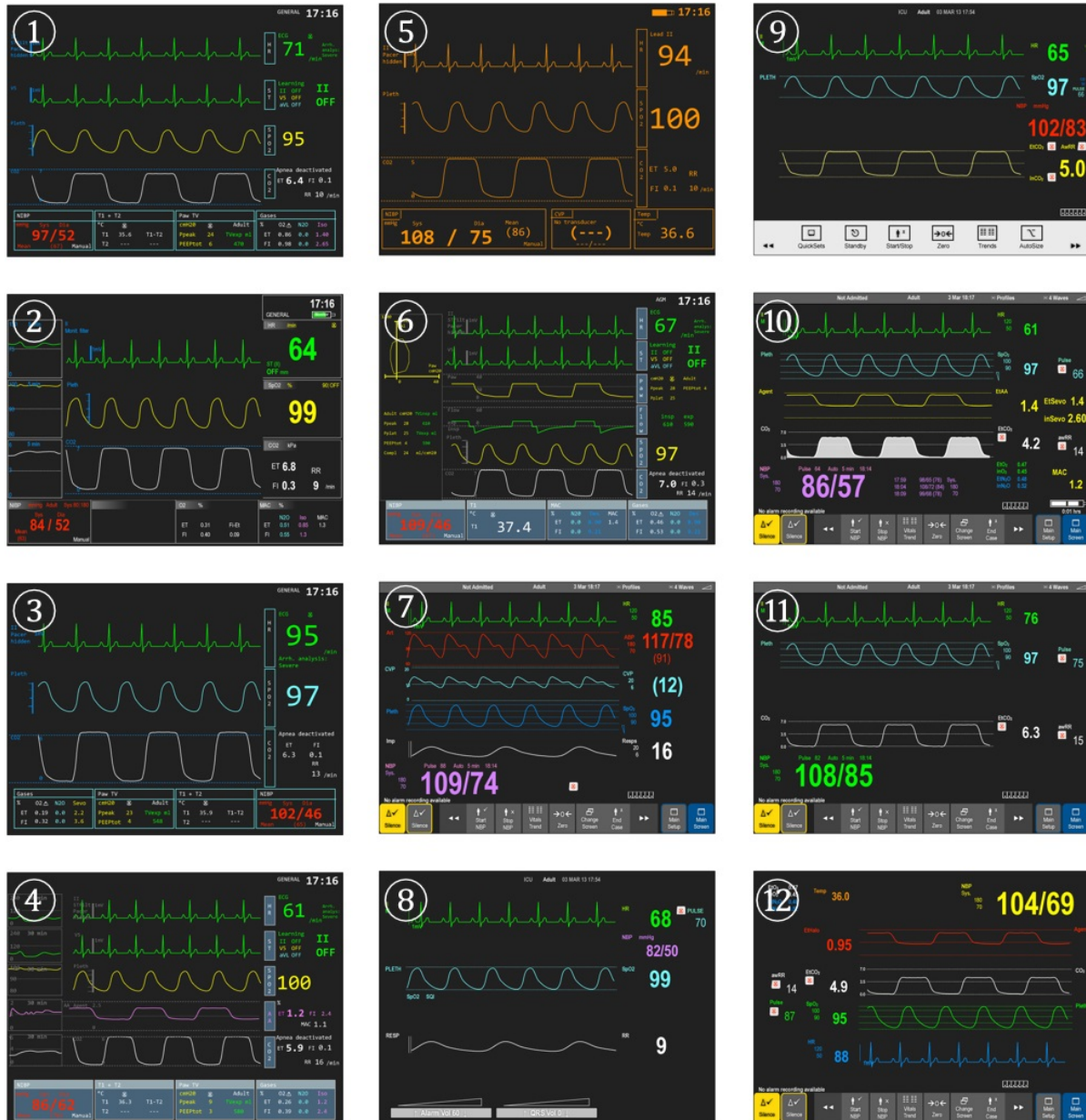


Figure 4-1: Illustration of the twelve monitor layouts used in the experiment; layouts 1-11 recreate real displays, but layout 12 is entirely fictional

Most monitors display heart rate derived both from ECG and pulse oximetry, and the ICU monitor (layout seven) displays both invasive and non-invasive blood pressures. In cases where parameters were available from multiple sources, each differed by one or two points so that it would appear which measure participants were consulting, without introducing implausible discrepancies. In either task, if participants were confident a parameter was not displayed on a layout they were instructed to mark this. Participants were given 60 seconds to complete their record for each display, but could

advance the test early at their discretion. After each display participants were required to rate their familiarity with the monitor.

During each task participants wore an eye tracking device (described in Appendix 10.2.3), and recorded target values on a case record form (CRF) in tabular format. This was chosen over an analogue of an anaesthetic- or 'track-and-trigger' chart as it was felt either of these formats would unfairly disadvantage the undergraduates, whereas a novel table would be equally unfamiliar to all participants. This CRF document also recorded participants' usual place of work.

4.3.5 Analysis

Continuous variables were analysed using repeated-measures multivariate analysis of variance (RM-MANOVA) (14), using monitor layout as a 12-level within-subjects factor. Professional background was used as the between-subjects factor. Outcomes of interest were the subject of a post-hoc Bonferroni test². These included: familiarity score, task duration, variance in fixation co-ordinates, total duration of monitor fixations, and number of monitor fixations. Post-hoc testing was included after it became clear early in analysis that *a priori* assumptions about the interaction between self-assessed familiarity and performance were false (see section 4.4.1). The Bonferroni test is a recommended measure for such post-hoc exploration of data (15). Proportional data were analysed using the Chi-squared test (or Fisher's exact test where appropriate).

4.4 Results

4.4.1 Familiarity Scores and Task Performance

There was a significant effect of professional background on the score given for familiarity with each display ($p < 0.01$). Overall, anaesthetists rated monitors significantly more highly than the

² This is distinct to the Bonferroni *correction*, a method of accounting for multiple statistical tests when assessing significance. The Bonferroni *test* is a method of assessing all pairwise combinations of levels of the within-subjects factors providing both confidence intervals and p -values for each comparison (see www.laerdstatistics.com for further information).

undergraduates ($p = 0.02$), but ICU staff did not significantly differ overall from anaesthetists ($p = 0.78$) or undergraduates ($p = 0.06$). These data are summarised in Table 4-1 and Figure 4-2.

Layout seven was unique in that it was rated more highly by ICU staff ($M = 90.5$, $SD = 21.55$). Significant differences were apparent versus both anaesthetists ($M = 61.30$, $SD = 29.39$, $p = 0.02$) and undergraduates ($M = 23.4$, $SD = 29.97$, $p < 0.01$). Layouts eight and nine were given non-significantly higher scores by ICU staff. Layout 12 did not replicate any real display pattern, and was accordingly rated as unfamiliar by all participants.

Mean familiarity scores given across all monitor layouts, by professional background

		Mean Familiarity Score	SD
Professional Background	Anaesthetists	46.60	17.87
	ICU Staff	38.80	19.54
	Undergraduates	21.65	18.87

Table 4-1: Mean familiarity scores by professional background

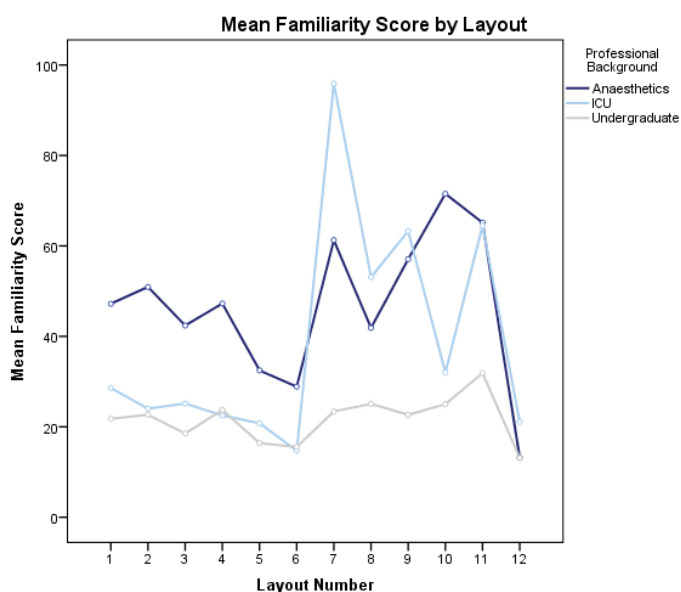


Figure 4-2: Mean familiarity score given to each layout, split by professional background

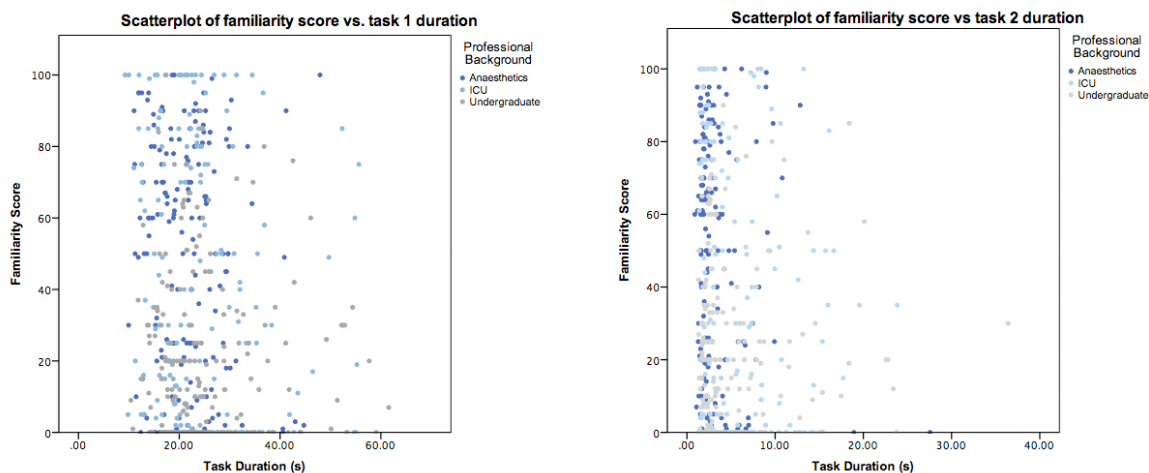


Figure 4-3: Scatterplots of task time against familiarity score

Figure 4-3 demonstrates that there were no simple correlations between familiarity score and task time in either task one (Spearman’s $\rho = -0.15$) or two (Spearman’s $\rho = -0.18$). There were similarly poor correlations between the VAS and total fixation duration, total fixation number, and mean fixation duration. These data are summarised in Table 4-2. Sub-analysis by professional group also failed to reveal any meaningful correlations.

Correlation between reported VAS score and eye-tracking measures

	Task block 1		Task block 2	
	Spearman’s ρ	p value	Spearman’s ρ	p value
Total task time	-0.15	<0.01	-0.18	<0.01
Duration of fixations	-0.19	<0.01	-0.18	<0.01
Number of fixations	-0.14	0.01	-0.03	<0.01
Mean duration of fixations	-0.11	0.13	-0.17	<0.01

Table 4-2: Eye-tracking correlations with familiarity score

A substantial number of professionally qualified participants were unable to correctly determine the layout from their own clinical area. This was significantly more frequent amongst the anaesthetists ($p = 0.01$), where eight of the 15 recruits did not rate their workplace layout as most familiar, compared with only one of 15 ICU staff.

Where an incorrect layout was chosen, the display selected was almost always from the same manufacturer as the device found in their own clinical area. Only two people selected a layout from a different manufacturer as their most familiar. Interestingly, one individual was a senior

anaesthetist working in a department with equipment supplied only by the other manufacturer. The second individual, also an anaesthetist, was working in an area where they could plausibly encounter monitors from both manufacturers.

In 19 cases (seven anaesthetists, six ICU staff, six undergraduates) the fictitious monitor was given a higher score than a real layout. Only three anaesthetists gave the fictitious layout a score of zero, compared with eight ICU staff and six undergraduates. One ICU nurse gave the fictitious display a score of 100. These data are presented in Figure 4-4. Of the qualified participants who gave the fictitious screen a non-zero score, all but two (both anaesthetists) worked most frequently with Philips monitors (which share iconography and typeface with the fictitious display).

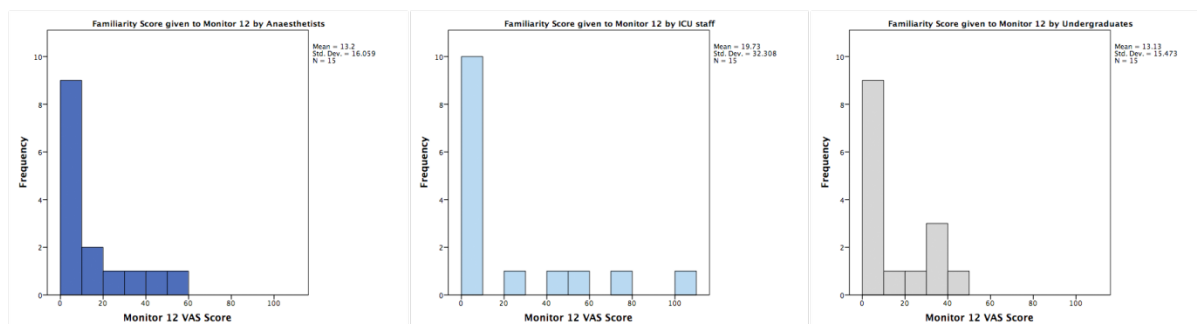


Figure 4-4: Familiarity scores given to the fictitious layout by each professional group

4.4.1.1 Display complexity and task performance

Experimental Layouts and Display Complexity						
Layout	Alpha-numeric characters	Waveforms	Text-objects	Total objects	Feature Congestion	Sub-band Entropy
1	295	4	69	73	4.65	2.86
2	212	6	67	73	4.69	2.69
3	255	3	65	68	4.51	2.75
4	377	10	90	100	4.54	2.99
5	148	3	38	41	2.92	2.89
6	429	7	109	116	4.73	3.10
7	271	5	37	42	3.94	2.72
8	96	3	19	22	3.02	1.92
9	150	3	20	23	3.84	2.58
10	396	4	47	51	4.88	3.04
11	274	3	25	28	4.17	2.57
12	266	4	30	34	3.72	2.43

Table 4-3: Measures of display complexity

Theatre monitors had higher feature congestion scores than monitors found in anaesthetic rooms ($p < 0.01$) or recovery/ICU ($p < 0.01$), but anaesthetic room monitors did not significantly differ from those in recovery/ICU ($p = 0.29$). There was no significant difference in the number of total objects between theatres and anaesthetic rooms ($p = 0.08$) or recovery areas ($p = 0.11$).

There was a correlation between the number of alphanumeric characters and both feature congestion score (Spearman's $\rho = 0.77$) and sub-band entropy (Spearman's $\rho = 0.657$). Correlations were also noted between total on-screen objects and feature congestion (Spearman's $\rho = 0.76$) and sub-band entropy (Spearman's $\rho = 0.77$)

Monitors manufactured by GE/Datex tended to be more complex, displaying significantly more text-based parameters ($p < 0.01$) and total objects ($p < 0.01$) when compared with monitors manufactured by Philips.

There were significantly fewer errors made reading displays manufactured by Philips ($p < 0.01$) compared with Datex displays. Some correlation was noted between total number of displayed

objects and both mean errors (Spearman's $\rho = 0.70$) and mean task one duration (Spearman's $\rho = 0.63$), as illustrated in Figure 4-5. Correlations were noted between total on-screen objects and total duration of fixations (Spearman's $\rho = 0.70$); variance in fixation co-ordinates (Spearman's $\rho = -0.39$ [x]/ -0.67 [y]); and total number of monitor fixations (Spearman's $\rho = 0.39$).

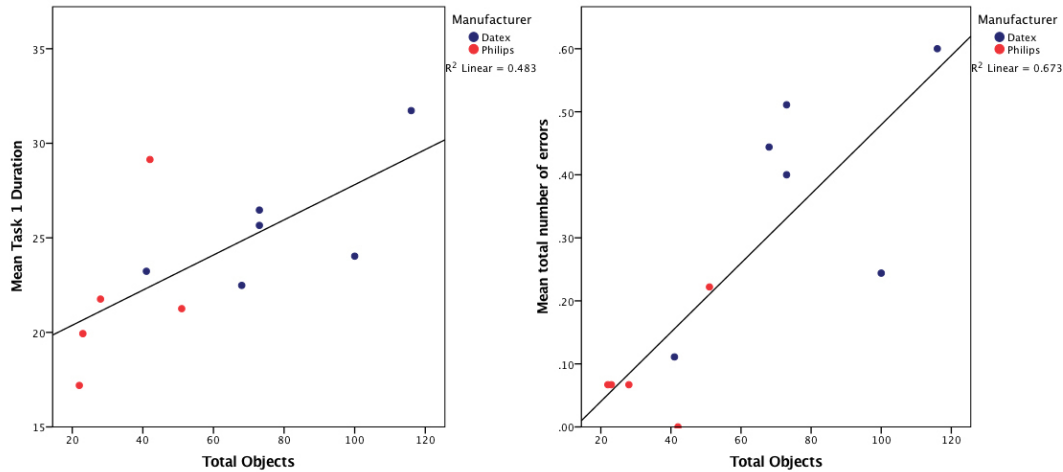


Figure 4-5: Correlation between task time and total number of on-screen objects (L) and mean number of errors (R)

4.4.1.2 Professional Background and Task Performance

Unlike self-reported familiarity, professional background did have some significant effects on task performance, notably on time to complete task two ($p = 0.04$), and duration of fixations made on monitors in both task one ($p < 0.01$) and task 2 ($p = 0.02$). There were no significant trends observed in task one time ($p = 0.08$), variance in fixation co-ordinates ($p = 0.40$ to $p = 0.95$), or total number of fixations in either task ($p = 0.24$ and $p = 0.98$). Task completion times are summarised in Figure 4-6, and fixation duration in Figure 4-7.

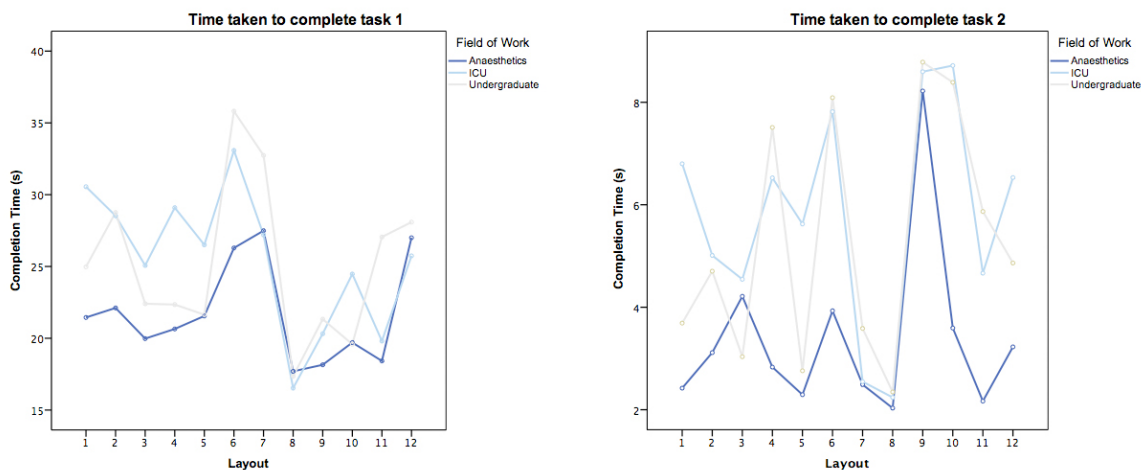


Figure 4-6: Time taken to complete each task by professional group

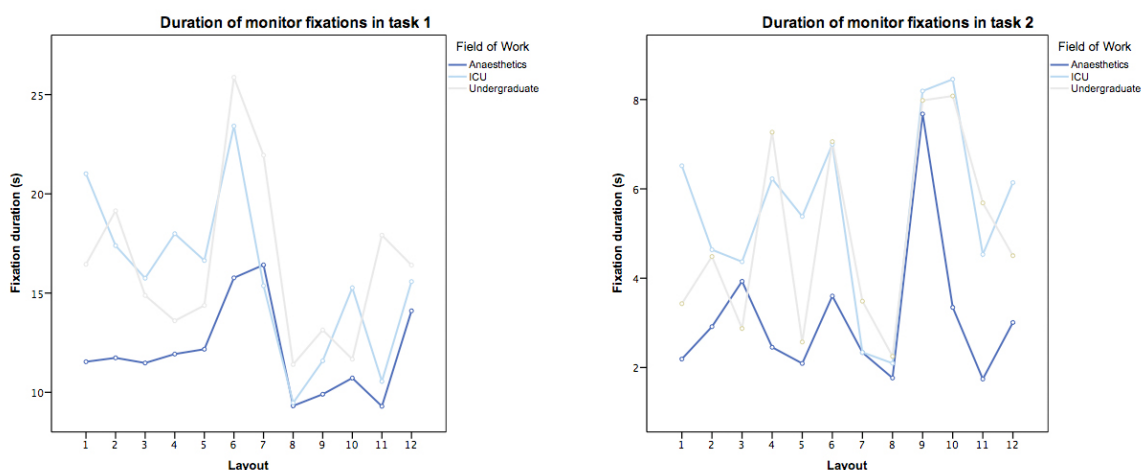


Figure 4-7: Total duration of fixations on monitor in task 1 and 2 by professional group

Overall the mean duration of task two was 4.76s ($SD = 4.47$). Post-hoc tests determined that anaesthetists were 1.9 to 2.4s faster than the other professional groups in locating the required value in task two ($p < 0.01$).

4.4.1.3 Error rates

Professional group had a significant effect on error rate ($p < 0.01$). Anaesthetists made fewer errors than both ICU staff ($p < 0.01$) and undergraduates ($p < 0.01$), as illustrated in Table 4-4 and Figure 4-8. ICU staff were more likely to miss displayed values ($p < 0.01$), whilst undergraduates were equally likely to record incorrect values or miss a displayed value ($p = 0.20$).

Task 1 errors by type and professional group

	Parameter not successfully located			Value of parameter incorrectly recorded			Total Errors		
	Mode	Median	Range	Mode	Median	Range	Mode	Median	Range
Anaesthetics	0	0	0-2	0	0	0-1	0	0	0-3
ICU	0	4	0-10	0	0	0-6	4	5	1-11
Undergraduates	0	0	0-8	0	0	0-8	0	3	0-11

Table 4-4: Errors by type and professional group

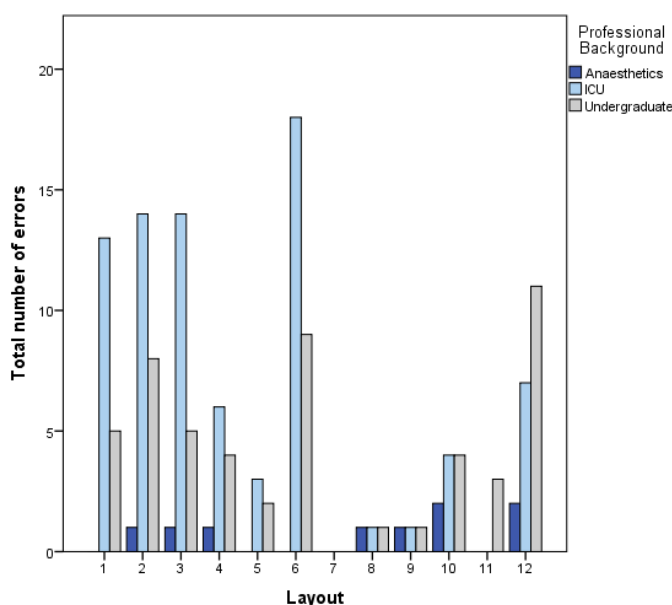


Figure 4-8: Total number of errors made on each display, grouped by display layout and professional background

E_tAA was the value most frequently omitted or incorrectly recorded. HR was never omitted, but was recorded incorrectly on five occasions. In four of these cases this error was made on layout 12. On a further two occasions this error was made by a participant who subsequently corrected their record. Professional background did not affect the likelihood of making this error ($p = 0.56$). These data are summarised in Table 4-5.

Number of occasions on which an error was made in recording each parameter

	Anaesthetics		ICU		Undergraduates		Total
	Not successfully located	Incorrectly Recorded	Not successfully located	Incorrectly Recorded	Not successfully located	Incorrectly Recorded	
HR	0	1	0	0	0	4	5
NIBP	1	2	40	1	0	2	46
SpO ₂	1	1	0	2	0	1	5
E _t CO ₂	1	0	5	3	1	4	14
E _t AA	1	1	53	13	16	25	109
Totals	4	5	98	19	17	36	179

Table 4-5: Error types by parameter and professional group

Search order also varied by professional group. Anaesthetists were much more likely to record items according to the order specified in the CRF form (i.e. HR, then NIBP, SpO₂, E_tCO₂, and finally E_tAA; $p < 0.01$), doing so on 124 of 180 occasions. Only 79/180 observations made by ICU staff and 66/180 observations made by undergraduates were recorded in this order.

In all cases the recorded heart rate was derived from the ECG, and not the pulse oximeter. Only non-invasive blood pressure was displayed on 11 monitors, however on layout seven (the ICU monitor) both an ABP and NIBP value were displayed. More than half of the anaesthetists (8/15) and undergraduates (11/15) recorded the invasive blood pressure when searching layout seven, compared with 5/15 ICU nurses. All five of these nurses subsequently amended their record, replacing the value with NIBP. While nine medical students also made this change, only three anaesthetists corrected their record.

4.4.2 Learning Effects

Figure 4-9 demonstrates that in task two ICU and undergraduate participants' performance improved over time ($p < 0.01$). There was a significant interaction between professional background and progress ($p = 0.03$). Anaesthetists were faster overall in almost all cases, however they did not improve between first and last layout to the same extent as ICU nurses and undergraduates, who both improved at similar rates. Figure 4-10 presents the same data, but excludes performance on monitor 9 from the analysis. Monitor 9 was uniquely a target-absent trial that participants were significantly slower to terminate than all other layouts.

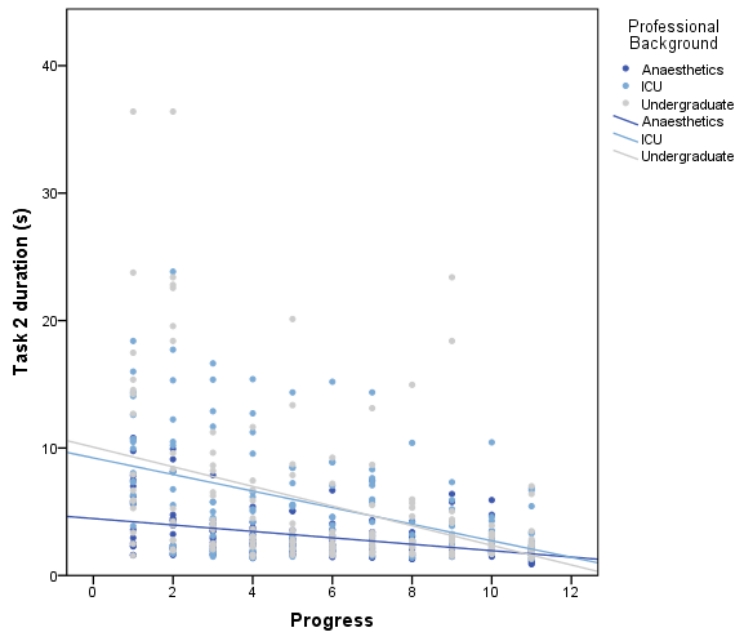


Figure 4-9: Mean task time as a function of progression through task

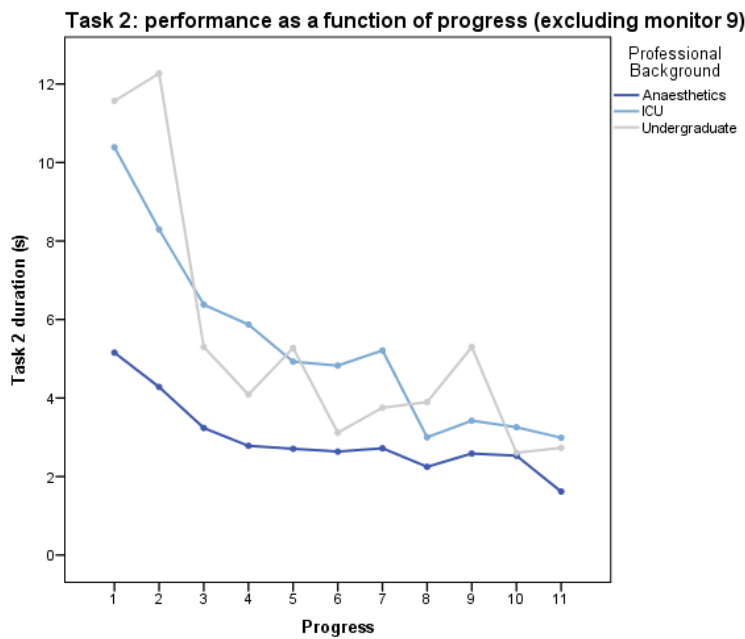


Figure 4-10: Mean task time as a function of progression through task (excluding monitor 9)

4.5 Discussion

Adequate monitoring of patients during anaesthesia or critical illness is an essential standard of care.

Good-quality situational awareness depends on integrating all available information, and equipment

must not impede this process. It is known that during crisis situations a great deal of attention is focused towards monitoring devices (3), and that these are also the scenarios in which errors are most likely to occur (16). Emergency situations are often those in which task-demands, time pressure, and mental workload are highest, and it is therefore essential that clinicians can acquire information accurately and rapidly.

Before information can be consciously used, level one situational awareness must be achieved. This requires the perception of relevant information. From a psychological perspective, the act of locating target parameters on the displays can be considered a visual search task. Much work has been done investigating how individuals search visual scenes, locate targets, and derive meaning from items therein. Some of these findings offer explanations of the pattern of results observed here, and these insights may be useful in the design of future clinical devices.

4.5.1 Visual Search

'*Guided search theory*' has been proposed to explain the mechanism by which individuals visually locate items of interest (17, 18). This is posited as a two-stage process in which basic information about objects is acquired rapidly and pre-attentively. This basic information generates a set of '*feature maps*', and items are prioritised for further attention using details from this map modulated by the individual's particular needs.

A bottleneck exists between the two stages of visual processing, meaning that search performance varies depending on the degree of cognitive input required. Guided searches can be described as either '*efficient*' or '*inefficient*', a distinction based on how response times vary with number of distractors (17). Efficient searches are very rapid, and remain so even in the presence of multiple distracting elements. These searches are relatively uncommon and only occur when seeking unique, primitive features (such as colour or size). It appears that the brain can make basic comparisons of such items at a pre-attentive stage, and that the capacity for such processing is essentially unlimited. As such, efficient searches do not need to pass through the bottleneck. Features that subjectively

appear to 'pop out' of scenes easily are probably doing so because they possess such distinguishing properties. Items that differ from the objects around them (are '*locally contrasting*') will also tend to be more prominent than those with close resemblance to their neighbours.

Inefficient searches occur where more complex cognitive processing is required to discriminate between objects. Examining items that possess subtle differences in features, identifying combinations of features amongst targets that share common characteristics, or seeking object features that are not amenable to pre-attentive processing all require inefficient searches. This is more computationally intensive as features of the target objects must be 'bound' together, thus processing is both slower and more resource dependant. The least efficient searches occur where no target features are suitable for pre-attentive processing, so each item must be individually located, evaluated, and compared.

Between these extremes are intermediate searches whereby targets are defined by conjunctions of basic features. Distractors that share no basic features in common with the intended target can be automatically disregarded (*'feature inhibition'*), and attention divided amongst the remaining items (19). Discarding certain distractors based on primitive features requires little processing, and is termed reducing the '*functional set size*' (11), since although they are still present, they require little cognitive effort to be discounted from a search.

Real-world searches are generally more complex than the tasks used in laboratory experiments, and will usually take the form of intermediate or inefficient searches. This would apply to the searches performed during this experiment. Such searches illustrate the interface between 'bottom-up' and 'top-down' guidance. Bottom-up guidance depends on intrinsic features of the item displayed, such as colour, shape, or distinctiveness. Top-down guidance is user-dependant and based on the current task (20), and appears more influential in real-world situations than in laboratory experiments (21). Searches in complex natural environments invariably involve large numbers of distractors, but an understanding of the elements that make up the scene confers additional search advantages

through '*scene semantic guidance*'. This involves using higher-level understanding and employing that knowledge to improve search-efficiency (22, 23).

As an example, consider searching the image in Figure 4-1 for a defibrillator. If you have some understanding of such devices, even if you have never encountered the specific model in this picture it is unlikely that your search will be random. You may predict, for example, that the device should be close to the patient, or attended by a member of staff, or perhaps even the type of information that should be displayed on its screen. Using this understanding it is possible to eliminate implausible locations and allocate search-priority to regions most likely to contain the object. The search therefore appears to occur more quickly, not because of any changes in the conspicuousness of the device, but rather through phenomena internal to the viewer.



Figure 4-11: Management of a simulated cardiac arrest

The example presented in Figure 4-12 can be used to draw some of these concepts together. The figure demonstrates two replica clinical monitors, both taken from theatre environments. In both cases the basic screen layout is the same, with identical typeface and display size, but the right-hand display recreates the monochrome monitor found in some anaesthetic rooms, whereas the left-hand

display replicates a colour screen. There are therefore differences in bottom-up features that may affect search performance.

If the two displays are searched for heart rate (recorded from ECG Lead II) there is likely to be a difference in response time between the two displays. With some experience of clinical monitors an individual is likely to be conditioned to expect heart rate to be displayed in green, and it is therefore reasonable to suggest the value is more easily located on the left-hand display. In this case it 'pops out' of the display, colour being a feature amenable to efficient search.

With no previous exposure to this display, or prediction of item colour, the search becomes more difficult. With some clinical experience the search might still be guided by top-down factors, as an expectation of the form in which heart rate should be displayed (usually a two-digit number, and often in the range of 60-99) would enable implausible values to be rapidly discarded: for example, the green value displayed for expired tidal volume (TV_{exp}) is 548, a number impossible to associate with heart rate. An individual with clinical knowledge might even use the electrocardiogram waveform fields to identify candidate screen-positions for heart rate. The search, although inefficient, would be somewhat faster than someone with no such experience to draw upon, who would need to resort to interrogating each data label individually.

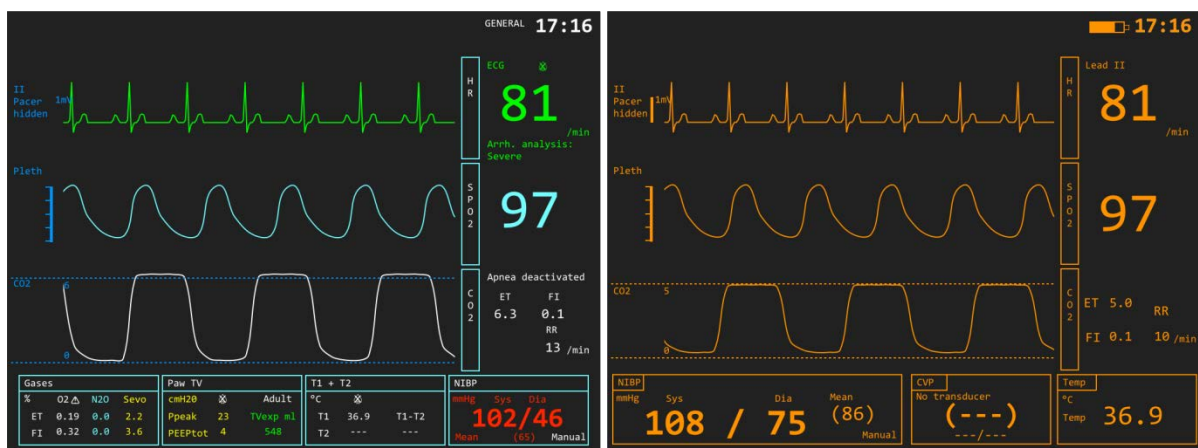


Figure 4-12: Example clinical monitors

Colour can have a profound effect on visual search, and it is known that familiar objects presented in unusual colours can impede search ability (24). Data from our experiment would seem to support this. A common mistake made on display 12 was the confusion of saturations with ECG. On every real display heart rate and ECG waveform was displayed in green. This was, in fact, the only common feature between the clinical displays. Users are therefore strongly conditioned to expect green values to represent heart rate. Monitor 12 disrupted this pattern by displaying heart rate in dark blue and oxygen saturation in green. Participants' colour preference overrode their noting of the parameter's label, and only two people (both nurses) who made this error subsequently corrected their record. It was therefore clear that few participants were cross-checking their data-selection against the on-screen labels. Since saturations were in the range of 94-98, plausible for HR, participants were denied a potential indicator that they had selected an incorrect value. It might be interesting to repeat this experiment using a more alarming or improbable value to test if a greater proportion of participants would question their initial, colour-based, selection.

Clinical consequences have arisen from the predominance of colour-preference over labelling. A report was written by American anaesthetists after a manufacturing fault meant a vaporiser entered clinical service with a mismatch between its labelling and colour-coding (25). Colour-coding for anaesthetic agents is internationally standardised, so anaesthetists are strongly conditioned to expect colour to denote device contents; in this case, a purple-coded vaporiser (indicating the contents should be isoflurane) was delivered with a sevoflurane label prominent on the case. The authors were unable to state with certainty how long the device remained in service before the error was detected.

As seen in these data, display complexity is also highly relevant. Consider the example presented in Figure 4-13, which also replicates a real clinical system currently in use. The software, Cerner Millennium®, forms the electronic patient record (EPR) used in many hospitals in the UK and US. This system is intended to provide a central repository of patient information, including laboratory test

results and clinical observation values. The figure presents the same basic range of observations as seen in Figure 4-12, but finding specific items such as heart rate becomes a much more onerous task. This is perhaps not an entirely fair comparison, because the typeface is considerably smaller than in Figure 4-12, however it does demonstrate how the presence of a great many distractor items with few distinguishing features makes searching for information much more difficult. Analysing the EPR screen for feature clutter and sub-band entropy suggest scores of 6.43 and 3.98 respectively. Using the correlations observed in the experimental data predicts a mean search time of approximately 30 seconds to record the dataset used in experimental task one. An informal test of this prediction with small sample of volunteers suggests this estimate may be a realistic approximation.

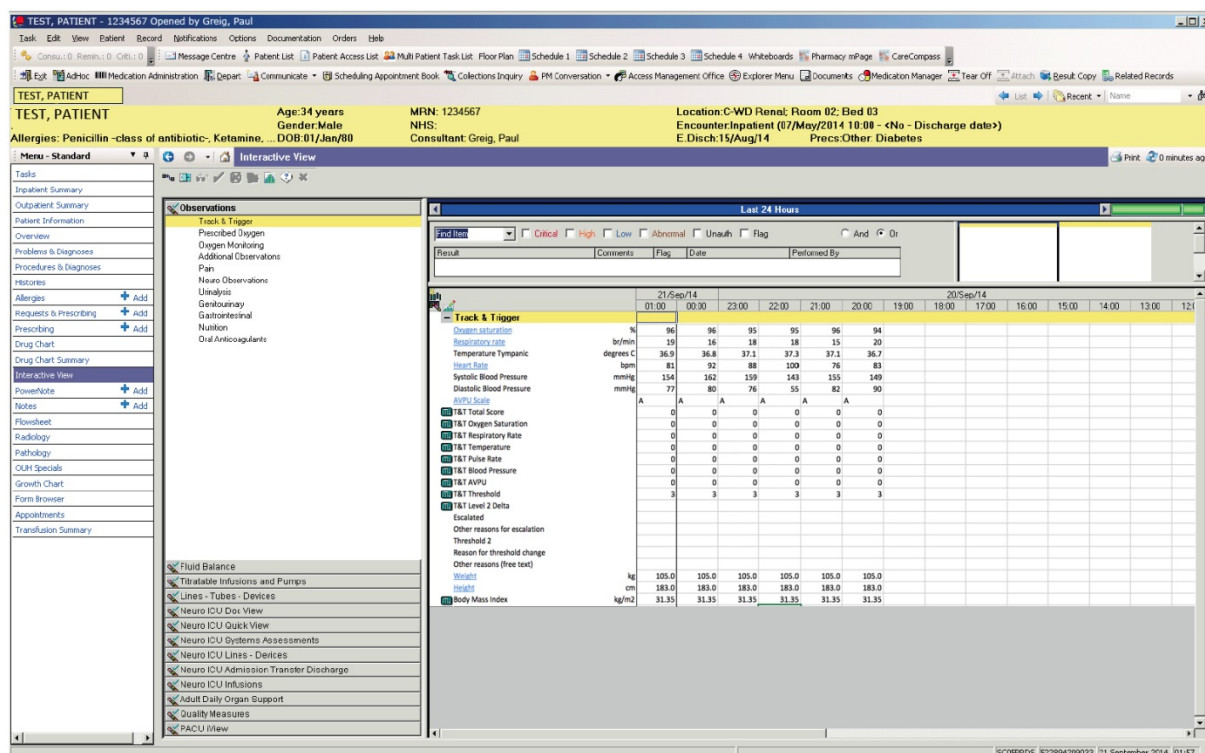


Figure 4-13: Screenshot of the Cerner Millennium® electronic patient record

Searching the monitors of Figure 4-12 is also somewhat simpler than that of Figure 4-13 for other reasons. Whilst searching for textual material English-speakers generally start searching from the top left of the screen, and scan line-by-line until the bottom right is reached (17). In the replica monitors seen in Figure 4-12 the spatial layout of the display facilitates ordering search in this

manner. The Millennium® display shown in Figure 4-13 has less linear structure, with vertical lines and changes in text-alignment disrupting horizontal scanning across the screen. Ergonomists suggest that searching such a display degrades performance and increases mental workload compared with a more streamlined interface (26). This screenshot demonstrates what would be termed 'high display-density' (many extraneous items must be examined to find the target, analogous to increasing set-size), and 'poor display-layout' (by degrading the usefulness of innate search heuristics).

This preference for searching monitors left-to-right and top-to-bottom explains why 21 of the 45 participants recorded ABP in preference to NIBP on the ICU display (monitor seven). Following this heuristic, ABP would be the first plausible value encountered to satisfy the need to record a blood pressure. All nurses and most of the medical students who recorded ABP later replaced this value with the NIBP-derived measurement, but very few of the anaesthetists made this correction. Search satisfaction is likely to be a factor in this. Having found an appropriate value it seems that anaesthetists simply stopped searching for an alternative. It is not clear why ICU nurses and students were less vulnerable to this effect. Anaesthetists may have been making a clinical judgement about perceived accuracy of invasive pressure measurement compared with NIBP, and therefore reached a decision that ABP was a better value to record (despite their instructions). This would go some way to explain the pattern compared to medical students, who perhaps were not attempting to interpret the values clinically as they encountered them. This does not satisfactorily explain the nurses' preference to record NIBP as instructed, as these staff routinely use both forms of measurement in practice, and arguably use invasive pressure monitoring more frequently than theatre anaesthetists.

One might speculate that the results would have differed if NIBP were to be displayed first (a possibility on some machines). Certain monitors do display the same information derived from different sources simultaneously. In most cases these measures are temporally related, but the specific case of blood-pressure measurement it is customary to cycle NIBP less frequently when an

arterial line has been inserted, potentially dropping from five-minute intervals to 30-minutes or longer. Should this value be displayed first it is not beyond possibility that an old and out-of-date value may be selected in a high-pressure situation, so it is crucial that secondary sources of information are not placed on-screen above the primary source.

The predicted duration of a visual search also differs greatly depending upon whether the target is present or absent (27). Determining that a target is absent is considerably slower than confirming an item is present. This is intuitive since target-present searches generally terminate immediately on successful detection and verification, however in target-absent conditions a sizeable proportion of distractors need to be inspected and discarded before a subject can satisfy themselves that there is genuinely no target to detect.

The time and accuracy of target-absent searches varies with target prevalence. A higher false-negative rate is observed when targets are rarely encountered, as users inspect fewer items before declaring a target absent. The time taken to terminate the search is therefore modulated by previous experience; Chun and Wolfe describe this as an '*internal staircase*' procedure (27), whereby the activation threshold changes with each trial. After a correctly concluded target-absent search their threshold changes and individuals become faster to terminate subsequent target-absent searches. This phenomenon has been linked to missed diagnoses in radiology (28).

This effect explains some of the data collected from layout nine in task two. This was uniquely a target-absent trial, as layout nine was derived from an anaesthetic-room monitor that did not feature respiratory rate in the display. By a considerable margin this was the longest duration task two trial for anaesthetists. As illustrated in Figure 4-6 anaesthetists took on average 8.22 seconds to determine that respiratory rate was not present, and progress to the next task. Monitor three had the next longest mean-task-duration, at 4.21 seconds. Although this same trial was also slowest for the other two professional groups, the differences were less striking. For ICU nurses the mean time of 8.54 seconds was only slightly slower than the next two slowest trials (monitor ten = 8.25s,

monitor six = 7.47s). Similarly medical students' mean time for monitor nine was 8.78 seconds, versus 8.39 seconds and 8.09 seconds for monitors ten and six respectively. These same monitors were amongst the most complex of the displays, showing the highest number of on-screen characters, the highest feature-congestion score, and the highest sub-band entropy scores. This raises the question that perhaps the non-anaesthetists were overwhelmed by the displays, and since no targets 'popped out' at them they were forced to resort to an individual inspection of many items even to detect target presence.

4.5.2 Familiarity and search performance

There is evidence that humans are able to extract contextual information rapidly from natural scenes (29). This allows determination of '*basic scene categorisation*' after views as short as 125ms (30). It appears that this categorisation may occur even before specific recognition of the scene's constituent parts (31, 32), and this occurs with little conscious effort.

This initial interpretation of a scene may be useful in visual search. Wolfe notes that 'typicality' may have an important effect on terminating searches, particularly when declaring a target absent (11). In one example it is suggested viewers may respond that a refrigerator is not present in a bathroom scene, not based (wholly) on a confirmatory visual search, but rather on the implausibility of the premise. Visual search is therefore modified by viewer expectation, and this may go some way to explaining the high number of responses incorrectly recording a value as not displayed observed amongst ICU nurses and medical students. ICU nurses were highly successful in correctly identifying the ICU monitor layout from the experimental set, and it is likely they possess a strong mental model of a 'typical' display. A prototypical ICU monitor is likely to incorporate ECG, NIBP, and respiratory rate, but rarely is it necessary to display anaesthetic gas measurements in ICU. This atypicality may lead ICU nurses to discontinue early their searches based on their expectations of what a monitor should display.

Previous exposure to target objects can also confer performance advantages through recognition-based effects. This has been explored in several ways using 'cueing' – showing individuals elements of the search task in advance. These cueing effects demonstrate an important role for memory in visual search. These have been tested using various permutations of accurate and inaccurate cues, cues presented before- or after display presentation, and cueing using specific versus generic target information. The relationship is not entirely straightforward however, and it appears that the nature of previous exposures is important in determining how well information is retained. A study by Vo and colleagues noted that simply looking at an item in a display does not create a representation that can be used to guide search (33). Instead improvements in response times are only seen when the specific object has actively been interacted with, for example by being the subject of a previous search. It appears therefore that specific, definite information about search targets is needed to gain maximum improvement in response time compared with more generic information such as textual cues, or category-matched but non-identical prompts (22).

It is generally found that more accurate foreknowledge of the target or environment confers greater improvements in search performance (34, 35), probably by enabling faster rejection of distractors (21). This informative experiment, by Malcolm *et al* in 2009, highlighted a difficulty in searching object-arrays. Unlike a naturalistic scene there is no established spatial relationship between items, limiting top-down guidance to foreknowledge only of the visual characteristics of the target. One might consider clinical monitoring devices to be closer to object arrays, as there is no 'natural' way to position items on screen. Users cannot necessarily predict object position on screen, and must instead depend on the target values' characteristics, or learned foreknowledge of their display. This goes some way to explaining why the performance advantage of anaesthetists disappeared when exposed to monitor 12, the fictitious display. Although this monitor shared iconography and typeface with real displays the specific layout of the display shared no common features with any display the participants could have encountered.

In this experiment, healthcare professionals' ability to recognise their equipment was very poor. At first glance this finding was somewhat counter-intuitive. Knowing that scene assessment involves rapid estimation of coarse layout, and this precedes extraction of specific visual information, it follows that these individuals must have been assessing layout familiarity based on early processing of generic features. Their decision is perhaps based on global properties such as iconography or typography, rather than specific parameter formatting or placement. This would be consistent with Rensink's theory that long-term scene memories are stored as 'scene schemata' (32), in which scene category is stored along with a subset of specific representations, but that *'items consistent with the schema do not need to be encoded in detail'*. This may explain why participants rated the fictitious display with non-zero familiarity scores – a generic 'Philips template' probably exists in the minds of staff exposed to such displays, so 'first impressions' lead them to erroneously conclude they have encountered the display before and this conclusion is not questioned when they later interact with the specific elements in the layout. Staff may then be falsely reassured when encountering new devices that share some generic features with a device they have used previously, and this may strengthen the case for better standardisation of equipment.

With hindsight, the poor performance of staff rating their familiarity with each display should have been predictable. There are existing data that suggest even highly familiar everyday objects such as coins or corporate logos are difficult to replicate from memory (36-38). It may be that disparate displays are so frequently encountered by anaesthetists (or infrequently encountered by students) it is not possible for them to create a generic schema; ICU nurses on the other hand have much stronger single-monitor schema due to repeated exposure to similar displays.

One surprising aspect of this is that the anaesthetists were faster in detecting targets across most layouts despite the poor reliability of self-reported familiarity; if visual cueing is responsible for this performance improvement then the priming image appears not to be the monitor layout as a whole (or at least not a consciously accessible replica of the layout). There is evidence that scene-guidance

is derived from implicit, rather than explicit, long-term memory data (39), so it is plausible to expect performance improvements in the absence of conscious recognition. The clinicians in our experiment do appear to have some consciously accessible memories for elements of monitor displays, but the difficulty recalling specific details such as how features interact on screen (particularly in respect of colour and screen position) suggests the features of these memory objects are not tightly bound to one another.

Although a direct performance-effect attributable to familiarity cannot be demonstrated by this dataset, the differences associated with professional background suggest that past exposure to these devices must have some effect on visual search. It is known that mental representations of environmental stimuli are held in working-memory (40), and that these representations can be searched in a manner similar to the original stimulus (41). There is also evidence that prior exposure to scenes enables individuals to ignore distractors more easily (11). Repeated exposure to the same scene modestly improves search for all items (42), but profound effects are seen when searching for a familiar object, with response times improved by up to one-third (10, 43). These advantages may however be relatively short-lived and can diminish as interval between first and second viewing increases (11).

Manginelli terms this learning effect *'the role of repeated context'* (44), and notes that such contextual cueing does not require conscious appreciation that the scene has been previously encountered (44, 45). Excessive spatial distractors reduced or abolished the advantage of previous exposure to a scene, but the 'preview benefit' can be up to 30% when both visual and spatial information about distractors is available (46). Spatial data stored in long-term memory can be used to orientate and optimise visual search (10, 47), so it is plausible to expect these findings to be relevant to screen design, where the colour palette is relatively limited but screens vary considerably in spatial organisation. It has been demonstrated that long-term memory cueing is sensitive to

reward (48), so it is interesting to speculate how these mechanisms might be manipulated to advantage in a clinical setting.

Fei-Fei and colleagues suggest that the same mechanisms underpinning poor recognition-performance may contribute to the causes of change blindness (49). They link this global scene assessment to Rensink's theories on poor change detection (50). Rensink noted that changes were less likely to be noticed if they did not change the 'gist' of the scene. Similarly Simons, in describing his 'door experiment' (51) suggested that variations in how identifying details about an individual were encoded contributed to the likelihood of the experimental subject noticing the person-swap. It was suggested that provided the change did not violate 'category membership' then it would likely pass unnoticed.

There are implications in this for clinical practice. There is a drive towards electronic record-keeping in theatre (52), but a common concern raised by anaesthetists when such systems are introduced to theatres is that they will find it more difficult to detect abnormalities. At first glance this appears counterintuitive, as electronic records are kept by a computer, recorded automatically and contemporaneously, and should be a great deal more accurate than a hand-written chart (53, 54). More advanced electronic systems can include decision-making support and tailored alarms, although these were not mandatory features required of the NHS electronic patient record implementation (55). While removing the need to maintain a paper record should reduce the anaesthetist's physical workload and free up attention for patient monitoring, anecdotally many anaesthetists express concerns they will miss trends as they develop. They value the cognitive input that is required to engage with values as they are transcribed on to a paper chart. There is surprisingly little clinical data addressing this point, but based on Vo's findings (33) without active cognitive engagement with the values it seems at least possible that changes may indeed not be noticed until extreme deviations occur.

The available data from this experiment and the wider psychology literature suggest that familiarity with a clinical monitor will confer advantages when searching for information displayed upon it. This would be expected to reduce mental workload, and although this experiment was not designed to test situational awareness, existing data on perceptual error suggest that risk should be reduced.

Based on these data it is possible to make recommendations for equipment provision. It is clear that specific cueing information provides the best search performance, and thus a case should be made to standardise screen position and colour schemes across clinical areas.

Display complexity should also be minimised. Over and above increasing the difficulty of visual searching there is a subconscious cost to processing distractor information, particularly in the spatial domain. Displays should be limited only to information of immediate relevance, although the increasing availability of touch-screen devices may afford opportunities to place secondary information on tabs that remain hidden until required. New automatic devices or smart-monitoring systems may even be employed to highlight information only when required.

Positioning of items on the screen also affects the likelihood of their being noticed, so parameters should be organised linearly to facilitate left-to-right scanning wherever possible. The search-satisfaction over blood pressure display observed on monitor seven suggests that the most critical values should be displayed topmost, and if confirmatory information derived from other sources must be displayed on the same screen, this secondary information should be placed below the primary source.

4.5.3 Summary

Familiarity undoubtedly improves both response time and accuracy when reading information from these displays. Overly complex displays with superfluous detail and poorly structured layouts are difficult to read, and likely represent latent risk factors.

Standardisation would offer benefits for staff moving between clinical areas, and would enable individuals to acclimatise more rapidly to a new workplace, and there are sound reasons to expect performance advantages to be gained.

This is relevant to the remaining experiments in this thesis, because mental workload is known to be a contributor to vulnerability to perceptual error, and therefore loss of situational awareness. The additional mental burden attributable to poor equipment provision may therefore represent a newly-recognised latent risk-factor in the clinical area, and represents a potential target for future intervention.

4.6 References

1. Endsley M, Garland D. Theoretical underpinnings of situation awareness: a critical review. *Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
2. Vincent C. *Patient Safety*. 2nd ed. Oxford: Blackwell; 2010.
3. Schulz CM, Schneider E, Fritz L, Vockeroth J, Hapfelmeier A, Brandt T, et al. Visual attention of anaesthetists during simulated critical incidents. *Br J Anaesth*. 2011;106(6):807-13.
4. Checketts MR, Alladi R, Ferguson K, Gemmell L, Handy JM, Klein AA, et al. Recommendations for standards of monitoring during anaesthesia and recovery 2015 : Association of Anaesthetists of Great Britain and Ireland. *Anaesthesia*. 2016;71(1):85-93.
5. Walsh T, Beatty PC. Human factors error and patient monitoring. *Physiol Meas*. 2002;23(3):R111-32.
6. Tappan JM, Daniels J, Slavin B, Lim J, Brant R, Ansermino JM. Visual cueing with context relevant information for reducing change blindness. *Journal of Clinical Monitoring and Computing*. 2009;23(4):223-32.
7. Curriculum Development Working Group. *CCT in Anaesthetics: Annex B - Basic Level Training*. London: Royal College of Anaesthetists; 2010.
8. Civil Aviation Authority. *CAP804: Flight Crew Licensing - mandatory requirements, policy and guidance*. London: The Stationery Office; 2015.
9. Air Accident Investigation Branch. *Report on the accident to Boeing 737-400 G-OBME near Kegworth, Leicestershire on 8 January 1989*. London: The Stationery Office; 1990.
10. Summerfield JJ, Lepsien J, Gitelman DR, Mesulam MM, Nobre AC. Orienting attention based on long-term memory experience. *Neuron*. 2006;49(6):905-16.
11. Wolfe JM, Alvarez GA, Rosenholtz R, Kuzmova YI, Sherman AM. Visual search for arbitrary objects in real scenes. *Attention, Perception, & Psychophysics*. 2011;73(6):1650-71.
12. Rosenholtz R, Li Y, Nakano L. Measuring visual clutter. *Journal of Vision*. 2007;7(2):17-.
13. Rosenholtz R, Li Y, Nakano L. Feature congestion and subband entropy measures of visual clutter: Massachusetts Institute of Technology; 2007 [Available from: <http://dspace.mit.edu/handle/1721.1/37593>].
14. Mayers A. *Introduction to Statistics and SPSS in Psychology*. Harlow: Pearson; 2013.
15. Maxwell S, Delaney HD. *Designing experiments and analysing data: a model comparison perspective*. New York, NY: Psychology Press; 2004.
16. Bion JF, Abrusci T, Hibbert P. Human factors in the management of the critically ill patient. *Br J Anaesth*. 2010;105(1):26-33.
17. Wolfe JM. Approaches to Visual Search: Feature Integration Theory and Guided Search. In: Nobre AC, Kastner S, editors. *The Oxford Handbook of Attention*. Oxford: Oxford University Press; 2014. p. 11-55.
18. Wolfe JM. Guided Search 4.0: current progress with a model of visual search. In: Gray W, editor. *Integrated models of cognitive systems*. New York: Oxford University Press; 2007. p. 99-119.
19. Treisman A, Sato S. Conjunction Search Revisited. *Journal of Experimental Psychology: Human Perception & Performance*. 1990;16(3):459-78.
20. Castelano MS, Henderson JM. Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*. 2007;33(4):753-63.
21. Malcolm GL, Henderson JM. The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*. 2009;9(11).
22. Castelano M, Heaven C. The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics*. 2010;72(5):1283-97.
23. Ehinger KA, Hidalgo-Sotelo B, Torralba A, Oliva A. Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition*. 2009;17(6-7):945-78.

24. Olivers CNL. Long-term visual associations affect attentional guidance. *Acta Psychologica*. 2011;137(2):243-7.
25. Rose GL, Eichorn J, Di Lorenzo A. Implications of a mislabelled vaporizer and the importance of color coding 2008 [Available from: http://www.apsf.org/newsletters/html/2009/winter/05_colorcoding.htm].
26. Moacdieh N, Sarter N. Display Clutter: A Review of Definitions and Measurement Techniques. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2015;57(1):61-100.
27. Chun MM, Wolfe JM. Just Say No: How Are Visual Searches Terminated When There Is No Target Present? *Cognitive Psychology*. 1996;30(1):39-78.
28. Evans KK, Birdwell RL, Wolfe JM. If You Don't Find It Often, You Often Don't Find It: Why Some Cancers Are Missed in Breast Cancer Screening. *PLoS ONE*. 2013;8(5):e64366.
29. Greene MR, Oliva A. Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology*. 2009;58(2):137-76.
30. Potter M. Meaning in visual search. *Science*. 1975;187(4180):965-6.
31. Oliva A, Torralba A. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*. 2001;42(3):145-75.
32. Rensink RA. The Dynamic Representation of Scenes. *Visual Cognition*. 2000;7(1-3):17-42.
33. Vo MLH, Wolfe JM. When Does Repeated Search in Scenes Involve Memory? Looking At Versus Looking For Objects in Scenes. *Journal of Experimental Psychology: Human Perception & Performance*. 2012;38(1):23-41.
34. Vickery TJ, King L-W, Jiang Y. Setting up the target template in visual search. *Journal of Vision*. 2005;5(1):8-.
35. Wolfe JM, Horowitz TS, Kenner N, Hyle M, Vasan N. How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*. 2004;44(12):1411-26.
36. Jones GV. Misremembering a common object: When left is not right. *Memory & Cognition*. 1990;18(2):174-82.
37. Jones GV, Martin M. Misremembering a familiar object: mnemonic illusion, not drawing bias. *Memory & Cognition*. 1992;20(2):211-3.
38. Blake AB, Nazarian M, Castel AD. The Apple of the mind's eye: Everyday attention, metamemory, and reconstructive memory for the Apple logo. *Q J Exp Psychol (Hove)*. 2015;68(5):858-65.
39. Johnson J, Woodman G, Braun E, Luck S. Implicit memory influences the allocation of attention in visual cortex. *Psychonomic Bulletin & Review*. 2007;14(5):834-9.
40. Lepsien J, Griffin IC, Devlin JT, Nobre AC. Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. *Neuroimage*. 2005;26(3):733-43.
41. Kuo B-C, Rao A, Lepsien J, Nobre AC. Searching for targets within the spatial layout of visual short-term memory. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*. 2009;29(25):8032-8.
42. Brockmole JR, Henderson JM. Using real-world scenes as contextual cues for search. *Visual Cognition*. 2006;13(1):99-108.
43. Hollingworth A. Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition*. 2009;17(1-2):273-91.
44. Manginelli AA, Geringswald F, Pollmann S. Visual Search Facilitation in Repeated Displays Depends on Visuospatial Working Memory. *Experimental Psychology*. 2012;59(1):47-54.
45. Astle DE, Nobre AC, Scerif G. Subliminally presented and stored objects capture spatial attention. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*. 2010;30(10):3567-71.

46. Persike M, Meinhardt-Injac B, Meinhardt G. The preview benefit for familiar and unfamiliar faces. *Vision Research*. 2013;87(0):1-9.
47. Summerfield JJ, Rao A, Garside N, Nobre AC. Biasing perception by spatial long-term memory. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*. 2011;31(42):14952-60.
48. Doallo S, Patai EZ, Nobre AC. Reward associations magnify memory-based biases on perception. *Journal Of Cognitive Neuroscience*. 2013;25(2):245-57.
49. Fei-Fei L, Iyer A, Koch C, Perona P. What do we perceive in a glance of a real-world scene? *J Vis*. 2007;7(1):10-.
50. Rensink R, O'Regan J, Clark J. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*. 1997;8(5):368-73.
51. Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998;5(4):644-9.
52. Birks RJS, Gemmell L, O'Sullivan E, Rowbotham DJ, Sneyd JR. Recommendations for standards of monitoring during anaesthesia and recovery. 4 ed. London: Association of Anaesthetists of Great Britain and Ireland; 2007.
53. van Schalkwyk JM, Lowes D, Frampton C, Merry AF. Does manual anaesthetic record capture remove clinically important data? *Br J Anaesth*. 2011;107(4):546-52.
54. Lees N, Hall R. Information technology in anaesthesia and critical care. *Continuing Education in Anaesthesia, Critical Care & Pain*. 2011;11(3):104-7.
55. Department of Health. Information for Health: An information strategy for the Modern NHS 1998-2005 - a national strategy for local implementation. London: NHS Executive; 1998.

5 Inattentional Blindness and Non-Clinical Stimuli

'Vision is the art of seeing what is invisible to others'

Jonathan Swift (1)

5.1 Chapter overview

In Chapter 3 it was established that the concepts of situational awareness and non-technical skills are not universally taught to doctors. Chapter 4 then established that unfamiliar equipment can contribute to an increased risk of error when gathering information.

These alone do not, however, tell us how perceptual errors affect clinical staff. The experiment described in this chapter therefore will address this question directly. Phenomena such as inattentional blindness are little studied in medicine, and few existing data can inform how to investigate these effects in a clinical context.

The following experiment therefore lays the foundation for future work by presenting a simple replication of two of Simons's original 'Invisible Gorilla' experiments (2, 3), this time using a clinical population.

This will test the hypothesis that clinical training does not offer any generalizable advantage in perceptual ability, and will then offer invaluable information useful in the design of future experiments testing for these effects in a healthcare-trained population.

5.2 Introduction

Clinicians, particularly those employed in critical-care specialities, are often required to work in high-pressure situations that depend upon the rapid assimilation and interpretation of information from

disparate sources. The ability to perform well in these circumstances has important safety implications. There is therefore a body of opinion within (and also probably without) medicine that demands staff be effective and reliable observers (4).

Inattentional blindness and change blindness are phenomena that have been variously explored using largely non-healthcare-trained participants. Expertise modulates the effect of perceptual errors (5), so how is one to generalise these findings to specific trained populations? Healthcare professionals, and doctors in particular, might be considered trained observers. One medical curriculum has, for example, the explicit purpose *'to provide the patient with a doctor trained as an attentive listener, a careful observer, an effective communicator and a knowledgeable and capable clinician'* (6). Other curricula are replete with learning objectives that require doctors to 'detect', 'recognise', and 'identify' conditions, and that places a duty on medical educators to train junior medics in the skills of information acquisition. After such training, they may therefore be differently susceptible to perceptual error than populations without this experience.

As described in Chapter 1, a classic demonstration of inattentional blindness involves an unexpected event that occurs whilst participants watch teams passing basketballs. This has become known as the Simons 'Invisible Gorilla' experiment (2). Simons observed approximately half of participants miss the unexpected stimulus, and risk varies with the difficulty of the concentration task performed (7).

This video sequence is well-recognised in both the scientific and lay literature, and has set a pattern for using video-sequences to explore inattentional blindness in naturalistic settings. To establish a baseline for the following experiments it was determined to begin by using Simons' methods to test the incidence of perceptual error amongst healthcare-trained staff, and to use these data to enable comment on the extent to which healthcare training produces a generalizable improvement in perceptual reliability.

5.3 Methods

5.3.1 Design

This is a prospective cohort study.

5.3.2 Subjects

Study participants were recruited from delegates at a national medical-education and simulation conference, as well as from both clinical and non-clinical departments within our institution.

A convenience sample of 16 participants per group was chosen to yield groups comparable in size to the original research published by Simons (2, 3).

5.3.3 Outcome Measures

The primary outcome was the noticing rate of the critical event (the presence of the 'gorilla' character) in the displayed video sequence. Secondary outcomes were the noticing rates of other change-events incorporated into the video sequence.

5.3.4 Study Procedures

This study was approved by the University of Oxford Central Research Ethics Committee (MSD-IDREC-C1-2012-139). All participants gave informed consent prior to enrolment.

Recruits were invited to watch the original perceptual-error videos created by Simons and Chabris (Viscog Productions, Champaign, IL). Participants were tested individually, watching on a 13' Macbook Air laptop.

After watching the video, participants were asked if they had witnessed anything unexpected. Those who answered positively were asked to describe what they had seen. Those answering negatively were given a list of items, which included some false-positives, and were asked to tick any they had noticed. These items are summarised in Table 5-1, and were inspired by a list that was provided to participants in the original Simons experiment.

Questionnaire administered after completion of the video sequence(s)

Did you see any of the following in the previous video sequence?

- A person in a gorilla costume
 - A person in a pirate costume
 - A person in a rabbit costume
 - A change in colour of the background
 - A change in the number of players
 - A change in the colour of the balls
 - A change in the number of balls
-

Table 5-1: Following completion of the video(s) participants were specifically asked if they had seen items on this list

Participants who had seen this video before, or who were aware of the ‘Invisible Gorilla’ experiments from other sources, were invited to watch a second of Simons’ videos (8). Although broadly similar in format to the original, this second sequence incorporates change events in addition to the gorilla. These additional events include: a change in the number of players, a change in the background colour, and the intrusion of a character in a gorilla costume. After watching the sequence, participants were again asked for their observations and presented with the same tick-list of features.

5.3.5 Analysis

Responses were stratified by level of healthcare training, and proportions analysed using Chi-squared tests. Comparisons were drawn between the proportions of participants who were aware of the unexpected event in each video sequence. Fisher’s exact test was used where appropriate, and two-sided p -values are quoted for each test.

Continuous data were tested for normality using the Shapiro-Wilk test. Positive skew was evident for age in some groups, and negative skew was evident in pass counts reported for both videos. Since data were not normally distributed in all groups comparisons were drawn using non-parametric tests. These included the independent samples Mann-Witney U test, or independent samples Kruskal-Wallis test as appropriate.

5.4 Results

The study recruited 48 participants, comprising 12 males and 36 females. Recruits included 16 doctors, 16 nurses and 16 people from other backgrounds. This was a heterogeneous group including administrative staff, research personnel and some undergraduates, summarised in Table 5-2.

		Participants by professional group		
		Gender		Total
		Female	Male	
Professional Background	Doctor	10	6	16
	Nurse	14	2	16
	Other	12	4	16
Total		36	12	48

Table 5-2: Summary of participants by professional group and gender

Demographic data are summarised in Table 5-3. There was no significant difference in age between genders ($p = 0.28$), or backgrounds ($p = 0.75$).

Demographic data and pass counts by professional group			
	Mean (95% CI)	Median	Range (Min-Max)
Age			
Doctor	40.94 (33.3 – 48.6)	34	46 (26 – 72)
Nurse	42.00 (38.1 – 45.9)	42	23 (31 – 54)
Other	35.44 (28.2 – 42.7)	30	36 (21 – 57)
Pass Count			
Video 1 (Naïve participants)	15.58 (14.9 – 16.3)	16	7 (12 – 19)
Video 1 (Non-naïve participants)	16.04 (15.4 – 16.7)	16	8 (11 – 19)
Video 2	15.29 (14.6 – 16.0)	16	8 (11 – 19)

Table 5-3: Summary of participant age by professional background

Half of the participants (24 of 48) were unaware of the experiment and watched only video one, with 14 of the 24 (58.3%) failing to report the gorilla. These results were identical to rates of inattentive blindness observed by Simons ($p = 1.00$). Neither gender ($p = 0.05$), profession ($p = 0.53$) nor specialty ($p = 0.71$) affected noticing rate. The number of participants in each group who noticed the gorilla is summarised in Table 5-4. Only a single participant incorrectly reported a ‘false-positive’. This participant reported seeing a change in the number of players in video one, and this

participant had seen this sequence before. This participant also falsely reported seeing a change in the number of balls in video two.

Results by profession and demographic groups

Video 1: Participants unaware of the experimental method (n=24)

		Event seen	Event not seen
Gender	Male	5	1
	Female	5	13
Professional Background	Doctor	2	5
	Nurse	3	5
	Other	5	4
Specialty of professional staff	Anaesthetics/ICU	2	5
	Medicine	2	2
	Surgery	1	2
	Non-clinical	5	10

Table 5-4: Numbers of participants reporting the change event(s) after prompting in video 1

The remaining participants had seen the sequence before, and therefore watched both video one and two. Only two of these 24 (8.3%) failed to see the gorilla in sequence one, and three (12.5%) failed in video two. None of the participants who failed to notice the gorilla in video one also failed in video two. In addition to the false-positive described above, there was one additional incorrect report in video two. This was a participant who reported seeing a change in the colour of the balls during the sequence. These were the only two false-positives in video two.

Results by profession and demographic groups

Video 2: Participants aware of the experimental method (n=24)

		All events seen	Event(s) not seen
Gender	Male	3	3
	Female	4	14
Professional Background	Doctor	2	7
	Nurse	3	5
	Other	2	5
Specialty of professional staff	Anaesthetics/Intensive Care	4	8
	Medicine	1	3
	Surgery	0	1
	Non-clinical	2	5

Table 5-5: Numbers of participants reporting the change event(s) after prompting in video 2

Seven of the 24 (29.2%) noticed every change-event in sequence two, a higher proportion than observed by Simons (3) ($p = 0.05$). No significant effects on noticing rate were observed for gender (p

= 0.31), profession ($p = 0.79$), or specialty background ($p = 0.91$). These data are also summarised in Table 5-5 and Table 5-6.

There was no difference in pass counts between the group who did notice the gorilla and those who did not ($p = 0.30$ [Video one] and $p = 0.65$ [Video two]). Similarly, the ‘noticers’ did not differ in age from the ‘non-noticers’ ($p = 0.89$ [Video one] and $p = 0.09$ [Video two]).

Noticing of specific events in video 2 by professional group

Event	Professional Background	Event Noticed		Total
		Yes	No	
Gorilla	Doctor	9	0	9
	Nurse	7	1	8
	Other	5	2	7
	Total	21	3	24
Background change	Doctor	3	6	9
	Nurse	4	4	8
	Other	2	5	7
	Total	9	15	24
Player Count	Doctor	1	8	9
	Nurse	2	6	8
	Other	1	6	7
	Total	4	20	24

Table 5-6: Noticing of specific events in video 2

5.5 Discussion

These data demonstrate that clinically trained personnel can experience perceptual errors, and that in this specific example their performance is no better than that of the general public.

Specialties such as anaesthetics place a degree of importance on training staff to be effective observers, specifying that trainees must become skilled in the acquisition and maintenance of situational awareness (9). Similar requirements exist in emergency medicine (10) and intensive care curricula (11). Although other medical specialties do not explicitly prioritise training in situational awareness, it is obvious that successful clinical practice depends upon careful observation and collection of data as part of the process of diagnosis and treatment-planning.

One might therefore expect healthcare staff, as people trained to acquire large amounts of complex information, to perform to a higher standard than those without such experience. Although vulnerability to perceptual error is likely to be a ubiquitous trait, it could be hoped that healthcare training might modulate the effects of inattentional blindness. These data refute that prediction, and the consistency with which these results parallel Simons' could be somewhat surprising. There is, however, some justification for observing high rates of inattentional blindness amongst our clinical participants.

It is known that expertise plays a role in protecting against inattentional blindness (5), probably by reducing the cognitive workload necessary to perform the given task (12, 13), but these expertise effects are highly specific (14). Whilst the detection of inattentional blindness in clinical personnel could therefore be anticipated, particularly amongst those naïve participants experiencing the videos for the first time, the generally poor performance by observers of the second video may be more informative (particularly since video two incorporates both inattentional- and change-blindness probes). Video two was only seen by participants who declared they had already seen the first, but even previous exposure did not protect two observers against perceptual error in video one.

A further three observers were unable to see the gorilla in video two, even after having just seen it during playback of the first video. This is surprising because one would expect watching the first video should attune the viewer's attentional set (15) such that the gorilla becomes more task-relevant and receives attentional prioritisation. Search conditions during video two should thus be optimal, at least for this specific stimulus. The perceptual failures observed here therefore confirm that even specific foreknowledge of an event is not sufficient to eliminate the possibility of error. This degree of specific cueing would never be available in a clinical situation where it is usual to have only general predictions of the course of a patient's illness, therefore the effects of this error can only be more commonplace during emergency care.

It might also be assumed that having watched the first video and answered questions about unexpected events, participants would watch the second in a state of generally heightened vigilance. It is here that these data diverge slightly from Simons' original results. It is likely that our 'aware' group were more familiar with Simons' work than the participants in his original experiment; many in this group were educators who volunteered that they used the 'Invisible Gorilla' sequence in their own situational awareness teaching. In the present experiment, this cohort noticed the change events in video two more frequently than Simons' results would predict, however this difference barely reached significance. The low rate of false-positives observed demonstrates that guessing was not a significant issue for participants in this experiment.

Simons suggests poor performance of observers seeing video 2, which he titled 'Monkey Business', may reflect an additional processing cost associated with actively searching for the gorilla (3). Conscious search adds to cognitive workload and narrows the attentional set, thus perceptual errors may become even more likely for events that cannot be predicted. Even without a conscious decision to search a display, subconscious processing of the scene still occurs and has measurable effects on search performance. The existence of a degree of subconscious processing of unattended stimuli is easily demonstrated by the *'cocktail party effect'* (16, 17). Although best recognised as an auditory phenomenon, a visual equivalent also occurs (18). Distractors are extremely common in the clinical environment (19), so it is likely that the failures of perception observed in this experiment are highly relevant.

Unfortunately, even after attention has been captured by a target, as seen here change detection is not fool-proof, and exogenous capture of attention can still be overridden by endogenous control. Watching the first video set a context, such that although the task instructions did not explicitly state a second gorilla would be present, it is highly likely the observers anticipated this event. Their endogenous attentional control may have down-regulated exogenous capture by the other probes.

5.5.1 Object features and event detection

Participants were instructed to observe the team as they passed basketballs. Unlike video one, by changing the team's composition video two introduced changes that directly affected the target of attention. It has been demonstrated that object identification is vulnerable to error, even when that object is the primary focus (20). Even fixating on a target does not guarantee that changes will be detected, even when the moment of change is cued (21).

This form of perceptual error has been termed '*feature-based inattentional blindness*' (22). It occurs when changes to an attended task-relevant object are not perceived, and represents a situation where, although the object featured in conscious awareness, details were omitted. Many clinical events require the detection of changes affecting only a small portion of an attended item, for instance a colour change in part of a patient's body, or a change in value on a monitor display, and these data demonstrate that such changes are not robustly detected even when attending to the object in question.

Even when cognitive- and perceptual-load are not taxed it is possible for people to miss major changes in one of only two objects in a display if the second object is not relevant to the primary task (23, 24). If experimental conditions are manipulated such that both objects become relevant, participants can again report changes in each. Perceptual overload may therefore not be the only route to perceptual failure, and this has been described as '*irrelevance blindness*'. This is consistent with notions of how the attentional set functions to tune out distractors, and explains why the gorilla of video one is generally not seen.

The nature and location of changes affect the likelihood of being noticed (25, 26). Changes to objects most crucial for determining the gist of a scene are observed more quickly (27), particularly where the change is one that could plausibly occur in the real world (28). This raises important questions about the ecological validity of using implausible events as experimental stimuli, at least when considering how to apply findings to realistic activities.

The spatial relationship of targets and distractors has other important effects with implications for real-world activities. It has been observed that background changes in proximity to important foreground figures are, somewhat counter-intuitively, less likely to be noticed than more distant events (29). Faulty search-strategy may contribute to this, as observers may erroneously conclude that the immediate vicinity of a target item has been well-searched because fine detail can be recalled about the primary object. Wiseman demonstrates this effect using a well-known television programme to show how viewers focus on the main actors in a scene, and pay lesser attention to background events even though those are spatially proximate to target objects (30). Such effects were observed in the current experiment, because changes in video two were frequently missed, in particular the colour-change in the background curtain, in spite of its proximity to the main events.

Another viable explanation as to why the background changes in video two were detected less successfully than the 'gorilla' can be offered. Attention appears biased towards animate objects, an effect attributed to the evolutionary advantage conferred by more effective threat detection (31). Surprisingly this advantage may also extend to inanimate objects with related features. Privileged processing can occur for schematic diagrams representing facial expressions, and that such processing can be used to guide subsequent attention and aid change-detection (32). Similar effects on change-detection have been demonstrated using Lego® 'minifigures' as proxies for animate objects (33), and animate critical stimuli are more likely to be noticed in word-based inattentional blindness tasks, although this effect is attenuated by perceptual load (34). If changes in humanoid objects can be processed more reliably at subconscious level there may be possibilities to take advantage of this by redesigning clinical monitors, perhaps using colour-coded homunculi in place of numerical displays.

A negative corollary of this also occurs, with some potentially informative data sources receiving less attention than might otherwise be warranted. Shadows, for example, are not afforded high priority by the visual system, and even meaningful changes in shadow composition are discarded when

processing natural scenes (35), and people are similarly poor at detecting changes in reflections (36, 37). It is suggested then that the visual system prioritises tangible objects and affords lower priority to transitory features such as shadows or reflections. These processing biases may be innate. If real-world objects are more deeply processed than 'virtual' objects, such as shadows, there may be implications for monitor display design to ensure that meaningful data are not discarded when critical events occur.

Some memory encoding does occur even in the absence of explicit recall of stimuli. It is a reasonably common finding that participants can suggest 'something' has altered without necessarily being able to specify details (38). Although the current experiment was not designed to capture such data, this was certainly something that several participants expressed when describing their observations. This implies that some degree of change-detection can occur implicitly, and indeed this may occur even before a coherent mental model of the object is available (39). Such effects have been demonstrated using recordings of event-related potentials to show implicit object processing (40), and also that memory traces for pre-change objects can persist even where those memories are not consciously accessible (41). Such implicit memory encoding has, in certain situations, also been demonstrated to trigger behavioural change in the absence of explicit awareness of an event (42).

A disconnection between behavioural influences and conscious appreciation of events superficially seems to conflict with the subjective experience of everyday life, however examples are easy to provide. Most drivers have the experience of successfully operating a motor-vehicle without retaining any memory of the locations driven through, and similar phenomena can affect pedestrians (43). That behavioural change can occur without conscious appreciation of events can also be demonstrated using effects such as the Ebbinghaus illusion (Figure 17). It is known that if instructed to pick up the central disks people will correctly adjust their hand position to grip the edges of each circle, whilst simultaneously reporting that the discs differ in size. Here, then, is

another discontinuity between subjective experience and the motor commands generated by the individuals' intention to act (44), and an effect that is modulated by cognitive load (45).

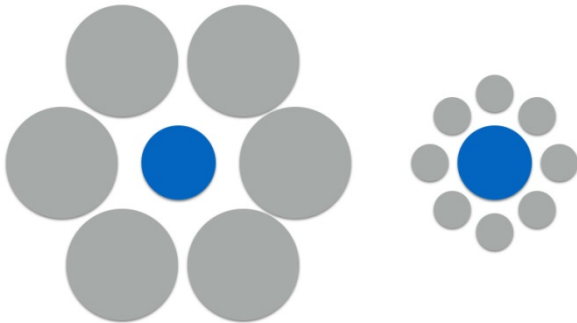


Figure 17: Ebbinghaus Illusion

Optical illusion and magic tricks are of interest when examining errors of perception. Magic tricks in particular depend on performers' deliberate manipulation of attention, and the 'science of magic' is sometimes used as an experimental model in the psychological literature (20, 46-50). Barnhart used a trick

that involved a 'disappearing' coin (in fact it was pulled across the table in plain sight) (46). Participants who previewed the trick's start-conditions were less affected by inattentional blindness and therefore better able to determine how the trick was performed. It was concluded that perceptual load was reduced by the preview condition, and since the initial view of the experimental scene was no longer novel the coin's movement was more readily detected. This may also be related to participants being better able to ignore the distracting event that the performer intended to draw attention away from the sleight of hand. If familiarity with the environment has a meaningful impact on perceptual error this could have implications for clinical practice, where emergency teams are frequently called to environments in which they do not regularly work.

5.5.2 Predicting inattentional blindness risk

It would be useful in clinical practice to be able to screen potential recruits for susceptibility to perceptual errors, and perhaps only appoint those with the most robust change-detection abilities to the most acute specialties. Unfortunately, little data exists to enable prediction of which individuals are at highest risk of perceptual error.

It is challenging to retest individuals' susceptibility to inattentional blindness, as the surprising nature of the critical stimulus is usually an important factor. A truly unexpected re-test is then

difficult to design. What evidence exists is inadequate and sometimes conflicting. Bredemeier explored predictors such as working-memory capacity using a visual and spatial '*n*-back' task, followed by a sustained-attention task involving tracking moving shapes. *N*-back tasks provide a measure of working memory capacity, and Bredemeier observed that higher working memory capacity predicted better performance in the shape-tracking task, but inattentional blindness risk was independent of memory performance (51). Conversely Richards *et al* found low working-memory capacity did predict increased risk of inattentional blindness in a similar shape-tracking activity (52). In the case of Richards, working-memory capacity was measured using an automated operation-span task whereby participants held a list of letters in memory whilst performing mathematical calculations (53). Richards also found different patterns of fixation between noticers and non-noticers; participants who missed the unexpected stimulus were significantly more likely to fixate on the distractors, and went on to suggest that they were less able to efficiently distribute their attention across the field.

More recent studies by Beanland (54) and Kreitz (55) have concluded that individual factors probably do not predict inattentional blindness risk. Kreitz observed that working-memory capacity weakly predicted performance only in certain types of static inattentional blindness task, but was not predictive of performance in divided-attention variants of the same task, nor in a dynamic task. Other factors, including breadth of attention and global/local attentional biases, were not predictive of performance in any condition. Kreitz's study is one of the few to attempt to correlate noticing between multiple inattentional-blindness tasks. They determined that there was no association between noticing in a dynamic inattentional blindness task versus noticing in a static task, and concluded that failing to notice critical stimuli is not a stable personal characteristic. The authors state that inattentional-blindness therefore appears to be determined by task-specific factors instead. These data were confirmed by a later investigation of cross-modal inattentional phenomena (56). In this experiment individuals were tested for inattentional deafness whilst performing a visual concentration task. Here again no measures of working memory correlated with task performance,

and high rates of inattentional deafness (64%) were noted as participants performed a visual discrimination task. This cross-modal failure also occurred when the tasks were reversed; 59% missed an unexpected visual stimulus when engaged by an auditory primary task. This adds to the evidence of Macdonald and Lavie that demonstrates inattentional deafness rates are modulated by the degree of visual load presented (57).

It is not just memory capacities that have been investigated in this regard; certain personality traits have been implicated in determining perceptual error susceptibility, with some claiming that the trait 'absorption' may be predictive of inattentional blindness risk (58). Kreitz expanded upon this by examining other personality traits, including openness to experience, schizotypy, and achievement motivation. Only openness was predictive, with participants who scored highly for openness being more likely to notice the events. This association was however very weak, and explained only 3% of the variance in noticing rate (59). Kreitz was unable to replicate the findings of Richards, identifying no link between absorption and perceptual error, attributing this discrepancy to differences in the tasks used in each experiment. The question of how to predict personal factors that determine perceptual error risk may therefore not be meaningful, and efforts should instead be aimed at mitigating the effects of the error.

If an individualised perceptual-error risk cannot be predicted, then examining the effect of training on perceptual reliability becomes even more important. Whether improved change detection abilities can be taught has been subjected to some research, but unfortunately only a very small number of papers suggest that it may be possible to improve performance with training. Even where successful it appears that training benefits only certain specific groups or task-types.

Some examples of a training-effect include improved change-detection amongst dysphoric individuals offered practice in performing the *n*-back concentration tasks who then performed more accurately in an array-based change detection activity (60). Improvements have also been seen in observers watching a specific type of magic trick (49), or in certain occupational settings such a

professional proof-reading (61). Regular players of video-games develop improved visual capacity (62), and mindfulness training protects working memory capacity during demanding concentration tasks (63). Such training may reduce inattentional blindness rates even after short periods (64), possibly by facilitating bottom-up processing of visual scenes thus limiting the biasing effects of top-down interference. The effect of this would be to loosen the attentional set, but the impact this would have on other cognitive processes is unknown. It could be predicted that more objects would be examined, and this would consume greater cognitive resources; performance in tasks that are highly demanding could thus be degraded.

The weight of evidence suggests however that significantly improved generalisable change-detection ability is not amenable to training (65). Performance in concentration tasks can be trained for, highlighting the expertise effect that has been described elsewhere (for example in perceptual error studies conducted amongst chess players (66), basketball players (67), and soccer players (12)), but these improvements do not transfer to either similar or dissimilar change-detection tasks.

In the medical context expertise does certainly confer other benefits; repeated and deliberate practice improves diagnostic skill (68), for example in radiograph or ECG interpretation. This is probably not, however, primarily due to enhanced perceptual ability, rather that the contents of long-term memory guide visual search in participants with previous exposure to certain stimuli (69).

5.5.3 Summary

Real-world perceptual errors have been explored using videos of simulated crimes and the effects on eyewitnesses (70), driving simulation (71), air-traffic control (72), CCTV-security (73), and medical radiology (74). The existence of perceptual errors amongst healthcare personnel in this experiment is therefore explicable, and consistent with predictions. The significance of the effect-size demonstrated by these data is however more difficult to determine. The experimental videos have no clinical relevance, and no expertise-effect applies. To what extent perceptual errors occur when delivering care remains an open question, and one that must be subject to further research. These

findings do however establish the plausibility of perceptual failure as a source of error in clinical practice.

Future research can now be directed towards establishing how widespread perceptual error is in medicine, which practitioners are at greatest risk, and under what circumstances errors are enabled. If this phenomenon exists in practice at similar rates, then it could be a major factor in determining situational awareness and therefore be a significant contributor to vigilance-related medical errors.

5.6 References

1. Swift J. *Miscellanies: Thoughts on various subjects*. London: Robert Dodsley; 1745.
2. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception*. 1999;28(9):1059-74.
3. Simons DJ. Monkeying around with the gorillas in our midst: familiarity with an inattentional-blindness task does not improve the detection of unexpected events. *i-Perception*. 2010;1(1):3-6.
4. Vincent C. *Patient Safety*. 2nd ed. Oxford: Blackwell; 2010.
5. Memmert D. The effects of eye movements, age, and expertise on inattentional blindness. *Consciousness and Cognition*. 2006;15(3):620-7.
6. Joint Royal Colleges of Physicians Training Board. *Specialty Training Curriculum for Audiological Medicine*. London: Royal College of Physicians; 2007.
7. Simons DJ, Jensen MS. The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review*. 2009;16(2):398-403.
8. Simons DJ. *The Monkey Business Illusion*: YouTube.com; 2010 [updated 10/5/2010. Available from: www.youtube.com/watch?v=IGQmdoK_Zfy].
9. Curriculum Development Working Group. *Curriculum for a CCT in Anaesthetics*. 2nd ed. London: The Royal College of Anaesthetists; 2010.
10. The College of Emergency Medicine. *Curriculum and assessment systems for core specialty training ACCS CT1-3 & higher specialty training ST4-6 training programmes*. Glucksman E, editor. London: College of Emergency Medicine; 2010.
11. The Faculty of Intensive Care Medicine. *Curriculum for a CCT in Intensive Care Medicine*. 3rd ed: The Faculty of Intensive Care Medicine; 2011.
12. Werner S, Thies B. Is 'change blindness' attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*. 2000;7(1-3):163-73.
13. White CB, Caird JK. The blind date: the effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accident; analysis and prevention*. 2010;42(6):1822-30.
14. Memmert D, Simons DJ, Grimme T. The relationship between visual attention and expertise in sports. *Psychology of Sport and Exercise*. 2009;10(1):146-51.
15. Most SB, Scholl BJ, Clifford ER, Simons DJ. What you see is what you set: sustained inattention blindness and the capture of awareness. *Psychological Review*. 2005;112(1):217-42.
16. Wood N, Cowan N. The Cocktail Party Phenomenon Revisited: How Frequent Are Attention Shifts to One's Name in an Irrelevant Auditory Channel? *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1995;21(1):255-60.
17. Moray N. Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*. 1959;11(1):56-60.
18. Mack A, Pappas Z, Silverman M, Gay R. What we see: Inattention and the capture of attention by meaning. *Consciousness and Cognition*. 2002;11(4):488-506.
19. Campbell G, Arfanis K, Smith AF. Distraction and interruption in anaesthetic practice. *Br J Anaesth*. 2012;109(5):707-15.
20. Smith TJ, Lamont P, Henderson JM. The penny drops: change blindness at fixation. *Perception*. 2012;41(4):489-92.
21. Smith TJ, Lamont P, Henderson JM. Change blindness in a dynamic scene due to endogenous override of exogenous attentional cues. *Perception*. 2013;42(8):884-6.
22. Persuh M, Gomez M, Bauer L, Melara RD. Feature-based inattentional blindness: loss of awareness to featural information in fully attended objects. *Attention Perception & Psychophysics*. 2014;76(8):2229-39.
23. Eitam B, Yeshurun Y, Hassan K. Blinded by irrelevance: Pure irrelevance induced "blindness". *Journal of Experimental Psychology: Human Perception and Performance*. 2013;39(3):611-5.

24. Eitam B, Shoval R, Yeshurun Y. Seeing without knowing: task relevance dissociates between visual awareness and recognition. *Annals of the New York Academy of Sciences*. 2015;1339:125-37.
25. Lin SH, Yeh YY. Attentional load and the consciousness of one's own name. *Consciousness & Cognition*. 2014;26:197-203.
26. Beck DM, Lavie N. Look here but ignore what you see: effects of distractors at fixation. *J Exp Psychol Hum Percept Perform*. 2005;31(3):592-607.
27. Saiki J, Holcombe AO. Blindness to a simultaneous change of all elements in a scene, unless there is a change in summary statistics. *Journal of Vision*. 2012;12(3).
28. Beck MR, Angelone BL, Levin DT. Knowledge About the Probability of Change Affects Change Detection Performance. *Journal of Experimental Psychology: Human Perception & Performance*. 2004;30(4):778-91.
29. Porubanova-Norquist M, Sikl R. Various categories of changes in the inducement of change blindness. *Studia Psychologica*. 2013;55(2):83-94.
30. Wiseman R, Watt C. And now for something completely different: Inattentional blindness during a Monty Python's Flying Circus sketch. *i-Perception*. 2015;6(1):38-40.
31. New J, Cosmides L, Tooby J. Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*. 2007;104(42):16598-603.
32. Lyyra P, Makela H, Hietanen JK, Astikainen P. Implicit binding of facial features during change blindness. *PLoS ONE*. 2014;9(1).
33. LaPointe MR, Cullen R, Baltaretu B, Campos M, Michalski N, Sri Satgunarajah S, et al. An Attentional Bias for LEGO People Using a Change Detection Task: Are LEGO People Animate? *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*. 2016:No Pagination Specified.
34. Calvillo DP, Jackson RE. Animacy, perceptual load, and inattentional blindness. *Psychonomic Bulletin & Review*. 2014;21(3):670-5.
35. Ehinger KA, Allen K, Wolfe JM. Change blindness for cast shadows in natural scenes: Even informative shadow changes are missed. *Attention, Perception, & Psychophysics*. 2016:No Pagination Specified.
36. Sareen P, Ehinger KA, Wolfe JM. Through the looking-glass: Objects in the mirror are less real. *Psychonomic Bulletin & Review*. 2014;22(4):980-6.
37. Chesney DL, Gelman R. What counts? Visual and verbal cues interact to influence what is considered a countable thing. *Memory & Cognition*. 2015;43(5):798-810.
38. Hughes HC, Caplovitz GP, Loucks RA, Fendrich R. Attentive and pre-attentive processes in change detection and identification. *PLoS ONE*. 2012;7(8).
39. Ball F, Busch NA. Change detection on a hunch: pre-attentive vision allows "sensing" of unique feature changes. *Attention Perception & Psychophysics*. 2015;77(8):2570-88.
40. Ball F, Bernasconi F, Busch NA. Semantic Relations between Visual Objects Can Be Unconsciously Processed but Not Reported under Change Blindness. *Journal of Cognitive Neuroscience*. 2015;27(11):2253-68.
41. Busch NA. The fate of object memory traces under change detection and change blindness. *Brain Research*. 2013;1520:107-15.
42. Gabay S, Avni D, Henik A. Reflexive orienting by central arrows: Evidence from the inattentional blindness task. *Psychonomic Bulletin and Review*. 2012;19(4):625-30.
43. Hyman IE, Jr., Sarb BA, Wise-Swanson BM. Failure to see money on a tree: inattentional blindness for objects that guided behavior. *Frontiers in Psychology*. 2014;5:356.
44. Aglioti S, DeSouza JF, Goodale MA. Size-contrast illusions deceive the eye but not the hand. *Current biology : CB*. 1995;5(6):679-85.
45. de Fockert JW, Wu S. High working memory load leads to more Ebbinghaus illusion. *European Journal of Cognitive Psychology*. 2009;21(7):961-70.
46. Barnhart AS, Goldinger SD. Blinded by magic: eye-movements reveal the misdirection of attention. *Frontiers in Psychology*. 2014;5:1461.

47. Kuhn G, Findlay JM. Misdirection, attention and awareness: Inattentional blindness reveals temporal relationship between eye movements and visual awareness. *Quarterly Journal of Experimental Psychology*. 2010;63(1):136-46.
48. Kuhn G, Tatler BW. Misdirected by the gap: The relationship between inattentional blindness and attentional misdirection. *Consciousness and Cognition*. 2011;20(2):432-6.
49. Rieiro H, Martinez-Conde S, Macknik SL. Perceptual elements in Penn & Teller's "Cups and Balls" magic trick. *PeerJ*. 2013;2013(1).
50. Smith TJ. The role of audience participation and task relevance on change detection during a card trick. *Frontiers in Psychology*. 2015;6:13.
51. Bredemeier K, Simons DJ. Working memory and inattentional blindness. *Psychonomic Bulletin and Review*. 2012;19(2):239-44.
52. Richards A, E MH, Vitkovitch M. Distracted by distractors: Eye movements in a dynamic inattentional blindness task. *Consciousness and Cognition*. 2012;21(1):170-6.
53. Unsworth N, Heitz R, Schrock J, Engle R. An automated version of the operation span task. *Behavior Research Methods*. 2005;37(3):498-505.
54. Beanland V, Chan EHC. The relationship between sustained inattentional blindness and working memory capacity. *Attention, Perception, & Psychophysics*. 2016:No Pagination Specified.
55. Kreitz C, Furley P, Memmert D, Simons DJ. Inattentional blindness and individual differences in cognitive abilities. *PLoS ONE*. 2015;10(8).
56. Kreitz C, Furley P, Simons DJ, Memmert D. Does working memory capacity predict cross-modally induced failures of awareness? *Consciousness & Cognition*. 2016;39:18-27.
57. Macdonald JSP, Lavie N. Visual perceptual load induces inattentional deafness. *Attention, Perception, and Psychophysics*. 2011;73(6):1780-9.
58. Richards A, Hellgren MG, French CC. Inattentional blindness, absorption, working memory capacity, and paranormal belief. *Psychology of Consciousness: Theory, Research, and Practice*. 2014;1(1):60-9.
59. Kreitz C, Schnuerch R, Gibbons H, Memmert D. Some see it, some don't: Exploring the relation between inattentional blindness and personality factors. *PLoS ONE*. 2015;10(5).
60. Owens M, Koster EHW, Derakshan N. Improving attention control in dysphoria through cognitive training: Transfer effects on working memory capacity and filtering efficiency. *Psychophysiology*. 2013;50(3):297-307.
61. Asano M, Kanaya S, Yokosawa K. Proofreaders Show A Generalized Ability To Allocate Spatial Attention To Detect Changes. *Psychologia*. 2008;51(2):126-41.
62. Green CS, Bavelier D. Action video game modifies visual selective attention. *Nature*. 2003;423(6939):534-7.
63. Jha AP, Stanley EA, Kiyonaga A, Wong L, Gelfand L. Examining the Protective Effects of Mindfulness Training on Working Memory Capacity and Affective Experience. *Emotion*. 2010;10(1):54-64.
64. Schofield TP, Creswell JD, Denson TF. Brief mindfulness induction reduces inattentional blindness. *Consciousness & Cognition*. 2015;37:63-70.
65. Gaspar JG, Neider MB, Simons DJ, McCarley JS, Kramer AF. Change Detection: Training and Transfer. *PLoS ONE*. 2013;8(6).
66. Reingold EM, Charness N, Pomplun M, Stampe DM. Visual span in expert chess players: evidence from eye movements. *Psychological science : a journal of the American Psychological Society / APS*. 2001;12(1):48-55.
67. Furley P, Memmert D, Heller C. The dark side of visual awareness in sport: Inattentional blindness in a real-world basketball task. *Attention, Perception, and Psychophysics*. 2010;72(5):1327-37.
68. Ericsson KA. Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. *Academic Medicine Research in Medical Education Proceedings of the Forty-Third Annual Conference November 7-10, 2004*. 2004;79(10):S70-S81.

69. Beck MR, Martin BA, Smitherman E, Gaschen L. Eyes-On Training and Radiological Expertise: An Examination of Expertise Development and Its Effects on Visual Working Memory. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2013;55(4):747-63.
70. Fitzgerald RJ, Oriet C, Price HL. Change blindness and eyewitness identification: Effects on accuracy and confidence. *Legal and Criminological Psychology*. 2016;21(1):189-201.
71. Pammer K, Bink C. Attentional differences in driving judgments for country and city scenes: Semantic congruency in inattentional blindness. *Accident; analysis and prevention*. 2013;50:955-63.
72. O'Brien KS, O'Hare D. Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*. 2007;50(7):1064-91.
73. Parasuraman R, de Visser E, Clarke E, McGarry WR, Hussey E, Shaw T, et al. Detecting Threat-Related Intentional Actions of Others: Effects of Image Quality, Response Mode, and Target Cuing on Vigilance. *Journal of Experimental Psychology: Applied*. 2009;15(4):275-90.
74. Drew T, Võ ML-H, Wolfe JM. The Invisible Gorilla Strikes Again: Sustained Inattentional Blindness in Expert Observers. *Psychological Science*. 2013;24(9):1848-53.

6 Inattentional Blindness and Clinical Stimuli

'...humans are fallible, and systems must be designed so that humans are prevented from making errors.'

van Beuzekom (1)

6.1 Chapter overview

It has now been established that clinicians are vulnerable to the effects of perceptual error, and that the training they receive does not confer any generalizable improvement in perceptual ability. Questions remain over how these risks affect staff when they are working within their field of expertise.

Expertise effects have been observed by groups using the Simons video amongst sportspeople, so it is plausible to expect that clinicians may not be as vulnerable to perceptual error within their own field as the experiment described in Chapter 5 might suggest.

The study described here has been designed to answer that question, using a bespoke video developed specifically for this purpose and inspired by the techniques used by Simons. This therefore offers an important opportunity to test for the first time the hypothesis that clinicians are vulnerable to inattentional- and change-blindness, even when viewing materials that are relevant to their workplace, and that less experienced clinicians would be more affected by perceptual error than subject-experts.

The aims of this experiment are therefore two-fold. Firstly, to determine if clinicians can be affected by perceptual error and secondly, to determine if this effect is modulated by expertise.

6.2 Introduction

The high frequency and impact of medical errors has been previously described (2-4), and it has been established that most errors in medicine arise from failures in non-technical skills, notably loss of situational awareness (5-8). Failure to achieve basic level one awareness (acquisition of relevant information) is responsible for up to three-quarters of situational-awareness related errors (9), and in road-traffic situations it has been noted that *'...improper allocation of attention constitutes the main collision factor leading to most accidents, due to a delay or absence in detecting danger'* (10). A common type of road-traffic accident can be described as 'looked-but-failed-to-see'. Such situations occur where a driver has appeared to check the road, but fails to notice a relevant hazard (thus experiencing a perceptual error). This may be the third most common type of traffic accident (11). If such failures are so common in a task as mundane as driving, one must then ask what risks do clinicians face?

The previous experiment established that clinicians possess no generalisable advantage in observational ability. An expectation therefore exists that perceptual errors represent a significant impediment to the achievement of situational awareness. Unfortunately, difficulties arise in the interpretation of these data, as the effect of expertise cannot be properly accounted for. Expertise has been shown to affect noticing of unexpected events (albeit with a very specific definition of 'expertise'), but our sample had no specific expertise with the experimental stimuli to draw upon. Testing clinicians using non-clinical stimuli therefore tells us only that generalised observational ability is not improved by clinical training, and cannot offer much insight into how perceptual errors affect the practice of medicine. Suitable experimental materials do not yet exist, so it is also required to develop new stimuli and experiments.

Defining expertise in a medical context is not entirely straightforward. A clinician's scope of practice, and the frequency with which particular situations are encountered, are at least as important as traditional metrics such as clinician grade or years of practice. This meant a setting for the new video

had to be chosen that would be instantaneously recognisable to most healthcare providers, and resuscitation was considered as a potentially suitable subject (12).

Most NHS employers require their staff to undergo regular update sessions in life support skills as part of their continuing professional development and appraisal process, and all doctors are required to render assistance in an emergency. National standards require resuscitation training to be delivered at least annually, and require front-line clinical staff in acute-care organisations to be able to deliver defibrillation where required within three minutes of collapse (13). Similar expectations broadly apply to primary care providers (14). Also, undergraduates are required to receive training in resuscitation and students must have been assessed as competent in resuscitation prior to graduation (15). Resuscitation is therefore one of the few areas of medicine that one can reasonably expect a universal degree of familiarity, making it a suitable topic for this study.

Definitions of resuscitation expertise could be variously based upon level of training, frequency of attendance at emergencies, seniority of clinician, or specialty of practice (as a proxy for the likelihood of encountering critically ill patients). Expertise might also be defined in more academic terms, such as familiarity with source materials and guidelines, prominence on resuscitation committees, or resuscitation-related research activities. No data exist to guide the choice; therefore a pragmatic decision was taken for the purposes of this study to stratify expertise by the highest level of training in adult life support that the clinician had received. This was a simple, objective, and self-reportable measure that permitted easy classification of participants and possessed good face-validity. Confounding variables, such as number of real-life experiences of resuscitation situations and specialty of practice were recorded, and were therefore available to confirm that practical experience correlated with training level.

This experiment then offers an important opportunity to test clinicians for perceptual error using stimuli with which expertise can be demonstrated. This was the first experiment to test for the

effects of inattentional blindness and change blindness using clinically-relevant stimuli in a dynamic task (16).

6.3 Methods

6.3.1 Design

This is a prospective cohort study.

6.3.2 Subjects

The study was open to medical undergraduates and healthcare-trained post-graduates. Participants were approached by email invitation, and were also free to pass the invitation on to colleagues at their own discretion. Volunteers were largely drawn from the University of Oxford and staff of the NHS South Central region.

The sample size was calculated to detect an effect size of 0.3, a level comparable to the expertise-effect demonstrated amongst athletes (17), and based on $\alpha = 0.05$ and $\beta = 0.8$. At least 105 participants were therefore required.

6.3.3 Outcome Measures

The primary outcome was the proportion of participants in each experience-group noticing the key event of oxygen disconnection in the video sequence. Secondary outcome measures were the proportion of participants noticing four other change- or inattentional blindness probes included in the sequence.

6.3.4 Study Procedures

This study was approved by the University of Oxford Central Research Ethics Committee (MSD-IDREC-C1-2012-139). All participants gave informed consent prior to enrolment.

This was designed as a web-based experiment. Study materials were hosted by SurveyMonkey (SurveyMonkey Europe SARL, Luxembourg; www.surveymonkey.com). Participants completed the survey in their own time and on their own computers. The survey began by collecting basic demographic data, including age and gender. Participants were asked to specify their highest level of resuscitation training, and to estimate their number of resuscitations attended annually.

Participants were instructed to imagine themselves in the role of a resuscitation team leader. They were informed that they were about to view a video of a simulated adult resuscitation, and asked to watch the team's performance carefully. They were told to expect questions about the conduct of the resuscitation after viewing the video.

Before playback began, participants were reminded to ensure that their computer volume was audible and set to a comfortable level, and instructions were presented textually. Questions were placed on-screen individually, and progress was self-paced.

Participants were shown a 50-second video created specifically for this experiment, which depicted a simulated adult resuscitation in progress into which was inserted several probes testing for perceptual error. Development of the video is described in Appendix 10.2.6. Participants initiated playback by clicking a button on-screen after reading their instructions.

Immediately after watching the video participants were invited to leave free-text on the sequence in response to the question 'please make any comments you wish about the team's performance'. They were then asked to report if they noticed anything unusual. If they answered in the affirmative, they were asked, again by free-text, to describe what they had observed. After collecting the unprompted free-text feedback, a series of specific events were provided. These were patterned after Simons' original gorilla experiment (18), and included a selection of false choices. From this list (Table 6-1) participants were required to tick any events they had noticed. Finally, each was asked at the end of the study if there were any technical difficulties encountered; those that reported problems in video playback were excluded from analysis.

After answering all questions participants were asked a final question to determine familiarity with the phenomenon of perceptual error, and whether they had previously seen the Simons 'Invisible Gorilla' video.

Event choices presented to participants at the conclusion of the video sequence

'Did you see any of the following in the previous video sequence? (Please tick any that apply)'

- 1) A change in the number of team members
- 2) The airway person changed
- 3) The CPR person changed
- 4) The cardiac rhythm changed
- 5) The patient's airway changed
- 6) The defibrillator malfunctioned
- 7) The oxygen malfunctioned
- 8) The team at the end were wearing different coloured tops
- 9) The team at the end were wearing different coloured hats
- 10) The airway person had acquired some new equipment by the end of the scene
- 11) The team leader had acquired some new equipment by the end of the scene

Table 6-1: List of change events from which participants were prompted to choose

6.3.5 Analysis

For the purposes of analysis participants were grouped according to resuscitation training experience. Participants with no training, or trained only to the level of basic life support (BLS) were assigned to the inexperienced group. Accredited advanced life support (ALS) providers were assigned to the experienced group. Those qualified to instruct ALS were designated experts.

Normality was assessed using the Shapiro-Wilk test. The continuous variables (age and number of resuscitations attended annually) were not normally distributed so non-parametric comparisons were made using the Kruskal-Wallis H test, and pairwise comparisons performed with Dunn's procedure and a Bonferroni correction. Chi-squared analysis (or Fisher's exact test where indicated) was employed to compare proportional data. Two-sided *p*-values are quoted for each test. Binary logistic regression was planned to explore the relationship between predictors of noticing rate.

6.4 Results

The study recruited 142 participants, comprising 72 males and 70 females. The professional background of volunteers and their resuscitation background are summarised in Table 6-2 and Table 6-3.

		Professional Background (n)					Total
		Doctor	Medical Student	Nurse	Resuscitation Officer	Non-clinical	
Level of Resuscitation Training	Basic/None	5	48	0	0	3	56
	Advanced	25	16	2	0	0	43
	Expert	22	0	6	15	0	43

Table 6-2: Summary of participants by professional background and resuscitation experience

		Participant demographics		Median and Interquartile Range			
		Males	Females	Age		Resuscitations annually	
Level of Resuscitation Training	Basic/None	26	30	22.5	20-24	0	0-1
	Advanced	28	15	31	24-44	2	0-6
	Expert	18	25	40	34.5-45	25	10-40

Table 6-3: Summary of continuous variables grouped by resuscitation experience

Age and number of resuscitations attended annually were compared by Kruskal-Wallis H test. This highlighted significant differences between groups ($p < 0.01$), thus pairwise comparisons were performed. The basic-trained group was both younger ($p < 0.01$) and had attended fewer resuscitations ($p < 0.01$) than the other two groups. Advanced and expert providers did not differ significantly in age ($p = 0.59$), but experts had attended significantly more resuscitations annually than either basic- or advanced-providers ($p < 0.01$).

The frequency with which participants spontaneously described noticing each of the events in their free-text responses are summarised in Table 6-4. There was a significant association only between experience and noticing the change in the patient’s airway.

Event noticing rates based on free-text responses, by participant experience						
	Noticing Rates				p value	
	Overall	Level of Resuscitation Training				
		Basic/None (n=56)	Advanced (n=43)	Expert (n=43)		
	Hat change	2 (1.4%)	0 (0.0%)	1 (2.3%)	1 (2.3%)	0.51
	Stethoscope change	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	N/A
Event	CPR provider switch	15 (10.6%)	3 (5.4%)	5 (11.6%)	7 (16.3%)	0.21
	Airway device switch	19 (13.4%)	2 (3.6%)	3 (2.1%)	14 (32.6%)	< 0.01
	Oxygen disconnection	11 (7.7%)	2 (3.6%)	3 (2.1%)	6 (14.0%)	0.16

Table 6-4: Summary of participants who, in free-text responses, described having noticed each event

Binary logistic regression was performed to explore the interaction between age, gender, resuscitation experience, foreknowledge of the inattentive blindness paradigm, specialty of practice, and number of resuscitations attended annually in predicting noticing of the events for which a significant association existed.

The logistic regression model for noticing the change in airway was statistically significant ($p < 0.01$; Nagelkerke $R^2 = 0.22$). The only significant predictor variable was resuscitation experience ($p < 0.01$), with experts being significantly more likely to notice the change (OR 13.03; 95%CI = 2.77 – 61.34).

Responses to the prompted list of changes are summarised in Table 6-5 and Figure 6-1. There was a significant association between experience and reporting the change in CPR provider, the change in the patient's airway, and the change in the stethoscope. There was no significant association between experience and noticing the oxygen disconnection or the change in hat colour.

False positives events were reported by 31 participants (21.8%). There was no association between reporting a false positive and professional background ($p = 0.86$) or resuscitation experience ($p = 0.74$). These data are summarised in Table 6-6. Of participants who reported a false positive, 24 reported one, six reported two, and a single participant reported three events. Again, there was no association with professional background ($p = 0.94$) or level of resuscitation training ($p = 0.73$).

Event noticing rates based on prompted responses, by participant experience

	Noticing Rates				p value
	Overall	Level of Resuscitation Training			
		Basic/None (n=56)	Advanced (n=43)	Expert (n=43)	
Hat change	11 (7.7%)	3 (5.4%)	5 (11.6%)	3 (7.0%)	0.50
Stethoscope change	24 (16.9%)	5 (8.9%)	7 (16.3%)	12 (27.9%)	0.04
Event CPR provider switch	36 (25.4%)	8 (14.3%)	14 (32.6%)	14 (32.6%)	0.05
Airway device switch	34 (23.9%)	7 (12.5%)	10 (23.3%)	17 (39.5%)	0.01
Oxygen disconnection	34 (23.9%)	11 (19.6%)	10 (23.3%)	13 (30.2%)	0.47

Table 6-5: Summary of participants who, on prompting, described having noticed each event

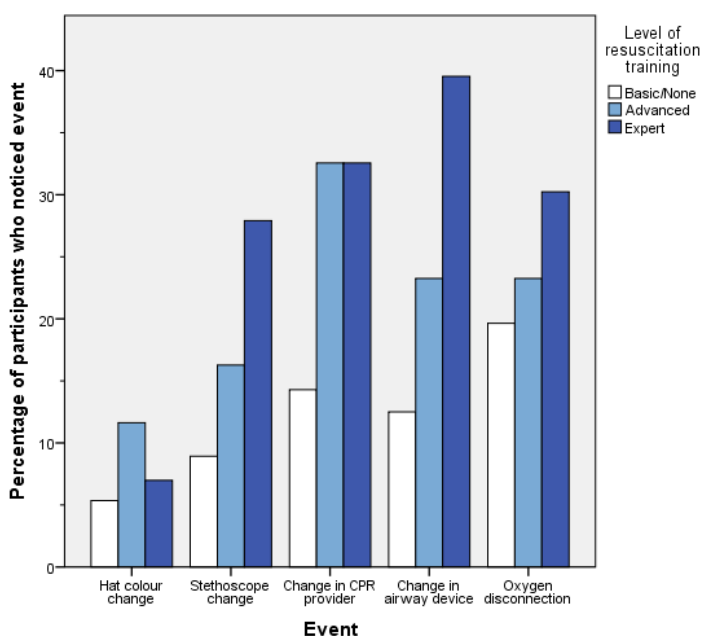


Figure 6-1: Graphical depiction of noticing rates for each event grouped by level of expertise

Participant prompted responses to false-positive items

	Reported	Not reported
Change in the number of team members	8	134
Change in the rhythm	9	133
Defibrillator malfunction	8	134
Top change	0	142
Leader equipment change	10	132

Table 6-6: Responses to false-positive items

Binary logistic regression was performed to explore the interaction between age, gender, resuscitation experience, foreknowledge of the inattentional blindness paradigm, specialty of

practice, and number of resuscitations attended annually in predicting noticing of the event for which a significant association existed.

As was seen in free-text responses, the logistic regression model for noticing the change in airway was statistically significant ($p = 0.02$; Nagelkerke $R^2 = 0.14$). The only significant predictor variable was resuscitation experience ($p=0.01$), with experts being significantly more likely to notice the change (OR 9.02; 95%CI = 2.14 – 37.98).

The logistic regression models for noticing the change in CPR provider ($p = 0.50$) and stethoscope change were not significant ($p = 0.25$). In both cases a significant model could be created using resuscitation expertise as the sole predictor. With regard to the change in CPR provider this limited model was only weakly predictive for noticing the switch ($p = 0.04$; Nagelkerke $R^2 = 0.06$), as was the case for the stethoscope change ($p = 0.05$; Nagelkerke $R^2 = 0.07$).

6.4.1 Existing knowledge of the experimental paradigm

After watching the video sequence and answering all questions, participants were finally asked if they were aware of 'inattentional blindness' or of the 'Invisible Gorilla' experiment. Several respondents offered no answer to this question ($n = 42$). The participants who responded did not differ in age ($p = 0.25$), gender ($p = 0.53$), professional background ($p = 0.49$), or level of resuscitation training ($p = 0.27$) from those who left no response.

Of the 100 participants who offered an answer, more than three quarters were already aware of the phenomenon (Table 6-7). Participants aware of the effect were no more likely to notice the unexpected events ($p = 0.60$), and expressed similar levels of surprise at the magnitude of the effects that they missed ($p = 0.61$).

Participant awareness of inattentional blindness or the ‘invisible gorilla’ experiment(s)

		Awareness of perceptual error (n = 142)			Degree of surprise at events that were not noticed (n = 100)				
		Aware	Unaware	Not answered	Totally Unsurprised	Somewhat Unsurprised	Neither Surprised nor unsurprised	Somewhat Surprised	Totally Surprised
Level of Resuscitation Training	Basic/None	29 (20.4%)	9 (6.3%)	18 (12.6%)	6 (6.0%)	4 (4.0%)	3 (3.0%)	16 (16.0%)	9 (9.0%)
	Advanced	25 (17.6%)	6 (4.2%)	12 (8.5%)	3 (3.0%)	11 (11.0%)	5 (5.0%)	7 (7.0%)	5 (5.0%)
	Expert	23 (16.2%)	8 (5.6%)	12 (8.5%)	6 (6.0%)	7 (7.0%)	5 (5.0%)	11 (11.0%)	2 (2.0%)
	Total	77 (54.2%)	23 (16.2%)	42 (29.6%)	15 (15.0%)	22 (22.0%)	13 (13.0%)	34 (34.0%)	16 (16.0%)

Table 6-7: Summary of experimental participants aware of the inattentional blindness paradigm, and the responses of participants when questioned as to their surprise at any changes they failed to notice

6.5 Discussion

It has been said that ‘one of the principle duties of a clinician is to detect and identify abnormalities’ (19). While it might be easy to dismiss some perceptual errors as camera trickery or stage-magic techniques, these data demonstrate for the first time that perceptual errors can and do affect healthcare staff. They can be demonstrated in dynamic situations, and can be triggered using stimuli directly relevant to clinical practice. Furthermore, they support the expectation that perceptual errors are prevalent even amongst expert care providers. These phenomena therefore directly impact on this primary duty of a clinician.

The video depicts events deliberately chosen to be of differing importance to clinical decision-making. Changes in clothing colour are seldom relevant to real practice, and are implausible within the context of an emergency situation such as this. At the other end of the spectrum lies the failure in the patient’s oxygen supply, something that would significantly impair the likelihood of successful resuscitation. The other events depicted fall somewhere between these extremes. The stethoscope, for example, is a device important to performing an adequate assessment of the patient, but its location is not immediately task relevant in this video. The switch in team-member compressing the chest is somewhat more relevant, insofar as the team leader should monitor their team’s activities to account for accumulating fatigue amongst providers, and the means of airway management has a direct impact on how CPR should be delivered to the patient (20). It was predicted that the least

relevant changes would be those noticed least frequently, and that is indeed the pattern observed in these data³.

The second prediction was that the expert group would demonstrate more successful event-detection than the novices, as the least experienced observers should be under the highest mental workload during the task. This expertise-effect was not, in fact, observed in all cases. Expertise played little part in the detection of the hat-change, and although there was a trend towards expertise modulating the noticing of the oxygen disconnection this did not reach significance.

The lack of expertise-effect in noticing the hat-change might be explained by its irrelevance to the primary task, as experience conferred no advantage in its detection. Furthermore, probable changes (those that are plausible or at least possible in a real-world scenario) are more readily detected than improbable events (21, 22); clothing should be invariate over the timescale depicted in this video, making a sudden change in attire highly implausible by this definition.

Attentional sets are highly relevant to detection during inattentional-blindness experiments, and it is noteworthy that these sets extend beyond physical features, incorporating the semantic categorisation of attended items (23). Moreover this categorisation effect may be independent of perceptual load (24). These data are important, because it means that semantic category can facilitate the detection of items that share no common physical features with attended targets (25). Hollingworth described this in terms of how semantically informative an object was, noting that changes to objects deemed important to interpreting the scene were more readily detected (26). Similarly, Simons observed in a real-world change-blindness experiment involving person-

³ Note that although the oxygen failure was the most clinically relevant event, this was depicted in a way to test for inattentional blindness, whereas the other events all probed for change-blindness. The fact this was not the most frequently observed event might therefore be explained by the different nature of the stimulus.

substitution that gist was highly important. Observers did not appear to record rich detail about the person with whom they were interacting, instead storing only top-level category information (27). Scene categorisation can happen within just a few hundred milliseconds of exposure to a new view (28), orientating first to the semantic categorisation of global information (29). It appears that changes affecting this high-level interpretation of objects or scenes facilitate detection, and that experts are more sensitive to these effects than novices (30). Levin suggests that violations of top-level categorisation may better prompt observers to compare the pre- and post-change scenes, facilitating change-detection (31). It is therefore possible to manipulate detection rates of critical stimuli using more complex conceptual information, and this may go some way to explaining why the airway change showed the strongest expertise effect.

One might reasonably predict this from a clinical perspective, as the airway-change necessitates an alteration in the way that CPR is delivered, moving from a 30:2 ratio to asynchronous compressions (20), and this is certainly something that the experienced and expert observers would be expected to recognise. From a psychological perspective this event could also be considered a category violation, moving from a starting point where no airway device is inserted to the patient having been intubated. A tracheal tube is considered a 'definitive airway' since it allows the development of higher inspiratory pressures and secures the respiratory tract against the possibility of aspiration, unlike mask ventilation. This extra contextual information may therefore have afforded observers in the current experiment a better chance to detect the switch. Swapping the airway in this manner changed the meaning of the scene for the experts, but not necessarily for novices unaware of these details.

The weaker than predicted effect of expertise on detecting the oxygen malfunction is somewhat more problematic. More than three quarters of the observers overall did not notice this malfunction, and experts did not significantly outperform the less experienced groups. A functioning oxygen supply is a critical piece of equipment during adult cardiac arrest situations, and the

application of high-flow oxygen is a pass/fail point during summative assessments at the conclusion of all adult resuscitation courses, on which the experts are specifically trained to examine (32).

This may represent a genuinely surprising result, but it is also possible that it occurred due to an artefact of the testing stimuli which advantaged the novice participants. This was the only inattentional blindness probe, and was associated with both a visual event, during disconnection, and a continuing audible change. Whereas all the other events included were one-shot change-blindness probes, this event was uniquely sustained from onset to the conclusion of the scene approximately ten seconds later. It is plausible that the experts may have noticed the event sooner than the less experienced groups, but that the continuous nature of the audible change gave the novices repeated opportunities to detect it. Longer exposure time has been found to facilitate detection of inattentional blindness stimuli in laboratory settings making this a plausible explanation (33). Unfortunately, the study was not designed to capture response-time data so it is not possible to comment upon whether experts did notice more rapidly.

It could be argued that other explanations, such as search satisfaction, would account for the weaker than predicted performance of the experts (34). The oxygen disconnection was the last event in the sequence, occurring very near to the end, and the observers may have stopped searching for new occurrences by this stage, so this is not an unreasonable postulation. The data however do not support this; only three of the 11 inexperienced observers who noticed the oxygen disconnection also reported seeing another event, compared to seven of 13 in the expert group. This suggests that if search satisfaction occurred, it affected the inexperienced observers to a greater extent.

Despite weaker than expected performance, expertise did remain the only predictor of inattentional blindness detected in this dataset. No association could be demonstrated between noticing events and age, gender, professional background, or foreknowledge of the inattentional blindness paradigm. These data suggest that a high proportion of clinicians are aware, at least at a basic level, of the phenomenon of inattentional blindness. More than half of the observers reported having

encountered Simons' 'Invisible Gorilla' sequence before, but that this group did not perform any better than the non-exposed group suggests that simply raising awareness of the phenomena is not sufficient to protect staff against their effects.

Although this is the first experiment to test for perceptual error using dynamic stimuli, other groups have considered the clinical implications of perceptual error with variable success (35, 36). Many of these existing reports lack sufficient detail to truly judge if perceptual error is a meaningful factor. Often other cognitive errors, such as a 'framing effect' or 'confirmation bias' (34) may be more likely contributors.

Some of the most robust work has been conducted in radiology, where accurate image interpretation has particular relevance. Drew for example observes that novice radiologists search images less efficiently than experienced staff (37), and has also used a modified version of the 'Invisible Gorilla' to test for inattentional blindness in this group (38). Drew inserted the silhouette of a gorilla into a series of slices taken from a CT chest scan, and asked observers to search the doctored images for pulmonary nodules (each considerably smaller than the gorilla image). Although an interesting experiment, it is not clear that this task definitively describes inattentional blindness. The result could be explained as search satisfaction, and while one cannot argue against the notion that a gorilla-shape in a CT scan is unexpected, one might reasonably ask why should it be relevant that it was missed? Perceptual-error risk is sensitive to the way in which task instructions are framed, with more detailed briefing increasing risk of inattentional blindness (39). In this case the attentional set of the participants will have been tuned to clinically relevant details: nodules are typically small, lightly coloured objects on a standard CT scan whereas the inserted shape was dark, and defined only by a faint outline (in a sense similar to lung vasculature). This is not just a theoretical effect, as clinicians are known to search radiographs for diagnostic information differently depending on the information they are given and the predictions they make about likely

diagnoses (40). One might therefore suggest that the expertise of the radiologists in this study contributed little more to this study than had they been shown Simons' basketball-based video.

6.5.1 Inattentional phenomena and other professional groups

While observation is a crucial skill in medicine there are other groups of professionals to whom these skills are equally important, and some lessons can be drawn from studies in these workplaces. Police forces and security personnel also depend on highly accurate observation to prevent and detect crime. These individuals receive overt, explicit training in surveillance, and research has been conducted to assess whether these groups are differentially affected by perceptual error. Evidence for a generalised improvement in perceptual error risk again is lacking, with some studies suggesting no difference between police officers and lay observers witnessing a simulated traffic incident. It is perhaps a matter of concern that, at the point of identifying the suspect, the police officers were highly confident in their identification but this was poorly (or even negatively) correlated with accuracy (41).

In a simulation of a tactical military scenario, Vachon and colleagues also observed a high degree of mission-critical perceptual failures (42). This experiment demonstrated that both central and peripheral inattentional blindness can occur during simulated real-world tasks. Detection was particularly poor if items were not fixated before or after the change (i.e. failure of attention), but interestingly nearly one in five missed events were fixated both before and after the change (demonstrating central failures). The authors concluded that success was therefore dependant on the distribution of attention across the display, as wider focus maximised the likelihood that targets would be fixated around the time of change. These findings then extended to dynamic stimuli previous data that had shown in static tasks fixation-location at the moment of change had a significant effect on the likelihood of successful detection (43). One conclusion to be drawn from this study is that systematically searching for data is a critical part of data acquisition, and this is certainly a point relevant to clinical environments.

The effect of security training was further investigated by Nasholm in a 2014 paper testing for inattentional blindness, again in a military setting (44). Trained observers were compared to novices in a closed-circuit TV monitoring task searching for security threats. The experimental video contained either a suspect package being left on the ground or a person crossing the scene wearing a pirate costume. A high proportion of respondents in both groups failed to notice the indisputably relevant suspect package, and experience did not appear a significant factor in noticing the event⁴. The authors do caution that the training of their 'expert' group may not have been sufficiently specific to confer benefit, as the monitored scene was not specifically military in nature, and this difficulty in matching expertise to situation is one that is also faced by medical investigators.

6.5.2 The inverse-U relationship between load and performance

Nasholm's study had an additional objective in that observers were also randomly allocated to watch their assigned video for different durations; one group observed for only 60 seconds before the critical moment, while the other monitored 42 minutes prior to the significant events. The authors observed that duration of observation was not a significant predictor of noticing.

This is somewhat surprising, because other studies using simulated CCTV-style vigilance activities do imply that 'time on task' is a relevant factor in noticing events, although potentially only where the task is perceptually demanding. In the worst conditions performance can suffer after only eight minutes of activity (45), and this result therefore seems to conflict with that of the suspect-package task described above. Intuitively one would expect noticing rates to fall as observation duration of activity increases due to boredom, fatigue, or deteriorating motivation. Investigations into industrial accidents occurring during sustained or monotonous tasks cite wavering attention as a common

⁴ Observers assigned to the 'pirate' group were significantly less likely to report the unexpected event, and the authors attributed this to perceptual error concluding that task-relevance is an important factor. It is not entirely clear that this is a valid conclusion, however, because it is not certain that failing to report the 'pirate' was caused by inattentional blindness. Having been briefed to evaluate security threats there is nothing to suggest that the costumed person should be any more sinister than other bystanders in the scene. It is possible that they were seen, evaluated, and dismissed.

contributing factor, and attentional focus can degrade over even relatively short periods of duty (46).

This effect of workload is often described as the 'Yerkes-Dodson Law', a relationship often taught on medical courses (47). In 1908 psychologists Robert Yerkes and John Dodson explored the effect of electrical stimulation on behavioural learning in mice (48). From this experiment an inverse-U relationship was derived, and later modified by Hebb (49). The relationship implies that peak performance occurs with an optimal level of stimulation. Performance degrades both with under-stimulation, due to boredom or disengagement, and also with overstimulation, as the individual becomes overwhelmed (Figure 6-2).

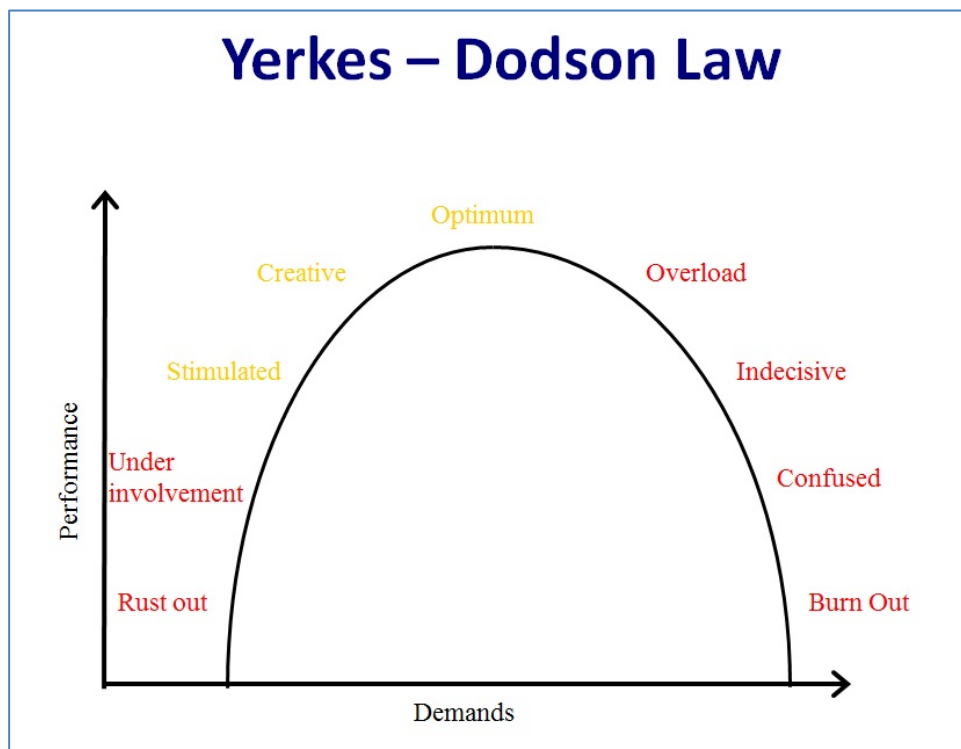


Figure 6-2: The Hebbian modification of the Yerkes-Dodson law; reproduced from the Royal College of Anaesthetists AaE: Anaesthetists' Non-Technical Skills course slide set (5th edition), by kind permission of Dr P Milligan

This association between workload and performance sits easily with perceptual load theories, which predict poor noticing rates in perceptually complex situations. Indeed, in most perceptual error literature, distractors are considered detrimental to task performance; some cognitive processing occurs on distracting information, even if it does not reach conscious awareness (50-53), and this has

been observed to impact on primary task performance (54, 55). Elsewhere it has been noted that subconscious processing of distractors can impair response times when they are incongruent to target items (56), thus semantic categorisation of unattended stimuli also consumes cognitive resources.

At the other end of the continuum there are occasional articles that demonstrate benefits associated with a degree of perceptual load. Studies have suggested that increasing perceptual load actually limits the impact of distractors, at least in certain groups of people classified as 'highly distractible' (57). Similarly, an optimal level of perceptual load can actually improve primary task performance, and this has been demonstrated using both auditory (58) and visual 'distractors' (59). The presentation of additional stimuli may improve performance by amending the attentional set, or alternatively may suppress internally-generated distractions and improve focus (60).

6.5.3 Expectation, memory, and change detection

Expectation has an important role in change-detection (61), and also impacts on what observers recall of a scene. Work by Mack demonstrated that observers can generate false memories of event in perceptual-error tasks, reporting that during critical trials in addition to a high degree of inattentional blindness to critical stimuli, participants would also claim to have seen items not actually displayed simply because they had been present in practice trials (62). A similar study used contextual cueing to manipulate the detection of an unusual object in a still photograph (63). This study used an image, taken in the 1930s, that coincidentally captured a woman who had jumped from a high window in mid-fall. They report that fewer than 2.2% of their 1500 participants noticed the woman⁵. The study manipulated priming information given to the participants, and the group who had read a vignette describing a depressed individual performed much better than others. They also describe a high false-positive rate when individuals were asked to describe elements in the

⁵ This may be a significant underestimate of noticing rate since a much higher proportion detected and reported a central object but were not able to correctly identify it; these observers were still counted as being inattentionally blind.

scene; many reported cars or figures that would be stereotypical in a street-scene, but that were not in fact present. While it is difficult to comment on detection of the incongruent object, it is clear that contextual information does influence recall of the scene and affects identification of objects within it (63).

Expectation can be exploited by stage magicians to make objects apparently vanish. After establishing a pattern by repeatedly manipulating an object such as a ball (64) or a coin (65), if the object is then hidden and the performer's hands simply mime actions one final time, up to two-thirds of the audience might be expected to claim they 'saw' the non-existent object in the final trick. These studies demonstrate that subjective experience is modifiable by top-down factors, and is not solely generated by stimulus factors. In essence, expectation can create false memories of an event.

These findings are a concern to medical practitioners who must carefully and accurately document their activities in medical records. In many critical situations notes are written retrospectively, thus are likely to be affected by this false-memory effect. In many situations the practical effects of this will be somewhat minimal, but at worst an inaccurate record could lead to inappropriate treatment decisions. Should an investigation of an event occur conflicting records might be seen as duplicitous, even though they were sincerely written accounts of a situation.

6.5.4 Non-visual inattentional phenomena

Whereas all previous perceptual error experiments in medicine have used visual stimuli, this study uniquely included an auditory probe. Around 77% of participants in this experiment failed to report the oxygen disconnection, so it is clear that inattentional deafness was relevant here.

The analyses of non-visual inattentional phenomena have received somewhat less attention than visual perceptual error, but these effects do occur in other sensory domains. Auditory perceptual errors are probably the most-studied alternative, and a suggestion of auditory inattention can be detected in some industrial accidents. Although not specifically referred to as such, the

investigations into the crashes of Eastern Airlines flight EAL401 into the Florida Everglades (66), Gulf Air flight GF072 (67), and Air France AF447 (68) all describe circumstances in which auditory alarms were apparently ignored by crew with catastrophic results. These are all situations in which inattentional deafness is likely to have been a major factor, linking the laboratory findings that cognitive load impairs auditory processing (69) with real-life workplace consequences, albeit in a non-medical context.

Such effects have been formally investigated using simulation (70). In an experiment recreating a routine landing in differing weather conditions, the operation was complicated by an indicated landing gear failure. Those flying in calm weather all noted and acted upon the fault, whereas nearly 40% missed the alert when flying in poor weather. Although total flying hours (a marker of experience) was not a predictor of noticing the alert, the group did observe a significant effect for scenario order; pilots who had experienced the calm weather flight first were much more likely to notice the alarm in difficult weather. The authors concluded that the priming effect of prior exposure to the alert sound was meaningful in tuning the participants' attention set. It is not clear from this study if simply encountering the alarm alone would produce the priming effect, or if the alarm would need to be heard in the context of an emergency to sufficiently prime the crew, but this may be an argument in favour of standardising equipment and alerts across workplaces.

Laboratory studies of auditory perception have also demonstrated inattentional effects (71, 72), one experiment even recreating Simons' 'invisible gorilla' in auditory form (73). Due to the complexity of auditory signals it has been difficult to systematically explore the mechanisms of auditory perceptual error, and the understanding of the underlying mechanisms is less advanced than for visual inattention (74), but inattentional deafness does appear functionally similar to inattentional blindness. Set size is a critical factor in determining probability of perceptual error (75), changes in sounds that are more distinctive (76) or that are the focus of attention (77) are more easily detected, and cueing effects are comparable to that seen with visual stimuli (78). Both workload and expertise

effects have been demonstrated in auditory change-detection (79), and spatial separation of sounds affects detection performance (80). As with visual stimuli, distractor effects have been demonstrated for auditory stimuli (71), and processing of sensory information from non-auditory sources affects sound-based event-detection (81), demonstrating limitation in auditory perception exists much as it does for visual information.

An auditory equivalent of change-blindness also exists, with substitutions of objects within the same semantic category less likely to be noticed than those which are incongruent (82). Semantic categorisation and verbal encoding can reinforce memory traces for an auditory stimulus and therefore modulate the risk of change-deafness. Changes in realistic sounds are more recognisable than alterations made in scrambled, unrecognisable noises, and where participants can name the sounds heard (demonstrating successful encoding of the stimulus) additional performance improvements are seen (83). This has led to the conclusion that listeners code a sound's semantic category rather than the specific physical detail of the noise. There is evidence that at least some change-detection tasks are easier if the object can be verbally encoded (84), although this is not an entirely consistent finding (85).

Finally, stimulus saliency is an important factor in both auditory and visual event-detection. It has been recognised for some time that attentional capture is contingent on a combination of object features and top-down factors related to task requirements (86-88). Participants are, for example, much more likely to notice their own name compared to other everyday words. Anthropomorphised shapes are also noticed more frequently than feature-matched objects with no human-like traits (89). Certain other templates might receive attentional prioritisation beyond their task-relevance, such as faces or animals (90, 91). These features appear more resistant to inattentional blindness than other objects, and these may be inherent attentional priorities (92).

Other priority targets might be learned. Pre-attentive selection has been observed in driving contexts, with hazards such as children detected more readily than adults, or similarly sized street

furniture (93). Frequency of exposure seems to be a relevant factor in developing these priority templates, as seldom encountered threats do not seem to generate the same attentional advantage as more commonplace items (90). This again then raises the tantalising possibility that perceptual-error risk might be amenable to training interventions if only suitable materials could be developed.

6.5.5 Summary

This experiment establishes a plausible hypothesis that perceptual failures are a cause of loss of situational awareness in practice. The participants in this study were conventionally trained to a standard appropriate to their clinical role. There is nothing here or in other literature to suggest that similar results would not be found were the study to be replicated elsewhere. Very few people are aware they may be vulnerable to these effects until they are exposed to materials such as this. Although these videos do not replicate the experience of clinical practice, they do depict a realistic scenario with which many healthcare personnel are faced. From these data, it appears likely that perceptual errors are a widespread and unrecognised source of error in healthcare, but new experiments using a better, more realistic activity is required to quantify this. Medical simulation offers an excellent opportunity to explore this.

6.6 References

1. van Beuzekom M, Boer F, Akerboom S, Hudson P. Patient safety: latent risk factors. *Br J Anaesth*. 2010;105(1):52-9.
2. Hoyert D, Xu J. Deaths: Preliminary Data for 2011. *National Vital Statistics Reports*; vol 61 no 6. Hyattsville, MD: National Center for Health Statistics; 2012.
3. Kohn L, Corrigan J, Donaldson M. *To err is human: building a safer health system*. Washington, DC: National Academy Press; 2000.
4. Vincent C. *Patient Safety*. 2nd ed. Oxford: Blackwell; 2010.
5. Singh H, Petersen LA, Thomas EJ. Understanding diagnostic errors in medicine: a lesson from aviation. *Qual Saf Health Care*. 2006;15(3):159-64.
6. O'Brien KS, O'Hare D. Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*. 2007;50(7):1064-91.
7. Walker GH, Stanton NA, Kazi TA, Salmon PM, Jenkins DP. Does advanced driver training improve situational awareness? *Appl Ergon*. 2009;40(4):678-87.
8. Singh H, Thomas EJ, Petersen E, Studdert D. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Archives of Internal Medicine*. 2007;167(19):2030-6.
9. Endsley M, Garland D. *Theoretical underpinnings of situation awareness: a critical review. Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
10. Koustanai A, Van Elslande P, Bastien C. Use of change blindness to measure different abilities to detect relevant changes in natural driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2012;15(3):233-42.
11. Pammer K, Bink C. Attentional differences in driving judgments for country and city scenes: Semantic congruency in inattention blindness. *Accident; analysis and prevention*. 2013;50:955-63.
12. General Medical Council. *Good Medical Practice*. London: General Medical Council; 2013.
13. Resuscitation Council (UK). *Quality standards for cardiopulmonary resuscitation practice and training (acute care)*. London: Resuscitation Council (UK); 2013.
14. Resuscitation Council (UK). *Quality standards for cardiopulmonary resuscitation practice and training (primary care)*. London: Resuscitation Council (UK); 2013.
15. General Medical Council. *Tomorrow's Doctors*. London: General Medical Council; 2009.
16. Greig PR, Higham H, Nobre AC. Perceptual Error in clinical practice: when seeing isn't believing. *Proceedings of the Anaesthetic Research Society Meeting: Merton College, Oxford, UK, April 11–12, 2013. Br J Anaesth*. 2013;111(2):318P-9P.
17. Memmert D. The effects of eye movements, age, and expertise on inattention blindness. *Consciousness and Cognition*. 2006;15(3):620-7.
18. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception*. 1999;28(9):1059-74.
19. Welsby PD. The neurophysiology of failed visual perceptions: some implications for medical teaching. *Postgrad Med J*. 2016;92(1084):74-7.
20. Nolan J, Soar J, Lockey AS, Perkins G, Pitcher D, Gwinnutt C, et al., editors. *Advanced Life Support*. 6th ed. London: Resuscitation Council (UK); 2011.
21. Porubanova-Norquist M, Siki R. Various categories of changes in the inducement of change blindness. *Studia Psychologica*. 2013;55(2):83-94.
22. Beck MR, Angelone BL, Levin DT. Knowledge About the Probability of Change Affects Change Detection Performance. *Journal of Experimental Psychology: Human Perception & Performance*. 2004;30(4):778-91.
23. Koivisto M, Revonsuo A. How meaning shapes seeing. *Psychological Science*. 2007;18(10):845-9.
24. Koivisto M, Revonsuo A. The effects of perceptual load on semantic processing under inattention. *Psychonomic Bulletin and Review*. 2009;16(5):864-8.

25. Most SB. Setting sights higher: category-level attentional set modulates sustained inattentional blindness. *Psychological Research*. 2013;77(2):139-46.
26. Hollingworth A, Henderson JM. Semantic Informativeness Mediates the Detection of Changes in Natural Scenes. *Visual Cognition*. 2000;7(1-3):213-35.
27. Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998;5(4):644-9.
28. Rensink R, O'Regan J, Clark J. To see or not to see: the need for attention to perceive changes in scenes. *Psychological Science*. 1997;8(5):368-73.
29. Conci M, Muller HJ. Global scene layout modulates contextual learning in change detection. *Frontiers in Psychology*. 2014;5:89.
30. Werner S, Thies B. Is 'change blindness' attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*. 2000;7(1-3):163-73.
31. Levin DT, Simons DJ, Angelone BL, Chabris CF. Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology*. 2002;93(Pt 3):289-302.
32. Soar J, Nolan JP, Böttiger BW, Perkins GD, Lott C, Carli P, et al. European Resuscitation Council Guidelines for Resuscitation 2015, section 3: Adult advanced life support. *Resuscitation*. 95:100-47.
33. Kreitz C, Furley P, Memmert D. Inattentional blindness is influenced by exposure time not motion speed. *Quarterly Journal of Experimental Psychology*. 2016;69(3):495-505.
34. Stiegler MP, Neelankavil JP, Canales C, Dhillon A. Cognitive errors detected in anaesthesiology: a literature review and pilot study. *Br J Anaesth*. 2012;108(2):229-35.
35. Dixon JB, Dixon BJ, Cobourn CS. Insights into proximal pouch dilatation following adjustable gastric banding--a form of inattentional blindness? *Obesity Surgery*. 2012;22(7):999-1000.
36. Zamir E. The bias of the question posed: A diagnostic "invisible gorilla". *Diagnosis*. 2015;1(3):245-8.
37. Drew T, Evans K, Vo ML, Jacobson FL, Wolfe JM. What can you see in a single glance and how might this guide visual search in medical images? *Radiographics*. 2013;33:263-74.
38. Drew T, Vö ML-H, Wolfe JM. The Invisible Gorilla Strikes Again: Sustained Inattentional Blindness in Expert Observers. *Psychological Science*. 2013;24(9):1848-53.
39. Memmert D, Furley P. "I spy with my little eye!": Breadth of attention, inattentional blindness, and tactical decision making in team sports. *Journal of Sport and Exercise Psychology*. 2007;29(3):365-81.
40. Suzuki S, Ikusaka M, Ohira Y, Miyahara M, Noda K, Kajiwara H, et al. Effect of diagnostic predictions combined with clinical information on avoiding perceptual errors of computed tomography. *Japanese Journal of Radiology*. 2013;31(11):731-6.
41. Smart SM, Berry MA, Rodriguez DN. Skilled observation and change blindness: A comparison of law enforcement and student samples. *Applied Cognitive Psychology*. 2014;28(4):590-6.
42. Vachon F, Vallieres BR, Jones DM, Tremblay S. Nonexplicit change detection in complex dynamic settings: what eye movements reveal. *Hum Factors*. 2012;54(6):996-1007.
43. Hollingworth A, Schrock G, Henderson JM. Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition*. 2001;29(2):296-304.
44. Nasholm E, Rohlfing S, Sauer JD. Pirate stealth or inattentional blindness? The effects of target relevance and sustained attention on security monitoring for experienced and naive operators. *PLoS ONE*. 2014;9(1).
45. Parasuraman R, de Visser E, Clarke E, McGarry WR, Hussey E, Shaw T, et al. Detecting Threat-Related Intentional Actions of Others: Effects of Image Quality, Response Mode, and Target Cuing on Vigilance. *Journal of Experimental Psychology: Applied*. 2009;15(4):275-90.
46. Edkins GD, Pollock CM. The influence of sustained attention on Railway accidents. *Accident Analysis & Prevention*. 1997;29(4):533-9.

47. Glavin R, Milligan P. *AaE: Anaesthetists Non-technical skills course Facilitators' Manual*. 5th ed. Larbert, UK: Scottish Centre for Simulation and Clinical Human Factors; 2016.
48. Yerkes RM, Dodson JD. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*. 1908;18(5):459-82.
49. Hebb DO. Drives and the C.N.S. (conceptual nervous system). *Psychol Rev*. 1955;62(4):243-54.
50. Ball F, Bernasconi F, Busch NA. Semantic Relations between Visual Objects Can Be Unconsciously Processed but Not Reported under Change Blindness. *Journal of Cognitive Neuroscience*. 2015;27(11):2253-68.
51. Ball F, Busch NA. Change detection on a hunch: pre-attentive vision allows "sensing" of unique feature changes. *Attention Perception & Psychophysics*. 2015;77(8):2570-88.
52. Busch NA. The fate of object memory traces under change detection and change blindness. *Brain Research*. 2013;1520:107-15.
53. Gabay S, Avni D, Henik A. Reflexive orienting by central arrows: Evidence from the inattentional blindness task. *Psychonomic Bulletin and Review*. 2012;19(4):625-30.
54. Mack A, Clarke J. Gist perception requires attention. *Visual Cognition*. 2012;20(3):300-27.
55. Bressan P, Pizzighello S. The attentional cost of inattentional blindness. *Cognition*. 2008;106(1):370-83.
56. Schnuerch R, Kreitz C, Gibbons H, Memmert D. Not quite so blind: Semantic processing despite inattentional blindness. *Journal of Experimental Psychology: Human Perception & Performance*. 2016;42(4):459-63.
57. Forster S, Lavie N. High Perceptual Load Makes Everybody Equal: Eliminating Individual Differences in Distractibility With Load. *Psychological Science*. 2007;18(5):377-81.
58. Beanland V, Allen RA, Pammer K. Attending to music decreases inattentional blindness. *Consciousness and Cognition*. 2011;20(4):1282-92.
59. Pammer K, Korrel H, Bell J. Visual distraction increases the detection of an unexpected object in inattentional blindness. *Visual Cognition*. 2015;22(9-10):1173-83.
60. Forster S, Lavie N. Harnessing the wandering mind: the role of perceptual load. *Cognition*. 2009;111(3):345-55.
61. Ward EJ, Scholl BJ. Inattentional blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*. 2015;22(3):722-7.
62. Mack A, Erol M, Clarke J, Bert J. No iconic memory without attention. *Consciousness & Cognition*. 2016;40:1-8.
63. Slavich GM, Zimbardo PG. Out of mind, out of sight: Unexpected scene elements frequently go unnoticed until primed. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*. 2013;32(4):301-17.
64. Kuhn G, Rensink RA. The Vanishing Ball Illusion: A new perspective on the perception of dynamic events. *Cognition*. 2016;148:64-70.
65. Tompkins ML, Woods AT, Aimola Davies AM. The Phantom Vanish Magic Trick: Investigating the Disappearance of a Non-existent Object in a Dynamic Scene. *Frontiers in Psychology*. 2016;7(950).
66. National Transportation Safety Board. *Aircraft Accident Report (File 1-0016)*. Washington D.C.: National Transportation Safety Board; 1973.
67. Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile. *Accident Investigation Report GF072*. Paris: Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile,; 1993.
68. Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile. *On the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF447 Rio de Janeiro - Paris*. Paris: Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile,; 2012.
69. Macdonald JSP, Lavie N. Visual perceptual load induces inattentional deafness. *Attention, Perception, and Psychophysics*. 2011;73(6):1780-9.

70. Dehais F, Causse M, Vachon F, Régis N, Menant E, Tremblay S. Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentional Deafness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2014;56(4):631-44.
71. Vitevitch MS. Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*. 2003;29(2):333-42.
72. Wayand JF, Levin DT, Varakin DA. Inattentional blindness for a noxious multimodal stimulus. *American Journal of Psychology*. 2005;118(3):339-52.
73. Dalton P, Fraenkel N. Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition*. 2012;124(3):367-72.
74. Dickerson K, Gaston JR. Did you hear that? The role of stimulus similarity and uncertainty in auditory change deafness. *Frontiers in Psychology*. 2014;5:1125.
75. Eramudugolla R, Irvine DRF, McAnally KI, Martin RL, Mattingley JB. Directed Attention Eliminates Change Deafness in Complex Auditory Scenes. *Current Biology*. 2005;15(12):1108-13.
76. Gregg MK, Snyder JS. Enhanced sensory processing accompanies successful detection of change for real-world sounds. *NeuroImage*. 2012;62(1):113-9.
77. Backer KC, Alain C. Orienting Attention to Sound Object Representations Attenuates Change Deafness. *Journal of Experimental Psychology: Human Perception & Performance*. 2012;38(6):1554-66.
78. Griffin IC, Nobre AC. Orienting Attention to Locations in Internal Representations. *Journal of Cognitive Neuroscience*. 2003;15(8):1176-94.
79. Koreimann S, Gula B, Vitouch O. Inattentional deafness in music. *Psychological Research*. 2014;78(3):304-12.
80. Gregg MK, Samuel AG. Change Deafness and the Organizational Properties of Sounds. *Journal of Experimental Psychology: Human Perception & Performance*. 2008;34(4):974-91.
81. Molloy K, Griffiths TD, Chait M, Lavie N. Inattentional Deafness: Visual Load Leads to Time-Specific Suppression of Auditory Evoked Responses. *The Journal of Neuroscience*. 2015;35(49):16046-54.
82. Gregg MK, Samuel AG. The importance of semantics in auditory representations. *Attention, Perception, & Psychophysics*. 2009;71(3):607-19.
83. Gregg MK, Irsik VC, Snyder JS. Change deafness and object encoding with recognizable and unrecognizable sounds. *Neuropsychologia*. 2014;61:19-30.
84. Simons DJ. In sight, out of mind: when object representations fail. *Psychological Science*. 1996;7(5):301-5.
85. Hollingworth A. Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*. 2003;29(2):388-403.
86. Folk C, Remington R, Johnston J. Involuntary Covert Orienting Is Contingent on Attentional Control Settings. *J Exp Psychol Hum Percept Perform*. 1992;18(4):1030-44.
87. Folk C, Remington R, Wright J. The Structure of Attentional Control: Contingent Attentional Capture by Apparent Motion, Abrupt Onset, and Color. *J Exp Psychol Hum Percept Perform*. 1994;20(2):317-29.
88. Braun J, Julesz B. Withdrawing attention at little or no cost: Detection and discrimination tasks. *Perception & Psychophysics*. 1998;60(1):1-23.
89. Mack A, Pappas Z, Silverman M, Gay R. What we see: Inattention and the capture of attention by meaning. *Consciousness and Cognition*. 2002;11(4):488-506.
90. New JJ, German TC. Spiders at the cocktail party: An ancestral threat that surmounts inattentional blindness. *Evolution and Human Behavior*. 2015;36(3):165-73.
91. Devue C, Laloyaux C, Feyers D, Theeuwes J, Bredart S. Do pictures of faces, and which ones, capture attention in the inattentional-blindness paradigm? *Perception*. 2009;38(4):552-68.
92. New J, Cosmides L, Tooby J. Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences*. 2007;104(42):16598-603.

93. Pammer K, Bairnsfather J, Burns J, Hellsing A. Not All Hazards are Created Equal: The Significance of Hazards in Inattentional Blindness for Static Driving Scenes. *Applied Cognitive Psychology*. 2015;29(5):782-8.

7 Insight into Perceptual Errors amongst Healthcare Personnel

'...we falsely believe that we have direct access to our own cognitive processes. Instead, we are only aware of these processes via inference...'

Levin (1)

7.1 Chapter overview

The previous chapter's experiment demonstrated that healthcare personnel are at risk of experiencing perceptual error when exposed to clinical stimuli. These effects may therefore pose potentially serious hazards to staff working in high-workload situations, and materials will need to be developed that educate them about these risks.

An important motivating factor for adult learners is that they see personal relevance in the material presented to them (2). If perceptual errors are a potential source of error in the clinical workplace, then it seems vital to establish to what degree clinicians recognise their personal vulnerability. If clinicians do not feel themselves vulnerable to these effects, which are after all not prominent features of normal experience, then they may dismiss these out of hand.

Experimental psychology is not widely taught in medical curricula. It does not feature at all in the undergraduate curriculum of the author's institution (3), and did not feature in the postgraduate curricula analysed for the experiment described in Chapter 3. Previous experiments on psychology undergraduates suggest that insight into these phenomena is poor, even amongst those who have been exposed to relevant educational materials.

It is hypothesised, therefore, that clinicians will significantly overestimate their perceptual reliability. Quantifying and understanding this will therefore play an important part in how these risks are communicated to healthcare staff in future.

7.2 Introduction

The day-to-day experience of perception is that it is rich, detailed, and automatic, but perceptual errors demonstrate that this phenomenological experience may be misleading. Moreover, a surprisingly large percentage of the general public fundamentally misunderstand how the visual system works. It has been demonstrated that more than 50% of respondents to questionnaires report that vision requires the emission of some form of wave from the eye (4), and that such beliefs can persist in spite of specific education (5).

Medical professionals obviously possess a much deeper understanding of the anatomy and physiology of the visual system than most lay people. There is, however, evidence that the cognitive aspects of visual perception are still incompletely understood, by even these trained personnel. The postgraduate curriculum for one medical specialty searched in Chapter 3, for example, repeatedly states that it essential that their trainees can *always* recognise emergency situations (6). A body of senior medical opinion clearly expects staff to be infallible observers. Given what is known about perceptual errors from existing psychological literature this learning objective is unachievable. This also raises questions about the degree to which front-line staff might individually overestimate their own perceptual reliability.

It has already been demonstrated that both scientists and the public lack insight into change blindness. Large overestimates of perceptual reliability appear commonplace, and are consistently held (7). For example, after presenting the famous 'Door Experiment' (8) to a cohort of psychology undergraduates Levin asked volunteers to predict their likelihood of noticing the switch in conversation partner (1). The participants were provided with before-and-after pictures of the

scene, and 98% of these individuals expected to detect the change even though fewer than half of the participants in the original study reported the switch.

Most people expect change-detection to be an automatic process, and do not understand that effort and attention are required (9). Subconscious attentional prioritisation also affects change-detection performance, by biasing the distribution of attentional resources towards more likely events (10). Although laboratory experiments demonstrate that performance varies with set-size (11), individuals' predictions are fairly constant when asked to consider their likelihood of detection change in scenes of varying complexity (12) or with differing intervals between exposure to pre- and post-change scenes. Subjects even expect successful change detection can occur up to an hour after exposure to the pre-change object (13). Levin terms the cumulative effect of these overestimates a 'metacognitive error' – an error arising from how people think about cognitive processes – and names this misunderstanding 'change-blindness blindness'.

Why such errors occur is less well understood. Levin notes that one of the difficulties people face when reflecting on perception is that successes are generally more memorable than failures, and this introduces recall-bias into their predictions (14). This may be particularly true in the field of failed perception. Events which are noticed may trigger a surprise response, aiding recall. Events that are missed, however could be termed 'unknown unknowns'. This term, derived from Luft and Ingham's Johari Square (15), describes items that an individual is neither aware of, nor has any appreciation that the subject in question exists at all. It would be impossible for an individual to keep an accurate tally of noticed versus unnoticed items, because there are no denominator data against which to compare.

Such false intuitions about perception may then lead individuals to make inappropriate predictions about change-detection performance. The danger is then that overconfidence leads to inadequate searches, and there are suggestions that this can have real-world implications. There is good

evidence, for example, that police 'line ups' (16) and eye-witness testimonies (13) are afforded a large amount of weight by juries, but are in fact more prone to error than commonly understood.

From these findings one can reasonably predict that medical personnel will be at risk of change-blindness blindness, and that it is plausible to expect this could impact on their clinical performance. Someone who expects to detect important information automatically might not make intentional searches, or lend only cursory glances to information displayed on clinical monitors. If medical personnel were to display a high degree of change-blindness blindness this would demonstrate an important opportunity to deliver training to improve insight, particularly of staff working in acute care environments.

This study was therefore designed to assess the degree to which clinicians with varying degrees of experience could predict their change-detection reliability in a series of commonplace resuscitation-based scenarios.

7.3 Methods

7.3.1 Design

This is a prospective cohort study.

7.3.2 Subjects

Volunteers were largely drawn from the University of Oxford and staff of the NHS South Central region. Participants were approached by email invitation, and were also free to pass the invitation on to colleagues at their own discretion.

A convenience sample of 142 was intended, paralleling the group sizes recruited to the experiment detailed in chapter 6.

7.3.3 Outcome Measures

The primary outcome for this study was the proportion of respondents predicting they would notice each event, stratified by their level of training and professional background.

7.3.4 Study Procedures

This study was approved by the University of Oxford Central Research Ethics Committee (MSD-IDREC-C1-2014-088). All participants gave informed consent prior to enrolment.

The study was conducted online, and materials were hosted by SurveyMonkey (SurveyMonkey Europe SARL, Luxembourg; www.surveymonkey.com).

The experiment was designed to parallel the methods employed by Levin (1), but using primarily medically-orientated scenarios. Participants were first asked to give basic demographic data, then describe their professional background and level of resuscitation training. They were then presented with three scenarios and invited to predict if they would notice a series of events occurring in each. The scenarios are listed in Table 7-1. Responses were collected using a four-point Likert scale between *'I am certain I would not notice this'* to *'I am certain I would notice this'*.

Scenario and events list

Scenario	Description	Events
Participating in a live medical scenario	You are present at a real-life resuscitation scenario; the patient is in VF cardiac arrest, and the team are providing CPR.	<ul style="list-style-type: none"> One member of the team leaves the area to get results from a blood gas; when they return they are wearing a different coloured hat One member of the team leaves the area to get a piece of equipment; a completely different person returns You look down to consult the patient's notes; when you look up the patient has been intubated You look down to consult the patient's notes; when you look up the rhythm on the patient monitor has reverted to sinus The oxygen supply becomes disconnected The patient's IV line becomes disconnected
Watching a recorded medical scenario	You are watching a video of a resuscitation situation; you have been instructed to monitor the team's performance, so are watching carefully	<ul style="list-style-type: none"> Two members of the team switch hats A stethoscope appears around one team member's neck The person delivering CPR is instantaneously swapped with another person in view The patient is instantaneously intubated The oxygen supply becomes disconnected
Participating in a live non-medical scenario	You are walking in an area you know well when you are approached for directions. Whilst giving the directions you are interrupted by two people carrying a door.	<ul style="list-style-type: none"> During the interruption the person you are talking with is switched for a completely different individual

Table 7-1: Scenarios and events

After considering the three scenarios enquiries were made to determine if participants had encountered perceptual errors, or previously seen either the Simons 'Invisible Gorilla' (17), or the Greig CPR video (18).

On completion of the questionnaire participants were then given additional information and a link to further reading, and were given a final opportunity to withdraw from the study.

7.3.5 Analysis

The only continuous variable (age) was not normally distributed so non-parametric comparisons were made using the Kruskal-Wallis H test, and pairwise comparisons performed with Dunn's procedure and a Bonferroni correction. Chi-squared analysis (or Fisher's exact test where indicated) was employed to compare proportional data. Two-sided *p*-values are quoted for each test. Binary logistic regression was used to explore the relationship between predictors of noticing each event.

7.4 Results

7.4.1 Participant Demographics

A total of 125 participants completed the study (59 males, 66 females). The study slightly under-recruited, despite twice gaining approval to extend the experiment's recruitment period. Most participants were medical students ($n = 74$) or doctors ($n = 45$). These data are summarised in Table 7-2 and Table 7-3. Medical students were significantly younger than doctors ($p < 0.01$) and nurses ($p < 0.01$), but doctors and nurses did not differ significantly in age ($p = 1.00$).

Participant background and resuscitation experience					
		Professional Background (n)			Total
		Doctor	Medical Student	Nurse/Resus Officer	
Level of Resuscitation Training	Basic/None	3	60	2	65
	ILS/ALS Trained	42	14	4	60
	Total	45	74	6	125

Table 7-2: Participants by professional background and resuscitation experience

Participant demographics					
		Males	Females	Median and Interquartile Range	
				Age	
Level of Resuscitation Training	Basic/None	27	38	23.0	22-24
	ILS/ALS Trained	32	28	27.0	24.3-32
	Total	59	66		

Table 7-3: Participant demographics by level of experience

Most participants ($n = 65$) had no resuscitation training, or were trained only to the level of basic life support; the remaining 60 were trained at least to the level of immediate life support, 56 of whom had received training on a Resuscitation Council (UK) accredited Advanced Life Support course. There was no association between participants' level of training and their predictions for noticing any of the study events, either in live or video-based situations (Table 7-4).

Somewhat surprisingly most participants had encountered some information about perceptual errors prior to their participation in this study (Table 7-5). Medical students were significantly more likely to have encountered this than the other professional groups ($p < 0.01$). Most participants who had encountered perceptual errors had done so via the Simons Gorilla video ($n = 92$), and this was

far greater than the number of people who had heard the specific term ‘perceptual error’ (n = 47). Only seven participants had seen the Greig CPR video before, two of whom had taken part in the original experiment. None of the participants had read the 2013 *Resuscitation* paper.

The predictions were next analysed based upon the participants’ foreknowledge of perceptual error. Foreknowledge was demonstrated by having encountered the terms ‘perceptual error’ or ‘invisible gorilla’, or having seen the Greig resuscitation video before. Again, there was no association between foreknowledge of perceptual error and predictions about noticing any of the specified events (Table 7-6).

Predicted noticing rates by participant experience					
	Predict would notice			p value	
	Overall	Level of Resuscitation Training			
		Basic/None (n=65)	ILS/ALS Trained (n=60)		
Events witnessed in a live clinical situation	Hat change	31 (24.8%)	17 (26.2%)	14 (23.3%)	0.84
	Team change	104 (83.2%)	55 (84.6%)	49 (81.7%)	0.81
	Intubation	123 (98.4%)	63 (96.9%)	60 (100.0%)	0.50
	Oxygen disconnection	91 (72.8%)	47 (72.3%)	44 (73.3%)	1.00
	IV disconnection	65 (52.0%)	36 (55.4%)	29 (48.3%)	0.48
	Change in rhythm	113 (90.4%)	58 (89.2%)	55 (91.7%)	0.54
Events depicted in a clinical video	Hat change	29 (23.2%)	16 (24.6%)	13 (21.7%)	0.69
	Team change	95 (76.0%)	50 (76.9%)	45 (75.0%)	0.83
	Intubation	118 (94.4%)	59 (90.8%)	59 (98.3%)	0.21
	Oxygen disconnection	100 (80.0%)	52 (80.0%)	48 (80.0%)	1.00
	Stethoscope drop	19 (15.2%)	10 (15.4%)	9 (15.0%)	1.00
Events witnessed in a live non-clinical situation	Person change	114 (91.2%)	59 (90.8%)	55 (91.7%)	1.00

Table 7-4: predicted noticing rates by participant experience

These data were confirmed using binary logistic regression. Each event was tested for the predictors: professional background, level of training, and awareness of perceptual error. Logistic regression confirmed that none of these factors predicted participant responses.

Participants were consistent in their predictions across the video- and live-scenarios. Chi-squared testing confirmed that there was an association between predictions for the four events that were depicted in both conditions (Table 7-7).

Summary of participant knowledge of perceptual error by professional background

		Professional Background			
		Overall	Medical Student (n=74)	Doctor (n=45)	Nurse/Resus Officer (n=6)
Foreknowledge of Perceptual Error	Unaware	20 (16.0%)	8 (10.8%)	8 (17.8%)	4 (66.7%)
	Aware	105 (84.0%)	66 (89.2%)	37 (82.2%)	2 (33.3%)

Table 7-5: Foreknowledge of perceptual error by participant background

Predicted noticing rates by participant knowledge of perceptual error

		Predict would notice			p value
		Overall	Foreknowledge of perceptual error		
			Unaware of perceptual error (n=20)	Aware of perceptual error (n=105)	
Events witnessed in a live clinical situation	Hat change	31 (24.8%)	7 (35.0%)	24 (22.9%)	0.27
	Team change	104 (83.2%)	18 (90.0%)	86 (81.9%)	0.52
	Intubation	123 (98.4%)	20 (100%)	103 (98.1%)	1.00
	Oxygen disconnection	91 (72.8%)	15 (75.0%)	76 (72.4%)	1.00
	IV disconnection	65 (52.0%)	12 (60.0%)	53 (50.5%)	0.47
	Change in rhythm	113 (90.4%)	18 (90.0%)	95 (90.5%)	1.00
Events depicted in a clinical video	Hat change	29 (23.2%)	6 (30.0%)	23 (21.9%)	0.38
	Team change	95 (76.0%)	14 (70.0%)	81 (77.1%)	0.49
	Intubation	118 (94.4%)	18 (90.0%)	100 (95.2%)	1.00
	Oxygen disconnection	100 (80.0%)	15 (75.0%)	85 (81.0%)	0.76
	Stethoscope drop	19 (15.2%)	4 (20.0%)	15 (14.3%)	0.77
Events witnessed in a live non-clinical situation	Person change	114 (91.2%)	19 (95.0%)	95 (90.5%)	0.36

Table 7-6: Predicted noticing by foreknowledge of perceptual error

Testing for an association between predictions in each scenario condition

		Predict would notice		p value
		In video scenario (n = 125)	In live scenario (n = 125)	
Event	Hat change	100	90	0.05
	Team-member switch	95	103	< 0.01
	Intubation	118	123	< 0.01
	Oxygen disconnection	29	31	< 0.01

Table 7-7: Chi-squared testing predictions of noticing the four matched events in a video-based and live scenario

The logistic regression models were therefore rerun adding response to the ‘video’ situation as a predictor when analysing their ‘live’ response, and vice versa. With the exception of intubation, which did not reach significance, these were the only significant factors in any regression models (Table 7-8).

Logistic regression model factors for 'live' versus 'video' based participant predictions

	Participant response	
	Live event predicting video event (<i>p</i> value)	Video event predicting live event (<i>p</i> value)
Noticing the hat change	0.01	< 0.01
Noticing the team-member switch	0.01	< 0.01
Event Noticing the patient's intubation	0.06	0.15
Noticing the disconnected oxygen supply	0.05	0.05

Table 7-8: Predicting participant responses in 'live' and 'video' conditions

7.4.2 Predictions of noticing versus actual noticing rates

These data offer a useful opportunity to compare predictions with reality, as many of the described events have been tested for in other experiments.

Simons and Levin conducted two iterations of the famous 'door experiment'. In the first, seven of 15 observers (46.7%) noticed the change in conversation partner, and four of 12 observers (33.3%) in the second (8). Levin followed this up by testing for insight into change blindness; 288 of 295 participants (97.6%) predicted that they would notice the switch (1), a highly significant overestimate.

Participants in the current experiment were also asked to predict their likelihood of noticing a switch in conversation partner under the circumstances depicted in the 'door experiment'. In this case 114 of 124 (91.2%) of respondents predicted that they would notice the change, again a significant overestimate compared with the actual findings of the 'Door Experiment' ($p < 0.01$). Note that a lower proportion of medical respondents predicted they would notice the change compared with Levin's psychology undergraduates ($p < 0.01$).

When considering predictions of noticing events in the video-based scenario respondents generally overestimated the likelihood of noticing events. These data are summarised in Table 7-9.

Predicted versus actual noticing rates in the video scenario

	Predict noticing in video scenario (n=125)	Actual noticing in video scenario (n=142)	p value	
	Hat change	29 (23.2%)	11 (7.7%)	< 0.01
	Stethoscope change	19 (15.2%)	24 (16.9%)	0.42
Event	CPR provider switch	95 (76.0%)	36 (25.4%)	< 0.01
	Airway device switch	118 (94.4%)	34 (23.9%)	< 0.01
	Oxygen disconnection	100 (80.0%)	34 (23.9%)	< 0.01

Table 7-9: comparison of predicted frequency of noticing each event portrayed in the video scenario compared with observed noticing frequencies

The pattern was less clear when considering the predictions versus results observed in simulation (Table 7-10). Participants were asked to predict their noticing of events that were to be included in the forthcoming simulation-based experiment, described in the next chapter.

Compared to the noticing rates observed in this experiment, participants here actually underestimated their likelihood of reporting the change in hats, but overestimated noticing the IV line disconnection, and were realistic about their chances of noticing the oxygen disconnection and switch in team member. The proportion of people who detected the person switch in simulation was 15 of 20 (75.0%), although this was highly sensitive to workload, as 10 of these were working in low-workload conditions; only 5 of 9 (55.6%) noticed the switch in high workload conditions. Noticing the person-switch was reported approximately three times more frequently in simulation than a similar event in the CPR video (75.0% vs 25.4%, $p < 0.01$), and slightly more than twice as frequently as the swap in the 'Door Experiment' (75.0% vs 33.3%, $p = 0.03$).

Predicted versus actual noticing rates in the live scenario

	Predict noticing in live scenario (n=125)	Actual noticing in live scenario (n=20)	p value	
	Hat change	31 (24.8%)	10 (50.0%)	0.03
	Team member switch	104 (83.2%)	15 (75.0%)	0.36
Event	IV line disconnection	65 (52.0%)	5 (25.0%)	0.03
	Oxygen disconnection	91 (72.8%)	12 (60.0%)	0.29

Table 7-10: Comparison of predicted frequency of noticing each event encountered in the live scenario compared with observed noticing frequencies

7.5 Discussion

Consistent with expectations, participants in this study did demonstrate inattention- and change-blindness for specific events that had been depicted in both live and video-based

scenarios. With the single exception of the stethoscope change, respondents demonstrated significant overconfidence in their likelihood of noticing the events in the video-based scenario. Approximately three to four times as many people expected to notice each event than did (Table 7-9). Clinicians therefore are overconfident in their ability to detect changes in medically-orientated scenarios.

There evidence is, however, less clear for change-blindness blindness in real-life situations. This was not because respondents were more pessimistic; in fact, participants were highly consistent in their responses and those who predicted they would (or would not) notice an event on video almost always gave the same response when making predictions about a similar event in a live situation. Rather it is because the participants in the simulated live scenario apparently over-performed when compared with those who undertook the video-based task. The reasons for this better-than-predicted event noticing are fully addressed in the discussion associated with that experiment (see Chapter 8).

The phenomenon of inattentional blindness seems reasonably well known amongst this medical cohort. Only 16% of respondents had never encountered perceptual errors before, although it is notable that while most of the participants in the study had previously seen the Simons 'Invisible Gorilla' video, only around half of them had heard any of the specific psychological terminology. It suggests that while the video is being used as a demonstration of how awareness can fail in surprising ways, little teaching is given about the mechanisms by which this can occur.

This is disappointing, as information about cognitive workload and its contribution to risk is highly relevant, particularly to the relatively inexperienced staff that often form the front-line of hospital medicine. Evidence suggests that junior medical staff do not always demonstrate strong non-technical skills. Situational awareness, team-working and communication errors are frequent contributors to critical incidents in practice (19). At an institutional level there is not always a recognition that specific training in such skills is necessary (20), and this can extend to individual

trainees. The author has personal experience of one trainee who dismissed non-technical skills training as *'touchy-feely c**p'*. A more thorough understanding of perceptual errors might therefore serve as a useful encouragement to better develop such skills, and to value the contribution of other team-members.

Why people seem to overestimate their perceptual reliability so consistently can be considered from an educational perspective. Learning can be described using Kolb's learning cycle, whereby an event experienced by the learner triggers reflection on that episode and the development of new understanding. The learner then experiments with new actions on which further reflection can occur (21). Taken in this context, the perceptual failures of daily-life remain unnoticed, and therefore cannot generate any reflection or behavioural modification. It is therefore easy to see why the lack of feedback available about perceptual failure in daily activities can then breed overconfidence, but there is evidence that this is amenable to change. Although daily experience does not give individuals an opportunity to reflect on the reliability of their perceptual abilities, such opportunities do exist in the laboratory. It has been demonstrated that people can modify their predictions on perceptual reliability based on specific feedback after completing change-blindness activities (9).

Evidence from this study supports the notion that education can bring about change in prediction behaviour. Even the low-level training that the current study participants have experienced did show a modest impact on their detection predictions. Although the study cohort was not accurate in their predictions for most of the events, they were at least less inaccurate than populations tested by existing experiments. Making a direct comparison between the population in the current study and those of Levin's original change-blindness blindness study (1), the clinicians sampled here were significantly more sceptical about their likelihood of noticing the change in conversation partner depicted in the 'door experiment', albeit that they still greatly overestimated when compared with the actual study results (8).

The population of this study was more familiar with perceptual errors than those of Levin’s work. This leads one to conclude that clinicians’ overconfidence is amenable to training and that better designed educational materials would be expected to calibrate clinicians’ expectations still more realistically. The non-clinical nature of stimuli such as the ‘Invisible Gorilla’ video is relevant in the design of teaching materials. It is a highly effective, reliable, and successful demonstration of perceptual error, but adult learning is most effective when the material can be made personally relevant (2). Materials specifically designed with clinical practice in mind such as the CPR video used in the previous study (18) (the only specifically clinical demonstration of inattentional- and change-blindness available) will be a more effective means of encouraging healthcare professionals to reflect on their own experience and practice.

Existing data demonstrate that probable events are more readily detectable than improbable events, although participants’ prediction of noticing events does not generally differ based on plausibility (22). The results of the current study, however, show that these clinicians did treat such events differently. In both the video-based and live scenarios participants predicted they would notice the most clinically-relevant events most often, and the more trivial changes least frequently (Table 7-11). One must conclude therefore that participants were considering their ‘spotlight of attention’ because focus was primarily on those details that would affect clinical decision-making.

Order in which events were predicted to be noticed		
	In a live clinical situation	Depicted in a video
Most frequently predicted to notice	Intubation of the patient	Intubation of the patient
	Change in cardiac rhythm	Oxygen disconnection
⇓ ⇓ ⇓	Team change	Team change
	Oxygen Disconnection	Hat change
Least frequently predicted to notice	IV line disconnection	Stethoscope drop
	Hat change	

Table 7-11: order of predictions for noticing events in each clinical scenario

These data differ from the non-clinical cohorts sampled by much of the existing psychology literature. The difference is not necessarily due to the clinical training of the cohort, but rather reflects that the participants were asked to consider noticing within the context of another active

task. The original change-blindness studies (1), and subsequent work by the same group (12), have largely investigated insight into perceptual error using more generic or laboratory-based activities. A criticism of this approach is that it denies people the opportunity to draw upon their own expertise or experience when making judgements, so in this regard the present study is perhaps closer to the findings of Smilek (23, 24).

Smilek's work has focussed on using real-world-relevant scenarios such as driving on a public road or witnessing a theft from an office environment. People questioned about these scenarios better understood the role of intention in change detection, and were more accurate in their predictions. Framing perceptual error predictions within specific, rather than generic, scenarios also makes judgements around scene complexity and intention more readily accessible (25). In the current study participants were given definitive instructions about viewing or participating in a resuscitation situation. This is an activity with which all participants have at least some familiarity from their clinical training, and is a situation in which they would expect to actively and intentionally search for information in the conduct of their duties. This seems to have been reflected in their predictions of noticing the critical events. An important conclusion to consider is that if change-blindness predictions are sensitive to the method by which the study is conducted it may be difficult to draw comparisons between studies of the phenomenon, and that some of the effect may be an artefact. Ecological validity is crucial, and at the very least one must be cautious about applying to the real-world conclusions drawn from purely laboratory-generated stimuli.

While video-based or computer-generated stimuli are highly convenient methods of testing individuals, it does appear from these data that their application to real-world scenarios may be more limited than previously thought. These methods can be strictly standardised, and undoubtedly have great value in the exploration of basic perceptual and cognitive mechanisms. These data do however suggest that to generate conclusions robustly applicable to real-world tasks, the best data

are generated using tasks most closely aligned to such activities. In this regard simulation is ideally placed to explore these issues in a standardised way.

7.5.1 Summary

This experiment was designed to explore clinicians' level of insight into perceptual error, and the accuracy with which they predict their own perceptual reliability. As expected, a high level of change-blindness blindness was observed.

It was surprising that a relatively high number of participants had encountered perceptual phenomena before, but it was interesting that prior exposure to the phenomena was not associated with better predictions. Simply exposing people to perpetual phenomena is therefore clearly insufficient to improve their performance, and specific educational materials are likely to be required.

Knowledge of people's expectations is crucial to developing such resources. As noted in chapter 2 with human factors, adults must understand the personal relevance of materials to motivate their learning, and therefore this must be accounted for when developing such articles.

7.6 References

1. Levin DT, Momen N, Drivdahl S. Change blindness blindness: the metacognitive error of overestimating change-detection ability. *Visual Cognition*. 2000;7(3):397-412.
2. Cantillon P, Hutchinson L, Wood D, editors. *ABC of learning and teaching in medicine*. London: BMJ Books; 2003.
3. Curriculum Development Committee. *Integrated Curriculum in Medicine and Surgery 2014-2015*. Oxford: University of Oxford Medical School; 2015.
4. Winer GA, Cottrell JE, Gregg V, Fournier JS, Bica LA. Fundamentally misunderstanding visual perception: Adults' belief in visual emissions. *American Psychologist*. 2002;57(6-7):417-24.
5. Gregg VR, Winer GA, Cottrell JE, Hedman KE, Fournier JS. The persistence of a misconception about vision after educational interventions. *Psychonomic Bulletin & Review*. 2001;8(3):622-6.
6. Joint Royal Colleges of Physicians Training Board. *Specialty Training Curriculum for Medical Oncology*. London: Royal College of Physicians; 2007.
7. Levin DT, Angelone BL. The visual metacognition questionnaire: a measure of intuitions about vision. *American Journal of Psychology*. 2008;121(3):451-72.
8. Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998;5(4):644-9.
9. Loussouarn A, Gabriel D, Proust J. Exploring the informational sources of metaperception: The case of change blindness blindness. *Consciousness and Cognition: An International Journal*. 2011;20(4):1489-501.
10. Angelone BL. *Incidental change detection and short-term memory load in a dual-task paradigm*. Dissertation Abstracts International: Section B: The Sciences and Engineering. 2004;64(9-B):4643.
11. Nobre AC, Kastner S, editors. *The Oxford Handbook of Attention*. Oxford: Oxford University Press; 2014.
12. Beck MR, Levin DT, Angelone B. Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition*. 2007;16(1):31-51.
13. Levin DT, Drivdahl SB, Momen N, Beck MR. False predictions about the detectability of visual changes: The role of beliefs about attention, memory, and the continuity of attended objects in causing change blindness blindness. *Consciousness and Cognition*. 2002;11(4):507-27.
14. Hearst E. Psychology and Nothing. *American Scientist*. 1991;79(5):432-43.
15. Luft J, Ingham H. *The Johari window, a graphic model of interpersonal awareness*. Los Angeles: University of California, Los Angeles; 1955.
16. Smart SM, Berry MA, Rodriguez DN. Skilled observation and change blindness: A comparison of law enforcement and student samples. *Applied Cognitive Psychology*. 2014;28(4):590-6.
17. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception*. 1999;28(9):1059-74.
18. Greig P, Higham H, Nobre A. Failure to perceive clinical events: an under-recognised source of error. *Resuscitation*. 2014;85:952-6.
19. Singh H, Thomas EJ, Petersen E, Studdert D. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Archives of Internal Medicine*. 2007;167(19):2030-6.
20. Greig PR, Higham H, Vaux E. Lack of standardisation between specialties for human factors content in postgraduate training: an analysis of specialty curricula in the UK. *BMJ quality & safety*. 2015;24:558-60.
21. Kolb DA, Fry R. Toward an applied theory of experiential learning. In: Cooper C, editor. *Theories of Group Process*. London: John Wiley; 1975.
22. Beck MR, Angelone BL, Levin DT. Knowledge About the Probability of Change Affects Change Detection Performance. *Journal of Experimental Psychology: Human Perception & Performance*. 2004;30(4):778-91.

23. Smilek D, Eastwood JD, Reynolds MG, Kingstone A. Metacognitive errors in change detection: missing the gap between lab and life. *Conscious Cogn.* 2007;16(1):52-7; discussion 8-62.
24. Smilek D, Eastwood JD, Reynolds MG, Kingstone A. Metacognition and change detection: Do lab and life really converge? *Consciousness and Cognition.* 2008;17(3):1056-61.
25. Beck MR, Levin DT, Angelone BL. Metacognitive errors in change detection: Lab and life converge. *Consciousness and Cognition.* 2007;16(1):58-62.

8 Workload and Perceptual Error during Simulated Cardiac Arrest

'The NHS lags unacceptably behind other safety-critical industries, such as aviation, in recognising the importance of effective team-working and other non-technical skills.'

Parliamentary Report into Patient Safety (1)

8.1 Chapter overview

Chapter 6 established that clinicians are susceptible to perceptual error, but the stimuli used in this experiment did not fully replicate the richness of a clinical environment. Although the experiment mirrored well-established psychology laboratory studies, questions might still be asked about how these data translate into practice.

Testing for perceptual error in a real clinical environment would be impractical. There would be implications for patient safety, were distractors or cues to be incorporated into a scenario where real tasks were being carried out. Also, it would be difficult, if not impossible, to adequately control events.

Accordingly, simulation was used to create a plausible scenario, but one in which stimuli could be inserted without jeopardising patient safety. Three hypotheses were available for testing. Firstly, it was expected that perceptual errors would occur, and secondly that their frequency would be at least that observed when using video materials. Thirdly, that the risk of perceptual error would be sensitive to both workload and expertise.

These results however are not guaranteed. The perceptual interaction with a three-dimensional environment is necessarily different to a two-dimensional video display. The visual scene is not fixed, as participants are free to move around or turn their heads. Also, there is a greater level of

involvement when physically present in the environment. Even a simulated emergency can be emotionally charged, and this might be expected to degrade performance, at least amongst those least used to coping with such situations.

This experiment was therefore designed to investigate the interaction between workload, experience, and incidence of perceptual error.

8.2 Introduction

Previous investigations have demonstrated that loss of situational awareness is a major contributor to critical incidents in medicine (2) and industry (3, 4). Good quality situational awareness requires the initial perception of information (5), but Simons' 'Invisible Gorilla' experiment demonstrated one way in which surprisingly large unexpected events can pass completely unnoticed (6).

Whilst an interesting demonstration, the Simons experiment tells us little about how these effects impact on clinical practice.

Some of this gap was addressed by developing a clinically-orientated video to explore perceptual error using medically-relevant stimuli (7). This was highly successful in demonstrating that even subject-experts experience high levels of inattention blindness to events within their field of expertise. That these events would meaningfully affect patient outcome should they happen in a real emergency is also relevant. This was the first experiment to demonstrate sustained inattention blindness in a dynamic task in a clinically-trained population. It established the plausibility of perceptual errors as a source of situational-awareness failure, and that the topic should be a subject of interest particularly to acute-care clinicians or those with training responsibilities.

It does not however fully answer the question of how such events occur in real-life. There are some data to suggest a divergence between performance in laboratory-based tasks and real-world activities. The complexity of performing in real-world tasks may mean that laboratory studies of perceptual function are not fully generalisable (8). Smilek notes that there is '*a critical theoretical*

issue in psychological research: Namely, it brings to the forefront the concern that the measuring tools and/or the experimental situations used to investigate cognitive phenomena can lead to conclusions that may misrepresent what is going on in the real world' (9).

If it is the case that laboratory studies, such as the previous experiments using video-based stimuli, do not accurately represent the occurrence of these phenomena in real-life then alternative lines of investigation must be considered. The 'gold-standard' method would be to observe and record events in a real clinical environment. Although possible, this would be impractical in most situations; issues of confidentiality and data-protection would make recordings challenging in busy ward environments where images may be captured of several patients or staff members. Also, although at the hospital-system level medical errors are commonplace, the individual practitioner will not experience errors on an everyday basis. This would mean many hours of recording would be required to capture a small number of errors, making such an experimental design impractical without large-scale resources and support.

What is therefore needed is a controlled environment that faithfully replicates a clinical area, and can be staffed by a group of clinicians who are happy for their actions to be recorded. Medical simulation is just such an environment, and offers the advantage that clinical emergencies can be recreated in a repeatable and predictable manner. This parallels recent work that has investigated perceptual error using high-fidelity simulators in non-medical scenarios, which successfully demonstrated high levels of inattention blindness and deafness during a driving task (10).

This experiment was therefore designed to test for perceptual errors using the most ecologically valid methods possible.

8.3 Methods

8.3.1 Study Design

A 2x2 repeated measures observational study with both between- and within-subjects factors.

8.3.2 Subjects

Volunteers for this study were recruited from clinical medical students⁶ attending the University of Oxford. Participants were approached via email.

The sample size was calculated to detect an effect size of 0.5, a level comparable to the expertise-effect demonstrated by Simons (6), and based on $\alpha = 0.05$ and $\beta = 0.8$. This predicted that at least 18 participants would be required.

8.3.3 Outcome Measures

The primary outcome was intended to be the noticing rate of the critical stimuli by level of participant experience. The secondary outcome was to determine variation in visual search strategy between experience- and workload-conditions.

This experiment also yielded data that could be used to test some of the predictions made by respondents in the experiment described in the preceding chapter (see Chapter 7). Also, these data allow comparison with the noticing rate for similar video-based events reported by the experiment described in Chapter 6. These comparisons were therefore included as tertiary objectives.

8.3.4 Study Procedures

This study was approved by the University of Oxford Central Research Ethics Committee (MS-IDREC-C1-2015-056). All participants gave informed consent prior to enrolment.

Participants were instructed to lead two cardiac arrest scenarios (Table 8-1), designed to vary in the effort demanded of the participant. These simulations were conducted in the University of Oxford Centre for Simulation, Teaching, and Research (OxSTaR) Centre.

⁶ The experiment was originally designed to also investigate the interaction between experience and workload, and the intention was to recruit from both medical students and qualified staff trained in adult life support. Unfortunately, the eye-tracking system developed a hardware fault that prevented completion of the study as planned. At time of writing (approximately 12 months after initial recruitment) this fault has not yet been rectified.

The simulation room was configured to represent a generic ward bed-space, and participants were informed in each case that the resuscitation team had been called, but had not yet arrived.

Summary of resuscitation scenarios		
	High Workload Scenario	Low Workload Scenario
Briefing	‘You are passing through the Emergency Department when you hear an alarm. When you arrive at the bedside you find a 69-year-old man collapsed in bed. CPR has been started by a ward nurse and an FY1. We would like you to lead this team, but your team is short-handed and you will also have to assist them in delivery of care to the patient.’	‘You are passing through the Emergency Department when you hear an alarm. When you arrive at the bedside you find a 69-year-old man collapsed in bed. CPR has been started by a ward nurse and two FY1s. We would like you to lead this team. There are enough staff to enable you to lead the team without having to perform any tasks personally.’
Team Present	Two actors present	Three actors present
Scenario Progress	Start conditions depict ventricular fibrillation (VF), which persists until the end of the scenario.	Start conditions depict asystole, which persists until the end of the scenario.

Table 8-1: Summary of resuscitation scenarios

Both scenarios began with the participant entering the simulation room to find CPR in progress. The actors playing staff members took the role of competent life support providers, and worked according to a script. These actors wore a radio ear-piece so that their actions or prompts could be standardised, as far as was practical within the context of the scenario. Resuscitation equipment was not initially available, although was delivered at the participants’ request.

A series of critical stimuli were designed into each scenario. Each scenario comprised four critical events. Certain items of equipment supplied to the team were engineered to fail on cue during each scenario, and the actors were also trained to perform certain actions at specified points. There were eight potential events, grouped into two sets.

Although participants always experienced the high-workload scenario first, the event set that occurred in that scenario was allocated to each participant using a random number generator, such that participants could experience event list A in the high workload scenario, and list B in the low workload scenario, or vice versa.

The triggers for each stimulus were event-based, rather than strictly time-based. The hat-change (list A) or person-switch (list B) for example always happened when a team member was instructed to take a blood gas sample to the analyser (Table 8-2).

Repeated inattentional blindness has been demonstrated in the same candidates within the same session, so it was realistic to attempt data collection using a repeated-measures design (11).

Summary of event triggers

		Description	Trigger
Event Set A	Shirt swap	Team member changes from blue to green shirt	Team member instructed to leave room and collect defibrillator
	IV flush	Team member administers a saline flush to the patient cannula without explanation	When CPR pauses for first rhythm check
	O2 disconnection	Oxygen disconnects from the wall-mounted flowmeter	Team leader lists potential reversible causes of cardiac arrest
	Hat swap	Team member changes hat from green to blue	Team member instructed to leave room to process blood gas sample
Event Set B	Stethoscope drop	Team member discards the stethoscope from around their neck whilst out of the room	Team member instructed to leave room and collect defibrillator
	Cannula disconnection	IV cannula in the patient's left arm, into which fluids are running, is pulled out	30 seconds following first rhythm check
	Defibrillator disconnection	The sternal-lead from the defibrillator is disconnected from the patient's chest	Team leader lists potential reversible causes of cardiac arrest
	Person switch	One team member leaves with a blood sample, while another (previously unseen) returns with the result	Team member instructed to leave room to process blood gas sample

Table 8-2: Summary of critical events

During each scenario participants were free to ask for advice from the actors, and all students were given a printed copy of the current Advanced Life Support guidelines to which they could refer.

Whilst managing the scenario each volunteer wore the wireless eye-tracking system described in Appendix 10.2.3. Recordings included measurement of fixations on specific regions of interest (Table 8-3), and variance in fixation co-ordinates.

Eye-tracking ‘Regions of interest’ specified in each scenario

ALS guideline printout
 Blood gas results printout(s)
 Clinical monitor
 Patient’s notes
 Patient’s chest

Table 8-3: Regions of interest

On completion of each scenario participants were asked to rate the scenario workload using the NASA Task Load Index (NASA-TLX) (12), as described in Appendix 10.2.9.

After rating the workload for the scenario participants completed a short questionnaire. To minimise the risks of participants guessing the true nature of the study these questionnaires included generic questions about the conduct of the resuscitation, as well as open-ended, free-text questions. Firstly, participants were asked to list all cardiac rhythms that were noticed, and all drugs given to the patient. They were then invited to describe any deviations from Resuscitation Council Advanced Life Support guidelines, any communication issues with their team, and any problems with their equipment.

After completing the scenario two questionnaires, an additional questionnaire was administered. In line with existing perceptual error research these were more specific (6), as it appears participants are better able to report changes on prompting than spontaneously (13). Participants were therefore presented with a tick list of items which included some false positives (Table 8-4). Responses were collected using a four-point Likert scale ranging from ‘certain that I did not notice’ to ‘certain I did notice’.

Event choices presented to participants after each scenario

'Did you see any of the following in the previous scenario?'

I had a different number of people on my arrest team at the end
 Members of my arrest team were different at the end
 Members of my arrest team were dressed differently at the end
 There were problems with the oxygen supply
 There were problems with the defibrillator
 There were problems with the IV line
 There were problems giving medications
 Items of equipment went missing

Table 8-4: List of change events from which participants were prompted to choose

On completion of the experiment each participant was fully debriefed, given additional information about the true nature of the study, and offered an opportunity to withdraw from the study. No participant wished to withdraw from the study after this debriefing.

8.3.5 Analysis

Continuous variables were analysed using the paired-samples t-test or a suitable non-parametric alternative as necessary. Proportional data were analysed using Chi-squared analysis, or Fisher's exact test as needed. Two-sided p-values are quoted for all tests. Logistic regression was planned to assess predictors of noticing for each event.

8.4 Results

This experiment recruited 20 students (12 males, 8 females), all of whom had recently undertaken clinical attachments in either emergency medicine/trauma or anaesthesia. The students had a mean age of 24.4 (SD 2.50), and all had been exposed to simulation-based teaching before. Eleven students received event set A, and 9 students were assigned event set B.

8.4.1 Workload Ratings

The NASA-TLX demonstrated that the workload manipulation had been broadly successful. The mean weighted score for the high-workload scenario was 70.2 (SD 11.8), and for the low scenario 61.1 (SD 16.0).

The weighted scores were normally distributed, as assessed by the Shapiro-Wilk test ($p = 0.54$). The fall in NASA-TLX scores between workload conditions was significant ($p < 0.01$), and mental demand was consistently the most significant factor in the students' workload ratings, followed by performance concerns. These data are summarised in Figure 8-1.

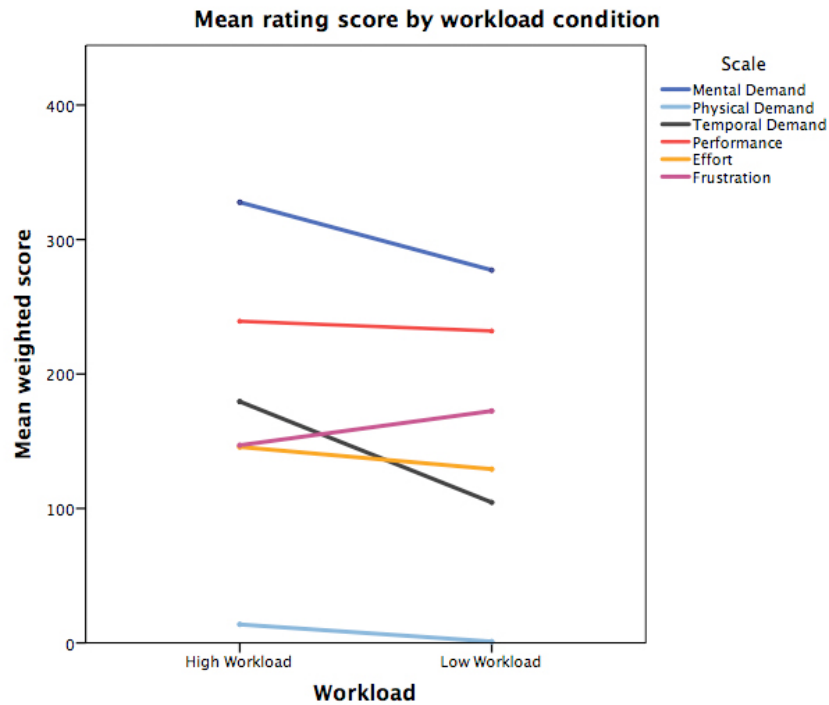


Figure 8-1: NASA-TLX scores

8.4.2 Event Timing

The critical events, although scripted, were triggered by participants' actions. The timing of each is summarised in Table 8-5. The mean duration of the high-workload scenario was 7m 55s (SD 35s); the low workload scenario was significantly shorter, at 6m 45s (SD 43s) ($p < 0.01$).

		Event trigger times (mm:ss)			
		Earliest	Latest	Mean	St Dev
Event Set A	Shirt swap	01:05	02:38	01:31	00:25
	IV flush	01:43	03:32	02:18	00:28
	O2 disconnection	02:21	05:03	03:30	00:48
	Hat swap	03:03	06:19	04:30	01:01
Event Set B	Stethoscope drop	00:33	01:38	01:00	00:18
	Cannula disconnection	02:02	04:12	03:05	00:47
	Defibrillator disconnection	02:08	06:02	03:50	01:17
	Person switch	01:57	06:03	03:54	01:35

Table 8-5: Summary of event timings

8.4.3 Noticing Rates

The noticing rate for each event, split by workload condition, is presented in Table 8-6. The most frequently reported event was the change in team member effected during blood gas processing in event set B, which was noticed by 15 of 20 people. No participants reported noticing the loss of the stethoscope from one team member in event set B. A binary logistic regression model was created to test the predictive value of participant age, gender, workload condition, and NASA-TLX scores. Workload was the only significant predictor in the regression model, and then only for detection of the clothing change, the oxygen disconnection, and the switch in team member. There were no significant predictors for the other events. These data are also summarised in Table 8-6. In total 25 of 80 (31.25%) events were detected under high load conditions, compared with 35 of 80 (43.75%) in the low load scenario.

Most events were detected more frequently in the low workload condition, or at least at the same frequency as in the high-load condition. The oxygen disconnection was unique in that it was the only event that was detected more frequently in the high-workload condition. False positive events were reported by five participants. One participant reported an oxygen disconnection in a set B/low workload scenario, and one reported an oxygen disconnection in a set B/high workload scenario. The remaining three participants all reported person switches in set A/low workload scenarios.

Noticing rates for each event by workload condition

		Seen in High Workload conditions (n=11)	Seen in Low Workload Conditions (n=9)	p value
Event Set A	Shirt swap	3 (27.2%)	7 (77.8%)	0.03
	IV flush	0 (00.0%)	2 (22.2%)	1.00
	O2 disconnection	9 (81.8%)	3 (33.3%)	0.04
	Hat swap	2 (18.2%)	2 (22.2%)	0.82
		Seen in High Workload Conditions (n=9)	Seen in Low Workload Conditions (n=11)	p value
Event Set B	Stethoscope drop	0 (00.0%)	0 (0.0%)	N/A
	Cannula disconnection	2 (22.2%)	3 (27.2%)	0.80
	Defibrillator disconnection	4 (44.4%)	8 (72.2%)	0.21
	Person switch	5 (55.6%)	10 (90.9%)	0.10

Table 8-6: Noticing rates for each event, by workload condition; p value is that derived from the logistic regression model using workload condition as a predictor

The experiment described in the preceding chapter collected data on clinicians' insight into the phenomenon of perceptual error. Participants in that experiment were asked to predict if they would notice many of the events included in the simulation scenarios used here. For reference these predictions are replicated in Table 8-7. The data from the simulated live study described in this experiment can now be used for comparison against the predictions made in the previous chapter.

Predicted versus actual noticing rates in the live scenario

		Predict noticing in live scenario (n=125)	Actual noticing in live scenario (n=20)	p value
Event	Hat change	31 (24.8%)	10 (50.0%)	0.03
	Team member switch	104 (83.2%)	15 (75.0%)	0.36
	IV line disconnection	65 (52.0%)	5 (25.0%)	0.03
	Oxygen disconnection	91 (72.8%)	12 (60.0%)	0.29

Table 8-7: Comparison of predicted frequency of noticing each event encountered in the live scenario compared with observed noticing frequencies

Finally, comparison can be made between the noticing of events in a live simulation scenario with the noticing-rate for events that were also depicted in the video-based experiment described in Chapter 6. In all cases the proportion of participants who noticed the shared events significantly differed (Table 8-8). Generally, the participants in simulation over-performed compared with the group that participated in the video-based experiment; the hat change, team-member switch, and oxygen disconnection were all noticed significantly more frequently. The stethoscope drop was the only event less likely to be noticed in simulation. As described in the previous chapter, the noticing

of the team-member switch was approximately twice as likely as noticing the change in conversation-partner during Simons' 'Door Experiment' (14).

Comparisons with other experiments may also be fruitful. Levin conducted an experiment in 2002 modifying the original 'Door Experiment' methodology. In this follow-up, the person switch was disguised in one of two ways (15). Overall 22 of 56 people noticed the person-switch across the two conditions⁷. The clinical cohort in the current experiment performed significantly better than the participants in Levin's study ($p < 0.01$).

Finally, an analogue of the stethoscope drop has also been attempted in a real-world non-clinical setting. Simons reported a series of single-object changes made to a person. In this case sports equipment of varying types and conspicuousness were removed from the items carried by a collaborator. Two simple object-removals were made across the experiments reported in this paper (13), and 17 of 18 participants overall noticed the removal of the balls. Six reported the loss spontaneously and 11 describing the event on cueing⁸. Again, this significantly differed from the proportion that noticed the stethoscope removal in this experiment ($p < 0.01$), although in this case the clinicians noticed the item-loss less frequently.

Noticing rates in the live scenario versus video scenario

	Actual noticing in video scenario (n=142)	Actual noticing in live scenario (n=20)	<i>p</i> value
Event			
Hat change	11 (7.7%)	10 (50.0%)	< 0.01
Stethoscope drop	24 (16.9%)	0 (0.0%)	0.03
Team member switch	36 (25.4%)	15 (75.0%)	< 0.01
Oxygen disconnection	34 (23.9%)	12 (60.0%)	0.02

Table 8-8: Comparison of predicted frequency of noticing each event in live scenario versus video-based stimuli

⁷ Two conditions were used in Levin's experiment; the first was a person-switch disguised by an office counter; in this case only five of 20 noticed the change. The second condition disguised the person switch using the act of taking a photograph; 17 of 36 noticed the change in this condition.

⁸ In the first experiment two of five reported the loss of a standard basketball spontaneously, and three of five described the change after cueing. In the second experiment, conducted with a more conspicuous ball, four reported the change spontaneously and eight required prompting.

8.4.4 Eye-tracking results

The eye-tracking system developed a fault in the radio receiver during the first recording session. It was sent for repair, but the fault persisted on its return. The eye-tracking results were therefore subject to a high degree of data-loss (Table 8-9).

The system successfully tracked the eyes for less than one third of the total recording time (31.1%). For only two subjects (recordings 07 and 09) was there a full recording of both high and low workload scenarios.

Eye-tracking: summary of data-loss						
Recording number	Scenario				Eye tracking validity	Comment
	High Workload		Low Workload			
	Scene	Eye	Scene	Eye		
01	✓	✓	✓	✗	40.9%	
02	✓	♦	✓	✗	14.6%	
03	♦	♦	✓	✓	56.4%	
04	♦	♦	✓	✓	57.4%	
05	♦	✗	♦	♦	15.9%	
06	♦	♦	✓	✓	49.3%	
07	✓	✓	✓	✓	90.0%	
08	✓	✓	✗	✗	18.1%	
09	✓	✓	✓	✓	80.2%	
10	✗	✗	✗	✗	-	System removed for repairs during this session
11	✗	✗	✗	✗	-	System removed for repairs during this session
12	✗	✗	✗	✗	-	System removed for repairs during this session
13	✗	✗	✗	✗	-	System removed for repairs during this session
14	✗	✗	✗	✗	-	System removed for repairs during this session
15	✗	✗	✗	✗	-	System removed for repairs during this session
16	✓	✓	✓	♦	49.9%	
17	✓	♦	♦	♦	20.6%	
18	♦	♦	♦	♦	17.9%	
19	♦	♦	♦	♦	49.3%	
20	♦	♦	♦	♦	65.3%	

Table 8-9: Summary of data collection using eye-tracking system; key: ✓ = full recording available; ♦ = partial recording available; ✗ = no recording available

All areas of interest were fixated more often in the low-workload condition (Figure 8-2), and fixation co-ordinates were more widely distributed in the x-axis (1599.8 vs 1231.0 pixels; $p = 0.354$). None of these trends reached significance, and given the extremely poor quality of data available little weight should be put on these comparisons.

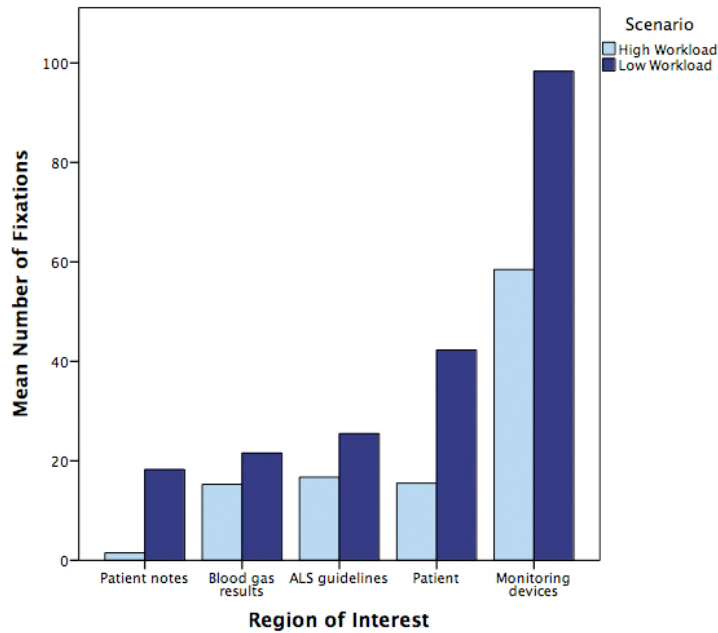


Figure 8-2: Fixations by region of interest

8.5 Discussion

This experiment demonstrates that perceptual errors do occur in live clinically-relevant scenarios, but these data suggest the incidence of such effects is less than might have been predicted based upon previously published data.

The workload manipulation was predicated upon the ALS algorithm. The arrest-rhythm in the high-workload scenario was pulseless ventricular tachycardia, which necessitates defibrillation. This activity requires planning and co-ordination of the team, adding to mental demands. A requirement to deliver the treatment, coupled with a need for the participant to manage the patient’s airway, greatly added to the physical and temporal demands. In addition, the high-workload patient was hyperkalaemic, an important precipitant of cardiac arrest that required additional treatment.

There were no ‘reversible causes’ of cardiac arrest to be addressed in the low workload scenario; this patient had asystole, and required only the delivery of CPR. The provision of a full resuscitation team also freed the experimental participant to provide leadership as their sole task. Although the workload manipulation was successful, evidenced by a significant fall in NASA-TLX score, there was

one measure that rose between conditions. Frustration scored more highly in the low-workload scenario, and it is likely this occurred in consequence of the very factors that made the physical and mental demands lower in the second scenario. In the first scenario, there were obvious targets for treatment, a shockable heart rhythm and a reversible cause, and therefore there was a reasonable expectation of a successful outcome. By contrast, the likelihood of survival from a non-shockable cardiac arrest with no reversible causes is likely to be under 10% (16), and this probably affected the students' ratings.

In general, all events were more likely to be noticed in the low workload condition. This was consistent with expectations, at least in direction, if not in magnitude. The one exception to this trend was the oxygen disconnection, which was far more likely to be noticed in the high workload condition. This surprising result may have arisen from the physical layout of the scenario room. It was necessary that the students be given the airway to manage during the scenario, as they were required to perform a continuous task throughout the resuscitation. It would not have been plausible to expect them to offer leadership whilst performing chest compressions, and whilst the students had all received training in automated external defibrillator (AED) use, training in manual defibrillation was not universal. A manual defibrillator was necessary to test if the participant noticed the disconnection because an AED would have verbally and continually prompted the participant to check pad connections as soon as the cable was loosened.

This placed the participant at the head of the bed, and by necessity next to the wall-mounted oxygen supply. It was observed that several volunteers were struck by the falling tubing as it disconnected from the flowmeter. Although change blindness has been demonstrated for tactile stimuli (17), it is likely that the combination of proximity (thereby increasing the perceived volume of the free-flowing gas) and the possibility of additional tactile transients made the detection more reliable.

The low rate of false-positive reports suggests that participants did not resort to guessing when completing the post-scenario questionnaires. The most frequently reported false-positive event was

to report a person switch in set A/low workload scenarios. It is possible that the participants noted the change in clothing, and associated that with a switch in team member's identity. This may have been more prominent because these specific individuals had already experienced a set B/high workload scenario in which there had been a change in personnel, and it is likely these participants were then primed to expect such an event.

Although these data demonstrate both change- and inattentional-blindness, the participants in this simulation study over-performed when compared to the previous video-based experiment. It is known that an intention to search for changes improves performance (18), and the nature of the experiment may have encouraged participants to adopt a strategy of data collection and comparison. Although they were not specifically cued to expect the perceptual error stimuli incorporated in the scenarios, cardiac arrests are dynamic and evolving scenarios, so to an extent the need to look for stimuli akin to change- and inattentional-blindness probes is implicit in the task briefing.

Contrast this with participants who watched the video resuscitation in the earlier experiment (7). These people were briefed to observe the team-members, rather than the patient, and to comment on their behaviour and decision-making. This made change-detection incidental to their primary task. One imagines that video-viewers were less involved in the scenario, perhaps more prone to distraction by task-unrelated thoughts.

Physical involvement in the scenario may also have modulated risk. Hutterman demonstrated that moderate levels of physical activity reduce the risk of inattentional blindness, although high levels of exertion increase risk again (19). This seems consistent with personal observation that the CPR provider, the team-member with by far the highest physical workload, often has the lowest situational awareness in both real- and simulated-emergencies.

There are other reasons to expect a difference between video- and simulation-based results. Differences have been observed in the way 2D and 3D scenes are examined, not least because 2D images are presented on a single plane, so there is no need to continually refocus on objects at different distances (20). Determining gist and extracting semantic information about a scene is more rapid when viewing naturalistic images (21), and a crucial human-interactive element is missing from most laboratory stimuli. Social cues do affect how people search for information (22); elements such as the gaze- or head-position of other people involved can, for example, successfully cue towards change-events (23). It is unfortunate the eye-tracking data available in this experiment is so limited. It offers tantalising hints that there are significant differences in the distribution of attention between workload conditions, but the high degree of data loss prevents meaningful conclusions being drawn from them.

These social cues, coupled with the realism of performing a natural task, may also give rise to difference in cognitive engagement between video and real-world tasks. The work of Vo demonstrated that simply looking at a scene (and even attempting to memorise its contents) does not improve search-performance, but searching for a specified target a second time in the display is substantially quicker (24). It was concluded that performing task-relevant actions with an object increased the likelihood of registering the item in memory.

Related effects have been observed in laboratory-based change- and inattention-blindness studies. Cueing with non-task-relevant words did not improve noticing rates (25) and fixations on-target, even at the moment of change, do not improve detection if the change is non-task-relevant (26). Also one of the few studies that has shown transferrable expertise effects may have demonstrated this indirectly. The experiment was conducted with professional proof-readers, who detected scene changes in photograph flicker-tasks more reliably than did general readers, and were also less prone to spatial biasing (non-proof-readers were more likely to miss changes in the lower half of the presented stimuli) (27). Asano attributed this to better attentional control amongst the proof-

readers, but one might speculate that cognitive engagement may also differ between groups. Reading passages generally requires a surprisingly superficial engagement with the text on the page – provided the start and end letters are correct, the internal letters of a word can be anagrammed⁹, the text remains understandable (28). The general reader requires only to get the meaning of the word – the gist of the scene, perhaps – so internal typographical errors and letter transpositions that do not alter the meaning of the word may pass unnoticed. Professional proof-readers must engage much more deeply with the text by examining every letter. Taken together it seems reasonable to conclude that participants watching a video, therefore more distanced from the event, interact with it more superficially, and are less likely to generate memories for inspected items, leading to more change-blindness in non-live situations.

Some of these effects may explain why complex scenes with multiple actors are more difficult to recall than scenes involving a single agent (29), an effect attributed to the need both to remember the distinguishing features of the actor, and also to attribute actions to a specific individual. Misattribution of actions has been termed '*unconscious transference*' (30), and this has been suggested as a potential source of error in eyewitness testimony. This might have been expected to contribute to those cases where the person-switch was missed. One can imagine clinical implications of misattribution from a practical perspective, in terms of which team members have been directed to carry out particular tasks, and also on a medicolegal level where individual identification may be important to reconstructing a sequence of events.

Another difference between video and live situations is that fixed camera positions are used in most perceptual error research (and all the popular demonstrations of inattention blindness). This may be a potential confounding factor. It is somewhat unrealistic to expect a person to maintain a fixed

⁹ For example: 'Raeding Wrods with Jubmled Lettres', taken from the title of a paper examining this phenomenon is understandable, albeit somewhat slower than reading plain text. A popular internet meme in the early 2000s was based on this type of letter transposition: <http://www.mrc-cbu.cam.ac.uk/people/matt.davis/cmabridge/> [sic]

head-position in such a situation, so their perspective of the scene will be constantly changing. There is some evidence from change-blindness studies that changing viewpoint does affect noticing; fewer eye-movements are needed to detect and confirm changes when they are superimposed, such as in a flicker task, than when they are presented adjacently (31). It is interesting therefore to consider how inattentional blindness might vary in video-based studies with a moving camera position.

It is known that the distribution of attention across a scene and the spatial location in which a change-event occurs is important in naturalistic tasks, and also that task-relevant experience has an effect controlling this distribution (32). Visual attention varies during actions, narrowing on task-relevant objects at critical moments, and much of this processing occurs in 'real time' as the data are needed (33). This means that the risk of perceptual error varies depending on whether the individual is personally performing the task, or simply observing someone else complete the same activity (34, 35). It was hoped that the eye-tracking system would permit some useful insights in this regard, but the high degree of data loss make these data difficult to interpret reliably. It is plausible to expect workload to have influenced attentional prioritisation, as when performing any task such as airway management or CPR it is necessary to devote some attention to the physical aspects of each activity. Opportunities for seeking other points of information elsewhere are then lost.

Finally, it is possible that some of the apparent over-performance is a result of the specific methodology of the study. These scenarios were six to seven minutes long, a relatively long duration for such studies, but required to make the scenario clinically meaningful. This meant that participants were exposed to the pre- and post-change objects for a protracted period, which might have offered better opportunities for registration or comparison of items. While perceptual errors may have occurred at the point of the event, the long duration of exposure afforded additional opportunities for reassessment leading to detection. Review of the scenario video-recordings supports this hypothesis; three of the twelve participants who noticed the defibrillator disconnection did so only at the point of a rhythm check when they saw an alert on the defibrillator

display screen. Although there are a few clinical situations where immediate response is required, in most routine situations delaying detection by a matter of a few minutes is unlikely to affect outcome significantly.

Kreitz observed that fast-moving but salient objects can still be missed, and concluded that exposure time is a key variable determining likelihood of noticing (36). Previous work has demonstrated that eye-movements do not appear to predict inattentional blindness risk, and fixation on target does not guarantee noticing it (37-39). Kreitz concluded that the exposure-time effect occurs because as time increases, the likelihood of covert attention chancing across the critical stimulus rises. In the present experiment participants had an average of one minute's exposure to the pre-change object, and two minutes' exposure to objects post-change; many of the events had substantially longer pre- and post-change exposure times than this. These far exceed the shorter observation times in most perceptual error studies, including the Greig CPR video (7), which is only around one minute in length.

Although perceptual errors did not happen as commonly as predicted, the risk was non-zero. Overall eight of the 20 participants missed the oxygen disconnection, a further eight did not see the defibrillator disconnect, and 15 missed the loss of the patient's cannula. Only two noted administration of the unexpected IV drug to the patient. Any of these events could materially impact on delivery of care. After good quality CPR, prompt defibrillation of a patient with a shockable arrest-rhythm is the only treatment for which there is robust evidence. Oxygen therapy is mandatory, and missing drug administration could result in doubled-up doses and excess toxicity.

A question then arises, what can be done to mitigate these effects? Attempts have been made to provide solutions to improve situational awareness in industrial settings using specifically designed tools intended to improve resilience to distraction (40). Evidence of efficacy is mixed, and some studies have even shown that awareness can be degraded by the use of such supports, probably because manipulating the tools is an additional mental burden (41).

Engineering solutions have also been attempted using heads-up displays or augmented reality systems to present task-relevant information. Again evidence is mixed; when a heads-up display was trialled in civilian airline pilots there was an increase in runway-incursions associated with use of the device (42). Head-mounted displays have been trialled in anaesthesia (43, 44), and in surgery (45-47), and in both cases there are data to suggest inattentive blindness risk actually increases. This phenomenon is termed 'cognitive capture', whereby attention is captured by the material presented on the display, at the expense of collecting information from the environment (48). Increasing information availability therefore does not always improve situational awareness, even when that information is task-relevant (49).

The quality of design in the display is crucial, and as a general principle to reduce the risk of perceptual error *'as much as possible information from different sources should be integrated before it is presented to the operator. This will lessen the need to shift attention across different information sources'* (50). Such principles are being adopted in military display design (51) but little evidence of this yet exists in healthcare (52). System-generated alerts must also be carefully managed to prevent 'alarm fatigue' (53), and excessive dependence on technology can generate 'automation complacency' (54), whereby users become so used to the system detecting hazards that they respond poorly to dangers the system has not highlighted.

It is likely that the best opportunity to mitigate the effects of perceptual error is therefore not technical. Nor as the previous chapter demonstrated, can it be simply highlighting the occurrence of perceptual error as tends to happen currently. Instead intervention probably must be based upon teamwork. Techniques developed by the military, such as snap-briefing and sit-rep updates (55), may have useful application in this regard. A 'snap brief' is a short, focussed pre-event briefing confirming crucial details and objectives, while a 'sit-rep' (situation report) is a brief pause in tasks to allow all participants to share information about progress and priorities. One can imagine that

deployment of such techniques, such as pre-programmed 'sit-rep' on completion of the primary-survey of a critically-ill patient, might well improve team- and shared-situational awareness.

8.5.1 Summary

This experiment has demonstrated that perceptual errors do occur in 'live' medical situations, albeit at lower levels than were predicted based on the video-based experiment described in Chapter 6.

The successful manipulation of workload conditions also demonstrated that perceptual errors affecting clinicians probably are sensitive to task demands, but that situational features, such as proximity to an event, may override this effect. It is disappointing that hardware failures prevented further investigation of the distribution of attention across the different workload and experience conditions.

Perhaps the most important point yielded by this experiment, of relevance to all researchers working in this field, is that the incidence of perceptual error is sensitive to the modality by which it is tested. The decreased incidence in simulation compared to using video-based materials suggest that exposure time, coupled with the additional prompts and stimuli available in a three-dimensional environment, are important to event-detection. This will have wide implications for any work in which researchers intend to generalise their findings to real-world situations.

8.6 References

1. House of Commons Health Committee. Patient Safety: 6th report of session 2008-2009. London: The Stationery Office; 2009.
2. Singh H, Petersen LA, Thomas EJ. Situational awareness in medicine. *Qual Saf Health Care*. 2006;15(5):384-.
3. Walker GH, Stanton NA, Kazi TA, Salmon PM, Jenkins DP. Does advanced driver training improve situational awareness? *Appl Ergon*. 2009;40(4):678-87.
4. O'Brien KS, O'Hare D. Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*. 2007;50(7):1064-91.
5. Endsley M, Garland D. Theoretical underpinnings of situation awareness: a critical review. *Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
6. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattentive blindness for dynamic events. *Perception*. 1999;28(9):1059-74.
7. Greig P, Higham H, Nobre A. Failure to perceive clinical events: an under-recognised source of error. *Resuscitation*. 2014;85:952-6.
8. Smilek D, Eastwood JD, Reynolds MG, Kingstone A. Metacognitive errors in change detection: missing the gap between lab and life. *Conscious Cogn*. 2007;16(1):52-7; discussion 8-62.
9. Smilek D, Eastwood JD, Reynolds MG, Kingstone A. Metacognition and change detection: Do lab and life really converge? *Consciousness and Cognition*. 2008;17(3):1056-61.
10. Murphy G, Greene CM. High perceptual load causes inattentive blindness and deafness in drivers. *Visual Cognition*. 2015;23(7):810-4.
11. Ward EJ, Scholl BJ. Inattentive blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*. 2015;22(3):722-7.
12. Human Performance Research Group. NASA Task Load Index (TLX) v1.0. NASA Ames Research Center, Moffett Field, California; 1988.
13. Simons DJ, Chabris CF, Schnur T, Levin DT. Evidence for preserved representations in change blindness. *Consciousness and Cognition*. 2002;11(1):78-97.
14. Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998;5(4):644-9.
15. Levin DT, Simons DJ, Angelone BL, Chabris CF. Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology*. 2002;93(Pt 3):289-302.
16. Nolan J, Soar J, Lockey AS, Perkins G, Pitcher D, Gwinnutt C, et al., editors. *Advanced Life Support*. 6th ed. London: Resuscitation Council (UK); 2011.
17. Van Hulle L, Van Damme S, Spence C, Crombez G, Gallace A. Spatial attention modulates tactile change detection. *Experimental Brain Research*. 2013;224(2):295-302.
18. Beck MR, Levin DT, Angelone B. Change blindness blindness: Beliefs about the roles of intention and scene complexity in change detection. *Consciousness and Cognition*. 2007;16(1):31-51.
19. Huttermann S, Memmert D. Moderate movement, more vision: Effects of physical exercise on inattentive blindness. *Perception*. 2012;41(8):963-75.
20. Welsby PD. The neurophysiology of failed visual perceptions: some implications for medical teaching. *Postgrad Med J*. 2016;92(1084):74-7.
21. Li FF, VanRullen R, Koch C, Perona P. Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*. 2002;99(14):9596-601.
22. Risko EF, Laidlaw KEW, Freeth M, Foulsham T, Kingstone A. Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*. 2012;6:143.
23. Sato W, Okada T, Toichi M. Attentional shift by gaze is triggered without awareness. *Experimental Brain Research*. 2007;183(1):87-94.

24. Vo MLH, Wolfe JM. When Does Repeated Search in Scenes Involve Memory? Looking At Versus Looking For Objects in Scenes. *Journal of Experimental Psychology: Human Perception & Performance*. 2012;38(1):23-41.
25. Kreitz C, Schnuerch R, Gibbons H, Memmert D. Some see it, some don't: Exploring the relation between inattention blindness and personality factors. *PLoS ONE*. 2015;10(5).
26. Smith TJ, Lamont P, Henderson JM. The penny drops: change blindness at fixation. *Perception*. 2012;41(4):489-92.
27. Asano M, Kanaya S, Yokosawa K. Proofreaders Show A Generalized Ability To Allocate Spatial Attention To Detect Changes. *Psychologia*. 2008;51(2):126-41.
28. Rayner K, White SJ, Johnson RL, Liversedge SP. Reading Words With Jumbled Letters: There Is a Cost. *Psychological Science*. 2006;17(3):192-3.
29. Fitzgerald RJ, Oriet C, Price HL. Change blindness and eyewitness identification: Effects on accuracy and confidence. *Legal and Criminological Psychology*. 2016;21(1):189-201.
30. Davis D, Loftus EF, Vanous S, Cucciare M. 'Unconscious transference' can be an instance of 'change blindness'. *Applied Cognitive Psychology*. 2008;22(5):605-23.
31. Josephs E, Drew T, Wolfe J. Shuffling your way out of change blindness. *Psychonomic Bulletin & Review*. 2016;23(1):193-200.
32. Zhao N, Chen W, Xuan Y, Mehler B, Reimer B, Fu X. Drivers' and non-drivers' performance in a change detection task with static driving scenes: is there a benefit of experience? *Ergonomics*. 2014;57(7):998-1007.
33. Hayhoe MM, Shrivastava A, Mruczek R, Pelz JB. Visual memory and motor planning in a natural task. *Journal of Vision*. 2003;3(1):6-.
34. Smith TJ. The role of audience participation and task relevance on change detection during a card trick. *Frontiers in Psychology*. 2015;6:13.
35. Wallis G, Bulthoff H. What's seen and not seen: Influences of movement and task upon what we see. *Visual Cognition*. 2000;7(1-3):175-90.
36. Kreitz C, Furley P, Memmert D. Inattention blindness is influenced by exposure time not motion speed. *Quarterly Journal of Experimental Psychology*. 2016;69(3):495-505.
37. Beanland V, Pammer K. Looking without seeing or seeing without looking? Eye movements in sustained inattention blindness. *Vision Research*. 2010;50(10):977-88.
38. Koivisto M, Hyona J, Revonsuo A. The effects of eye movements, spatial attention, and stimulus features on inattention blindness. *Vision Research*. 2004;44(27):3211-21.
39. Memmert D. The effects of eye movements, age, and expertise on inattention blindness. *Consciousness and Cognition*. 2006;15(3):620-7.
40. St John M, Smallman HS. Staying Up to Speed: Four Design Principles for Maintaining and Recovering Situation Awareness. *Journal of Cognitive Engineering and Decision Making*. 2008;2(2):118-39.
41. Vallières BR, Hodgetts HM, Vachon F, Tremblay S. Supporting Change Detection in Complex Dynamic Situations: Does the CHEX Serve its Purpose? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2012;56(1):1708-12.
42. Crawford J, Neal A. A Review of the Perceptual and Cognitive Issues Associated With the Use of Head-Up Displays in Commercial Aviation. *International Journal of Aviation Psychology*. 2006;16(1):1-19.
43. Liu D, Jenkins SA, Sanderson PM. Patient monitoring with head-mounted displays. *Current Opinion in Anaesthesiology*. 2009;22(6):796-803.
44. Liu D, Jenkins SA, Sanderson PM, Watson MO, Leane T, Kruys A, et al. Monitoring with head-mounted displays: performance and safety in a full-scale simulator and part-task trainer. *Anesthesia and Analgesia*. 2009;109(4):1135-46.
45. Dixon BJ, Daly MJ, Chan H, Vescan AD, Witterick IJ, Irish JC. Surgeons blinded by enhanced navigation: The effect of augmented reality on attention. *Surgical Endoscopy and Other Interventional Techniques*. 2013;27(2):454-61.

46. Dixon BJ, Daly MJ, Chan HH, Vescan A, Witterick IJ, Irish JC. Inattention blindness increased with augmented reality surgical navigation. *American Journal of Rhinology & Allergy*. 2014;28(5):433-7.
47. Marcus HJ, Pratt P, Hughes-Hallett A, Cundy TP, Marcus AP, Yang GZ, et al. Comparative effectiveness and safety of image guidance systems in neurosurgery: a preclinical randomized study. *Journal of Neurosurgery*. 2015;123(2):307-13.
48. Varakin DA, Levin DT, Fidler R. Unseen and Unaware: Implications of Recent Research on Failures of Visual Awareness for Human-Computer Interface Design. *Human-Computer Interaction*. 2004;19(4):389-422.
49. Marusich LR, Bakdash JZ, Onal E, Yu MS, Schaffer J, O'Donovan J, et al. Effects of Information Availability on Command-and-Control Decision Making: Performance, Trust, and Situation Awareness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2016;58(2):301-21.
50. Durlach PJ. Change Blindness and Its Implications for Complex Monitoring and Control Systems Design and Operator Training. *Human-Computer Interaction*. 2004;19(4):423-51.
51. McFarlane DC, Latorella KA. The Scope and Importance of Human Interruption in Human-Computer Interaction Design. *Human-Computer Interaction*. 2002;17(1):1-61.
52. Greig PR, Higham H, Nobre A, editors. The use of eye-tracking with simulated monitors: a case for standardisation? *Changing Behaviours: Association for Simulated Practice in Healthcare Annual Meeting; 2014 11/11/2014 - 13/11/2014; Nottingham*.
53. Hagenouw R. Should we be alarmed by our alarms? *Curr Opin Anaesthesiol*. 2007;20(6):590-4.
54. Ruscio D, Ciceri MR, Biassoni F. How does a collision warning system shape driver's brake response time? The influence of expectancy and automation complacency on real-life emergency braking. *Accident Analysis & Prevention*. 2015;77:72-81.
55. Arul GS, Pugh HEJ, Mercer SJ, Midwinter MJ. Human factors in decision making in major trauma in Camp Bastion, Afghanistan. *Annals of the Royal College of Surgeons of England*. 2015;97(4):262-8.

9 Conclusions

9.1 Summary of findings

There can be little doubt that medical error is one of the great intractable problems in modern healthcare. Chapter 1 demonstrated the scale of the problem, and that the associated human and financial costs are significant. Up to 35,000 people every day in the UK may be harmed by errors. Between the additional care costs that errors incur, and the sums consumed by legal actions, the cost to the National Health Service is measured in the billions. That considerably more than 1% of the total service budget must be diverted away from front-line care at a time when the service can little afford it is a matter of concern.

If a disease process were causing this degree of harm it would likely generate a great deal of research funding and interest from the profession, but it has taken some time for patient-safety to reach the level of prominence that it deserves. Some of the difficulty is that there is no single cause for medical error and, as described by Reason (1), a diverse range of individually relatively minor incidents generally combine to create a major adverse outcome. This means both that no single answer will solve the problem of error, and (because individual clinicians are often relatively distanced from the consequences of error) that the motivation to engage with the patient safety agenda can be weaker than hoped for. To the individual practitioner, it does not feel as though errors are a major problem. To many healthcare personnel, errors are still something that should not happen to professionals and if a mistake is made then the person involved must be at fault. This creates a culture where errors are hidden, and blame becomes the objective of post-incident investigation.

A more productive response, and one that high-reliability organisations adopt, is to accept that risk and error are inevitable, and to design systems that trap errors as early as possible. This requires an honest investigation into the causes of error, and reflection on the factors discovered.

It has been identified that human factors, and non-technical skills contribute to most medical errors. Prominent safety-incidents, such as the death of Elaine Bromiley (2), demonstrate that well-meaning professionals can fail with tragic results, in spite of appropriate training and equipment. A hallmark of this case was the apparent failure of the staff involved to recognise the seriousness of the situation they were attempting to manage, despite a wealth of information being available to them from a variety of sources. Other staff members, monitoring devices, and notably the patient herself, were all showing signs that the situation was much graver than the team-leaders understood.

Why then should these senior clinicians have apparently ignored all of this critically important information? Some might suggest surely they were deficient in some way? This has been the response of some investigators into other critical incidents; doctors have been convicted of manslaughter and spent considerable time in custody over medical errors¹⁰. There are however reasons other than negligence or inadequacy to expect situational awareness to fail in critical situations. Risks cannot be mitigated until they are understood, and the experiments presented in this thesis were therefore designed to shed light on some of the factors that might contribute such incidents.

A loss of situational awareness contributes to a great many medical errors. Developing an accurate mental model of a situation fundamentally underpins all medical decision-making, and management decisions suffer when clinicians' models diverge from reality. The psychology of situational awareness has attracted research interest, and influential authors such as Micah Endsley have

¹⁰ In the case of Wayne Jowett, who died after inappropriate administration of chemotherapy, one of the doctors was convicted of manslaughter and spent 11 months in custody. Reviews of this subsequently demonstrated a large number of staff- and process-failures in this patient's care that led to the doctors directly caring for the patient making the final, catastrophic error.

written a breadth of useful material on the subject. Endsley describes three levels of situational awareness (3); basic perception of information, integration of the data into a mental model, and application of the model to predict future events and guide decisions.

Acquiring relevant information is fundamental to the achievement of high-quality situational awareness. Most people conclude, based on daily experience, that this process is automatic, richly detailed, and robust. Chapter 2 demonstrated that this is not necessarily true, and a wide body of literature demonstrates that not only is perception highly fallible, but that few people recognise how limited their perception is. Classic demonstrations of the effect, such as Dan Simons' 'Invisible Gorilla' (4) and 'Door Experiment' (5) are well-known demonstrations of inattentional- and change-blindness that almost always elicit surprise when individuals first encounter it. Difficulty arises when trying to predict what effects these phenomena have on clinical activities, and there has not been a great deal of work conducted on medical populations that addresses this question.

The first question to be considered was how the medical profession encourages its trainees to acquire the skills necessary to achieve good quality situational awareness. Given that situational awareness is a fundamental skill, one might expect it to appear prominently in training objectives for junior medical staff. Unfortunately, it was shown in Chapter 3 that although certain specialties, notably those engaged in critical-care activities such as anaesthetics, emergency medicine, and intensive care, prioritise training in these skills, most medical disciplines barely mention them. Why these specialties in particular should have engaged with non-technical skills training earlier than most other medical disciplines may relate to a recognition that crisis situations represent particularly high-risk environments. It may also reflect the nature of critical-care teams and the environments in which they practice.

This is not to say that other specialties do not require such skills. All doctors face unexpected emergencies, and many doctors in medical and surgical specialties are faced with a need to assess acutely unwell patients. It is clear from the training syllabi that there are stark differences between

how these skills are taught to medical trainees. Many specialties still adhere to the traditional model that these professional traits can be acquired passively through the observation of suitable role-models. Few formal learning objects centred on non-technical skills exist outside of the three critical-care areas, and even fewer consider whether, and how, these skills should be formally assessed. It is also not clear that the medical profession can even agree on what terminology should be used to describe the breadth of non-technical skills required for professional practice.

Based upon the training curricula, one would assume that anaesthesia has widely adopted best-practices in human factors, but there are questions that can still be asked about how well understanding the theoretical aspects of human factors translates into the practical day-to-day working life. Chapter 4 demonstrated that a surprising lack of equipment standardisation exists even in vital theatre equipment. The anaesthetic monitor is the centre of safe practice and it is likely that high-quality monitoring is one of the biggest contributors to the improvements in anaesthetic safety seen over the last 20-30 years.

Despite the central role played by our monitoring equipment there is almost no standardisation in layout or colour scheme for the display of vital signs. This experiment successfully demonstrated that errors occur even when reading from familiar displays, and that clinicians can develop innate biases towards certain colours, one example being a conditioned expectation that green values denote heart-rate derived from ECG. These heuristics can work well until the clinician is exposed to a display that uses different colouration; to draw a parallel, most travellers would be somewhat discomfited to find out that the pilot of their aircraft had only discovered how their instrument panel was arranged when preparing for flight, yet that is a frequent condition in which anaesthetists find themselves at the start of an operating list.

Interestingly, the range of devices was such that very few anaesthetists were even able to recognise the equipment with which they worked on a regular basis. Despite this, an expertise effect was demonstrable in this experiment. From this one can conclude that non-standardised equipment

does contribute an additional cognitive burden. Users responded more slowly to unfamiliar displays, and made more mistakes. This is significant, because mental workload and expertise are fundamental contributors to errors of perception.

Given that all clinicians should be trained to understand situational awareness, Chapter 5 investigated whether such training conferred any generalizable perceptual advantages. This was tested using an observation task based on Simons' 'Invisible Gorilla', replicating the methods of the original experiment for a clinical population. This experiment demonstrated that not only did clinical background make no difference to noticing the critical stimulus, the study cohort performed identically to those recruited to Simons' original experiment. This clearly established the plausibility that clinicians can be affected by perceptual errors in dynamic tasks, and that there was no transferrable advantage conferred by their clinical training. The following experiment, described in Chapter 6 was the natural progression from this non-clinical study, and extended these findings in several important ways.

A specifically clinically-orientated video was designed to test for change- and inattention blindness. This sequence had to use a clinical scenario with which most healthcare professionals would have at least a basic understanding, so could not come from a highly-specialised branch of medicine. The experimental video was therefore based upon a resuscitation scenario, as all clinicians (including undergraduates) must undergo regular training updates in life support. Standards in life-support training are set nationally by the Resuscitation Council (UK), leading to a convenient, practical, and objective way to define expertise.

This video incorporated events with varying clinical significance, which offered a useful opportunity to comment on the effect of expertise. As predicted, the most experienced observers noticed the most clinically-relevant events but there was no effect of experience on the clinically trivial changes that were incorporated. This then demonstrated clearly that perceptual errors can affect clinicians using medically-orientated stimuli, and that the risk of error is sensitive to expertise and clinical

relevance. Notably even experts – instructors in the provision of advanced life support – demonstrated high levels of perceptual error for events that would meaningfully affect their patient. This was the first time such effects have been documented in a clinical setting using a dynamic task, and had one of the largest sample sizes of any perceptual error study to date.

The consistency with existing psychological data was reassuring and demonstrated that a real effect was being observed, but still questions could be asked about the effects of perceptual error in the workplace. Clinical practice is, after all, quite different to watching a video.

While planning the experiment to test for perceptual errors in more realistic scenarios an opportunity arose to investigate clinicians' insight into these failures. The data presented in Chapter 7 demonstrated that understanding of perceptual errors is relatively poor, and this aligned with predictions based on previous studies of 'change-blindness blindness'. This has been well described by psychologists, and it is known that both members of the general public and psychology students with deeper understanding of cognitive processes greatly overestimate their likelihood of noticing change- and inattentional-effects. This finding can now be extended to healthcare personnel.

Although many medical students and junior clinicians are now being exposed to the 'Invisible Gorilla' video at various stages in their training this does not appear to be associated with any teaching of the mechanisms that underpin it. Understanding of the cognitive processes driving perception appears superficial, and does not translate into accurate predictions of perceptual ability. It was clear that the clinicians forming the experimental cohort in Chapter 7 did possess some advantage over those participating in Levin's original change-blindness blindness experiment (6) in that their predictions were more pessimistic than previously observed. It would be misleading however to suggest that the group were accurate in their predictions, rather it should be said they were less inaccurate than those polled in the existing studies.

To comment on how perceptual effects impact on the clinical workplace it was necessary to explore perceptual error using the most realistic methods practicable. There are good reasons why differential performance rates might be expected between laboratory and live stimuli. Some have termed this the 'reel to real' effect (7), and Chapter 8 did demonstrate an important divergence between the detection of events in simulation with predictions based on the data from the video used in Chapter 6. It was clear that although perceptual errors did occur for a sizeable proportion of the candidates, with potentially meaningful effects, they significantly out-performed the cohort in the video-based study. Not only that, but detection of events such as clothing-changes and person-swaps were better detected than in previous real-world studies such as Simons' 'Door Experiment'. In the long term the importance of this finding may in fact outweigh the knowledge that clinicians are vulnerable to perceptual error, as these data can be generalised across the whole of perceptual research.

9.2 Methodological limitations of the programme of study

The exploration of perceptual error was deliberately designed with an incremental approach in mind. Predicting that clinicians would have limited understanding of perceptual errors it was felt most appropriate to take small steps, starting from establishing vulnerability using proven methods, and then stage-by-stage adapting those methods until the simulation experiment replicated a realistic clinical situation that the participant could experience directly.

The first video-based experiment, using methods and materials developed by Simons et al and described in Chapter 5, was conducted by recruiting volunteers to a booth set up at a national simulation conference. Although research staff did not choose the individuals who participated, it may be that there was a systematic bias predicated upon the type of clinician who would attend such a conference.

Despite this it is unlikely that the results were significantly skewed. One would predict that as clinicians with an educational and human-factors interest these clinicians should be the most aware

of performance limitations. That they performed exactly as the participants in Simons's original work suggests the data are reliable.

Greater numbers were required in the follow-up experiment using the bespoke clinically-orientated video. This required the move to on-line data collection. This has been used successfully in other fields, but it does carry the risk that the experimenter cannot control all the variables in the way they could if participants attended a laboratory. Comparisons of lab-based and online experimentation suggest that web-based data collection can yield results comparable to laboratory samples (8), but certain caveats must be considered. In this experiment, we must depend upon the honesty of participants, in their reporting of their observations, but also that they obeyed instructions, and adequately described any technical failures they encountered. One artefact of online data collection encountered in this study was that the design of the study materials appears to have provoked a post-completion error in the audience.

A post-completion error is one where the primary task is completed before all necessary steps have been conducted (9). In the case of the experiment, there was a page of explanatory text describing perceptual error before the final question, asking if participants had heard of these phenomena before, was presented. This was a design choice, because it was predicted few of the study respondents would have encountered terminology such as 'perceptual error' or 'inattentional blindness' before, but may have seen demonstrations before. This assumption was, of course, later confirmed by the experiment described in Chapter 7. It was intended that the description set a context for the final question, but it became apparent that several users had assumed the experiment was over at the explanatory text and logged off without moving to the final page. There were no significant differences between those participants who responded, and those who did not. It is therefore unlikely that the data-loss significantly biased the responses to these final questions.

Two of the events depicted in the video sequence used in this experiment might be expected to occur in a real cardiac arrest; a change in CPR provider or airway device is often effected during the

rhythm check, and therefore it might be suggested that some experimental participants may have noticed either but chose not to report it as a change.

This possibility was accounted for in the experimental design. The video was edited to ensure there was insufficient time depicted in which to actually perform these switches. There were no audio cues to suggest the team leader had requested such changes, nor were there any of the associated noises that one would expect when such tasks are carried out.

Regardless of the scripting, when collecting responses from observers the first questions encountered by the participant (before the nature of the experiment became apparent) requested free-text comments on the conduct of the resuscitation. At analysis, if the free-text comments reflected the participant having noticed either event, they were recorded as having noticed, irrespective of how they recorded responses on the 'tick list'.

Finally, if underreporting of these events were a significant problem, this activity would be most familiar to experienced providers who had managed cardiac arrest in real situations. This being the case, one would predict that any under-reporting of these events would be most prevalent in the experienced groups, but the results demonstrate this group were *more* likely to report the changes. It seems, therefore, that two events can still be considered valid cues to test for perceptual error.

Another issue arose in the experiment described in Chapter 8. A common difficulty in interpreting findings from perceptual error studies is examining the role of memory. Many experiments, including this one, tested the participants retrospectively. The final 'tick list' of specific events was presented for both experiments at the end of the second simulation. This was felt necessary, because to administer this questionnaire between scenarios the list would inappropriately cue participants before scenario two.

This had implications on data interpretation since a failure to describe an event may represent memory failure rather than perceptual error. To mitigate the effect of this, a broad definition of

perceptual error was adopted; a participant was recorded as having noticed an event if any of the following were true: there was a behavioural response evident during the scenario, the event was alluded to in any free-text response, or the appropriate box on the final questionnaire was ticked. This was the most pragmatic solution available to a thorny issue in studies of perceptual error.

The incremental approach to the overall programme of experiments carried a cost in time. The final simulation experiment was subject to a high degree of data loss from the faulty eye-tracking system. Had it been conducted earlier in the process there may have been time to repeat data collection with replacement hardware. It is unfortunate that these data were lost, because the eye-tracks that were available offered tantalising glimpses of the variations in visual search across workload conditions.

Considering the importance of ecological validity, one might ask whether such experiments should have been conducted in a real clinical environment. Although this would have yielded data most applicable to practice, there are several reasons to believe this would have been impractical.

Apart from the obvious confidentiality issues involved in recording data from a clinical space, it would be difficult to adequately consent patients for research participation in the acute situations in which perceptual errors are most likely. The large number of people involved in caring for such patients also adds many confounding variables that would be difficult to control for.

Finally, although perceptual errors may be relatively commonplace, it would be difficult to insert a controllable cue into a real clinical scenario that would not jeopardise the team's ability to care for their patient.

Despite these issues, one can only speculate if there remains a 'reel to semi-real to actually-real' effect. It is plausible to expect there may yet be differences between simulation and direct clinical care, because there is an added layer of emotional involvement brought by caring for an actual human being that a plastic manikin can never fully replicate.

9.3 Clinical and policy implications of these data

Non-technical skills acquisition is too important to be left to the vagaries of the 'hidden curriculum'. Formal learning objectives, where clearly stated, allow for trainers to set measurable standards, and permit the development of assessment tools suitable to the task. This creates objectivity, and makes the learning and assessment process fairer for trainees navigating the curriculum. It also ensures that everyone in a training programme understands non-technical skills, rather than just a few enthusiasts or those fortunate enough to be paired with supervisors who value these activities. This approach is evident in anaesthesia, which enjoys the benefit of high-quality non-technical skills rating systems that were developed by experienced clinicians in concert with industrial psychologists. Such systems exist for anaesthetists (10), anaesthetic practitioners (11), surgeons (12), and scrub teams (13), but they are most fully integrated into the training systems of anaesthesia. Having a well-designed system, that has been widely agreed, greatly facilitates the delivery of training by creating a shared understanding of non-technical skills, and arms practitioners with a vocabulary with which to describe them. High-reliability organisations should encourage multidisciplinary teams to train together, but without this shared understanding (or even a shared belief that such skills are important) it is likely that training will remain relatively patchy.

A similar degree of standardisation is required in equipment provision. Ergonomists have investigated human-machine interactions, and recommended four principles of interface design important to preserving situational awareness, and providing resiliency against interruption (14). These are: changes should be detected automatically and highlighted to the user; alerts should not detract from other mission-critical data or tasks; the display should provide a clear overview and prioritise information; and non-critical information should not clutter displays, rather it should be accessible on-demand. Clinical displays achieve the first (at least where changes cause parameters to fall outside pre-defined limits) and second points, but the other two principles are violated by most clinical displays.

Given the wealth of existing data, little research is needed to make improvements in both these fields, and therefore these are highly amenable to immediate intervention. The solutions will, however, require direction from national bodies. Professional groups such as the AAGBI, and organisations such as the Royal Colleges, can set standards within their respective professional groups, but these issues cut across specialties. Accordingly, top-level bodies such as the Department of Health or General Medical Council will need to set, and importantly enforce, the required standards. Very few clinicians have adequate understanding of ergonomics to make these decisions robustly, and therefore if left to individual institutions it is likely that simple commercial interests will continue to predominate.

Regarding the clinician individually, it is useful to explore how they misunderstand their perceptual reliability for two reasons. One is that if clinicians are excessively confident in their ability to detect important events they may be more prone to error in the workplace. They will search for data less effectively, using inappropriate search-strategies, or prioritising the wrong targets. An obvious consequence of this is their failure to then achieve basic situational awareness, but it may also lead the clinicians to undervalue the sharing and comparison of this basic-level information within the healthcare team during an emergency.

The second finding of importance from this study was that simply exposing people to demonstrations of perceptual error, whilst interesting and entertaining, does not appear to significantly improve their performance in change-detection tasks. A consistent feature of all the experiments presented herein is that simple awareness of perceptual error is not associated with better rates of noticing critical stimuli. This is a finding that has also been replicated in non-clinical scenarios (15). Non-clinical demonstrations lack the face-validity and personal relevance needed to provoke the meaningful reflection required to drive adult learning in this context, and they seem to have made little impact on the people tested here. The video developed for the experiment described in Chapter 6 may therefore prove to be a highly useful and readily deployable educational

tool, situating the learning experience within a clinically relevant and relatively commonplace setting.

When training clinicians to manage emergency situations educators must now raise awareness of vulnerability to perceptual errors, and this will lead them naturally to highlight the importance of good communication in maintaining team situational awareness. During an emergency, at any given time it is likely that individual members of the team will be experiencing loss of situational awareness. The probability is however that not all members of the team are affected equally or simultaneously. Isolating clinicians from their colleagues by heavy reliance on technology or equipment might therefore exacerbate the problem, and it seems likely that a more effective target for mitigating the effects of perceptual error would be to train teams to communicate situational awareness information explicitly.

The results of this programme of experiments show that inattention blindness is lessened with longer exposure times and where opportunities exist for individuals to reassess their situation frequently. If information can be shared freely, the team as a whole can maintain situational awareness, and this is consistent with data presented in some of the few industrial and military investigations to consider the impact of perceptual errors on team-tasks. Even the presence of a single assistant may be enough to greatly improve noticing rates.

Adopting the military model of snap-briefing followed by regular situation reports may be an effective technique, and would require relatively little expenditure to achieve. The assessment of critically ill patients requires a sequential assessment characterised by a primary survey designed to detect immediately life- or limb-threatening conditions and treat them immediately. This assessment follows a stereotyped ABCDE pattern where Airway is assessed first, followed by Breathing, Circulation, Disability, and finishing with a generalised Exposure seeking other injuries. Following the primary survey a secondary survey is conducted comprising a 'top to toe' examination

of the patient for less immediately threatening injuries, or more subtle signs of significant illness. On completion of the secondary survey there is usually a handover to other teams for definitive care.

The structure of this assessment is familiar to most doctors, and the ABCDE assessment should be universally recognised by clinicians of all grades. This is therefore an easy target for intervention, and the natural structure of the assessment lends itself well to refinement. The author suggests that a snap-brief should occur prior to commencement of the primary survey, with a 'sit-rep' at a minimum after the primary and secondary surveys. This is something that excellent teams already perform. Experienced trauma teams in busy Emergency Departments or military settings practice such techniques, but in-hospital emergency teams are often staffed by junior clinicians. These teams are much more variable in their skills, and might benefit most from this practice. Such a change would be amenable to future research using the methods described in Chapter 8.

Military experience suggests these techniques should add little time to the overall assessment (16) and might even save time overall by encouraging team efficiency both in data collection and in task management. The advantage of this model is that it will improve team-situational awareness regardless of whether perceptual error or other factors contributed to its loss, and it does not represent a revolution in resuscitation practice. Rather it represents a natural evolution of existing processes that are familiar to most practitioners. These techniques are easy to teach, particularly through high-fidelity simulation, and any effects on clinical care would be easy to measure.

An improved understanding of the role perceptual error plays in loss of situational awareness may also be of immediate assistance to staff investigating and learning from critical incidents. The author has personal experience of conducting root-cause analyses where co-investigators lack familiarity with the issues raised in this thesis, and this seems to represent a fundamental weakness in the process, as perceptual failures likely are a relevant latent risk factor in healthcare.

Perceptual errors feature little in enquiries into serious untoward incidents at present, and nothing can be done to understand their role in real-world medicine until a better appreciation of their incidence is gained. With training, it may then become easier for investigators, clinicians, and legal teams to understand just how clinicians can and do miss mission-critical information. Perceptual errors can affect anyone, and healthcare should move beyond attaching blame. No longer should we be asking ‘how on earth did they miss *that?*’ and instead recognising ‘I could have missed that too’. No matter how well-trained or experienced they are, clinicians must still function within the processing limits of the human brain.

9.4 Future opportunities for research

Psychologists will have interest in further exploration of the ‘reel-to-real’ effect observed in the final experiment. Many existing psychological experiments have used video depictions of naturalistic scenarios, and it would be relatively straightforward to revisit these experiments using live versions of the same stimuli.

It is essential to understand better why such a significant performance-gap exists between the video-based study and the simulation exercise, and to consider what might this mean for future studies. One can firstly conclude high ecological validity is vital to generating conclusions that are applicable to real-world situations, as the clinician ‘at the sharp end’ will primarily be concerned about the practical implications of these data.

Discovery of this effect must also prompt a reevaluation of some existing studies, including some by the author (17). The work of Drew *et al* represents a sizeable portion of the perceptual error work that has been conducted amongst clinicians (18-20). Much of Drew’s work is predicated upon image-change detection amongst radiologists. This is a convenient group to study; image-interpretation is undoubtedly an ecologically valid activity, and the nature of the work lends itself to straightforward adaptation of existing experimental paradigms such as flicker tasks. One could suggest however that the stimulus against which the volunteers were tested does sometimes lack validity (18) and

therefore some of Drew's conclusions about the incidence of perceptual error amongst trained medical personnel should be treated with a degree of caution.

Given that the generalisability of experimental data is of utmost interest to clinical staff, if large discrepancies exist between laboratory- and real-world experiments not only will time and effort be wasted inappropriately adapting training materials, but also there is the risk of alienating the target audience with 'false alarms'. Laboratory experiments are essential for determining the mechanistic explanations for these perceptual effects, but there is a vital and urgent role for investigations using more realistic stimuli. Laboratory tasks with their option to tightly control stimulus presentation and therefore isolate cognitive processes cannot be dispensed with, but one must consider the objectives of a particular study. If the intention is to draw real-world conclusions with any certainty, then it seems obvious that the tasks under study must be designed to replicate naturalistic interactions as far as is practical. This will of course come at the cost of lessening the experimenter's control over the data, thus the two approaches are highly complementary.

Finally, one must consider what can be done about perceptual error themselves. These effects represent a common end-point arising because of combinations of different cognitive and situational mechanisms. Inattention blindness can for example be classified as central (stemming from excessive cognitive demands limiting perceptual processing), peripheral (caused by attentional misdirection), or irrelevance-based (an attended feature is processed but discarded due to apparent non-relevance to the primary task). Change blindness similarly can arise from different stages of cognitive processing, and memory lapses may also play a role in apparent change-detection failures. It seems unlikely therefore that a single simple solution will exist to negate these effects.

Once the real-world effects of perceptual errors have been strictly defined, the exact means by which teams adapt to this new understanding can be subjected research. There are also useful implications for staffing of emergency teams to optimise workload. Understaffed teams will be at

higher workload, and therefore more likely to miss critical events, while overstaffing wastes resources.

The interaction between workload and experience remains something of an open question. Completing the final simulation experiment with a working eye-tracking system and recruits from a different experience-level would yield important insights into how clinicians seek information in emergency situations, and could further inform future training and staffing guidelines. It may be that the time when standards can be set on resuscitation team composition, and investigations should be urgently conducted to understand the most robust skill mix appropriate to predictable emergency situations.

Given that it is improbable that individuals can be 'immunised' against perceptual error, other solutions should be considered. Technical approaches such as head-mounted displays have been trialled in industrial and healthcare settings. These have appeal; it seems intuitive that presenting crucial data continuously within the practitioner's field of view should improve situational awareness. Evidence of benefit is however lacking, and some data even suggest these devices actually increase the risk of inattention blindness. Another issue with highly technical solutions is that they tend necessarily to be devised for individuals, rather than the teams that medical professionals must operate within. Devices, particularly those with poorly designed or integrated interfaces, then might tend to draw individuals away from their colleagues and impede intra-group communication. Human-factors intervention at the team-level, around for example communication, is likely to be more efficacious, but also quicker and cheaper to develop and investigate.

Resuscitation is apt for this type of study. Emergency responses are highly-protocolised, and the scenarios are sufficiently commonplace that they are suitable for examination in large numbers. Testing the effectiveness of inserting 'sit-reps' or 'snap briefs' into resuscitation management would be straightforward, and if effectiveness were demonstrated the existing resuscitation training infrastructure operated by groups such as the Resuscitation Council would offer a conduit to deliver

this easily and in volume. Furthermore, national bodies such as the Resuscitation Council are well-placed to make recommendations on team composition, and would have an interest in future research on the necessary skills mix at cardiac arrest situations.

9.5 Final Comments

The implications of these data in this thesis are wide reaching, of broad interest, cutting across both clinical medicine and experimental psychology. Non-technical skills contribute to many adverse events in healthcare, but they are not universally taught or understood by clinicians. Perceptual errors are even less well-recognised amongst healthcare personnel, but there is solid evidence to conclude that these effects do impair situational awareness in medicine. These effects appear universal, and while the long-term acquisition of expertise reduces vulnerability it is not clear that educational packages delivered over short periods can greatly reduce risk. Training interventions must therefore be delivered aimed at improving overall team performance, accepting that individual components of the team will from time to time lose situational awareness during the performance of their duties.

9.6 References

1. Reason J. Human error: models and management. *British Medical Journal*. 2000;320(7237):768-70.
2. Harmer M. Independent Review on the care given to Mrs Elaine Bromiley on 29 March 2005 London: Clinical Human Factors Group; 2005 [Available from: <http://www.chfg.org/articles-films-guides/elaine-bromiley-report>].
3. Endsley M, Garland D. Theoretical underpinnings of situation awareness: a critical review. *Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1-24.
4. Simons DJ, Chabris CF. Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception*. 1999;28(9):1059-74.
5. Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*. 1998;5(4):644-9.
6. Levin DT, Momen N, Drivdahl S. Change blindness blindness: the metacognitive error of overestimating change-detection ability. *Visual Cognition*. 2000;7(3):397-412.
7. Risko EF, Laidlaw KEW, Freeth M, Foulsham T, Kingstone A. Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*. 2012;6:143.
8. Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB. Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*. 2012;19(5):847-57.
9. Byrne MD, Bovair S. A working memory model of a common procedural error. *Cognitive Science*. 1997;21(1):31-61.
10. Flin R, Maran N, Glavin R, Patey R. *Framework for Observing and Rating Anaesthetists' Non-Technical Skills*. Aberdeen: University of Aberdeen; 2003.
11. Rutherford JS, Flin R, Irwin A, McFadyen AK. Evaluation of the prototype Anaesthetic Non-technical Skills for Anaesthetic Practitioners (ANTS-AP) system: a behavioural rating system to assess the non-technical skills used by staff assisting the anaesthetist. *Anaesthesia*. 2015;70(8):907-14.
12. Crossley J, Marriott J, Purdie H, Beard JD. Prospective observational study to evaluate NOTSS (Non-Technical Skills for Surgeons) for assessing trainees' non-technical performance in the operating theatre. *Br J Surg*. 2011;98(7):1010-20.
13. Mitchell L, Flin R, Yule S, Mitchell J, Coutts K, Youngson G. Evaluation of the Scrub Practitioners' List of Intraoperative Non-Technical Skills (SPLINTS) system. *Int J Nurs Stud*. 2011.
14. St John M, Smallman HS. Staying Up to Speed: Four Design Principles for Maintaining and Recovering Situation Awareness. *Journal of Cognitive Engineering and Decision Making*. 2008;2(2):118-39.
15. Levin DT, Simons DJ, Angelone BL, Chabris CF. Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology*. 2002;93(Pt 3):289-302.
16. Arul GS, Pugh HEJ, Mercer SJ, Midwinter MJ. Human factors in decision making in major trauma in Camp Bastion, Afghanistan. *Annals of the Royal College of Surgeons of England*. 2015;97(4):262-8.
17. Greig P, Higham H, Nobre A. Failure to perceive clinical events: an under-recognised source of error. *Resuscitation*. 2014;85:952-6.
18. Drew T, Võ ML-H, Wolfe JM. The Invisible Gorilla Strikes Again: Sustained Inattention Blindness in Expert Observers. *Psychological Science*. 2013;24(9):1848-53.
19. Drew T. Searching for Feature-Based Surround Suppression in Inattention Blindness. *Journal of Vision*. 2015;15(12):443.
20. Drew T, Evans K, Vo ML, Jacobson FL, Wolfe JM. What can you see in a single glance and how might this guide visual search in medical images? *Radiographics*. 2013;33:263-74.

10 Appendices

10.1 Conflict of interests

There are no conflicts of interest to declare in relation to the work presented here. The author has no commercial relationships that have an interest in the findings of this work, nor does he have any consultancy activity relevant to these data.

10.2 Materials & Methods

10.2.1 Literature search strategy for chapter 1

To generate the overview presented in chapter 1 a search was performed in January 2013. Medline (including in-process and non-indexed citations) and Embase were searched using the terms noted in Table 10-1, for literature published up to (and including) December 2012. A manual search of the references provided in relevant paper was also performed. The library of human-factors literature collected by the University of Oxford Centre for Simulation, Teaching, and Research was also consulted, as well as papers recommended by human factors trainers within the department.

Search Term	
1	human factors.ti,ab
2	non-technical.ti,ab
3	nontechnical.ti,ab
4	healthcare.ti,ab
5	clinical.ti,ab
6	emergency.ti,ab
7	critical.ti,ab
8	anaesthetic.ti,ab
9	anesthetic.ti,ab
10	operating.ti,ab
11	surgery.ti,ab
12	medicine.ti,ab
13	5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13
14	1 OR 2 OR 3
15	13 AND 14

Table 10-1: Chapter 1 search terms

10.2.2 Literature search strategy for chapter 2

To generate the overview presented in chapter 2 a search was performed in in January 2013. Medline (including in-process and non-indexed citations), Embase and PsychInfo were searched using the terms noted in Table 10-2, for literature published up to (and including) December 2012. A manual search of the references provided in relevant paper was also performed, and papers, textbooks, and presentations recommended by Professor Nobre were also consulted.

Search Term	
1	Perceptual error.ti,ab
2	Inattentional blindness.ti,ab
3	Change blindness.ti,ab
4	1 OR 2 OR 3

Table 10-2: Chapter 2 search terms

10.2.3 Materials used for capturing eye-tracking data

All experiments generating eye-tracking data used the Dikabilis Essential wireless mobile tracker (Ergoneers GmbH, Manching, Germany). The system as used required participants to wear a headpiece weighing approximately 80g linked to a transmitter pack worn at the waist.

The system sampled eye data at 50Hz, and transmitted to a receiver station in an adjacent room. Participants were free to move their head as they wished, and there were no limitations in the tasks they could perform whilst wearing the device.

Eye-tracking data was recorded and processed using D-Lab version 3.1 (Ergoneers GmbH, Manching, Germany). Where relevant, 'Region of Interest' (RoI) markers could be placed around predetermined items (Figure 10-1) to enable automated analysis of fixations within each specific region.



Figure 10-1: Image of a patient monitor configured for the simulation experiment described in chapter 8. Note the 'Region of Interest' markers affixed to either side of the screen to enable the system to track target regions.

10.2.4 Materials used for statistical analyses

Unless otherwise specified, all statistical analysis was conducted using SPSS (IBM, Portsmouth, UK) version 21.

10.2.5 Materials used in generating replica displays described in chapter 3

The design of the clinical displays used in the tasks was informed by a pre-study survey that was conducted across theatres and critical care areas of Oxford University Hospitals NHS Trust in mid 2013. A full description of this survey is presented in Appendix 10.3.

This survey collected information about the provision of monitoring equipment in theatres and intensive care areas across the Trust. Each monitor was assessed for its default layout, screen size, the manufacturer's details, and the specific model of each device.

These data enabled recreations to be created that matched the colour scheme and screen positions of the target monitoring parameters. Six monitors were recreations of GE/Datex displays and five replicated Philips displays. Monitors were drawn from a mixture of venues, including anaesthetic rooms, operating theatres, recovery bed spaces, and intensive care.

Replica displays were designed to match the real devices in colour scheme and layout, thus could contain any or all of the following values: ECG, HR, NIBP, SpO₂, fractional inspired carbon dioxide concentration (F_iCO₂), E_tCO₂, fractional inspired anaesthetic agent (F_iAA), E_tAA, respiratory rate (RR), fractional inspired oxygen concentration (F_iO₂), end-tidal oxygen concentration (E_tO₂), fractional inspired nitrous oxide concentration (F_iN₂O), end-tidal nitrous oxide concentration (E_tN₂O), temperature, flow-volume loops, trend displays, airway pressures, spirometry, central venous pressure (CVP), or invasive arterial pressure (ABP) as appropriate to the monitor.

Only ECG, NIBP, SpO₂, E_tCO₂, E_tAA, and RR were relevant to the experimental tasks. These fields were populated with values generated by a random-number generator, within pre-set ranges appropriate to a healthy anaesthetised patient (Table 10-3).

Value ranges used to populate monitor parameters		
Parameter		Value Range
ECG		60-100
NIBP	<i>Systolic</i>	<i>80-110</i>
	<i>Diastolic</i>	<i>50-80</i>
SpO ₂		95-100
E _t CO ₂		3.5-6
E _t AA	<i>MAC¹</i>	<i>0.7-1.5</i>
	<i>Desflurane</i>	<i>4.5-9.5</i>
	<i>Isoflurane</i>	<i>0.8-1.7</i>
	<i>Sevoflurane</i>	<i>1.4-3.0</i>
RR		8-16

Table 10-3: Value ranges for each task-relevant parameter

¹For MAC values less than 1, the monitor displays E_tN₂O 0.50

[Table 10-4: Values used to populate monitor displays](#)

Any non-task-relevant parameters that were required to recreate a display were fixed, again to values consistent with a healthy anaesthetised patient. In some instances, clinical values could be available from different sources.

While layouts one to 11 are faithful replicas of displays found in clinical use, layout 12 is entirely fictitious. This display used a colour scheme not found clinically and, other than using the basic

iconography and typefaces found on Philips machines, shared no common features with real displays.

All monitor recreations were presented on a monitor of 17", with an aspect ratio of 4:3 and a screen resolution of 1920x1080.

10.2.6 Materials used to develop the video sequence described in chapter 6

The video was scripted by the author, and recorded in the Oxford Centre for Simulation, Training, and Research (OxSTaR) with the assistance of centre faculty¹¹.

The video incorporated several change blindness probes and a single inattentional blindness probe. After 12 seconds has elapsed the team pause to evaluate the patient's rhythm and the camera cuts to the defibrillator. During this cut several instantaneous changes occur – the provider delivering CPR is swapped with a bystander, the patient is instantaneously intubated, two team members swap hats, and one team member acquires a stethoscope around their neck. These changes are illustrated in still images taken from the sequence (Figure 10-2).



Figure 10-2: Still images taken from the experimental video sequence. The two images represent before (left) and after (right) the cut to the defibrillator. Four change events have occurred: the CPR provider has switched with the individual in the background, the patient has been intubated, the two female team members have swapped hats, and the airway provider has acquired a stethoscope.

¹¹ For a list of the individuals who assisted filming of the video, please see acknowledgements at the end of this work

After CPR resumes, a shock is delivered and a dose of adrenaline given. During administration of the adrenaline (40s elapsed) the oxygen supply becomes disconnected from the wall. The release of the oxygen tubing is both visible and associated with a distinct change in the video soundtrack (Figure 10-3 & Figure 10-4). The spectrographs demonstrate two main contributors to background noise during the scene. Firstly, the clinical monitor alarm sounds continuously throughout the clip and secondly there is a constant high-pitched noise generated by gas flowing from the oxygen supply. After disconnection, the pitch of this flow rises, with a greater proportion of higher-frequency sounds. This was confirmed using the software utility Praat v6.0.19 (Boersma & Weenink) (1), which demonstrated that the central frequency of the noises associated with the gas flow rose from 3108.6Hz, pre-disconnection to 5406.3Hz post-disconnection. These changes persist to the end of the sequence. Higher frequency noises are generally perceived as louder than equivalent lower frequency sounds, so this should be clearly audible and would be expected to draw a response (2).

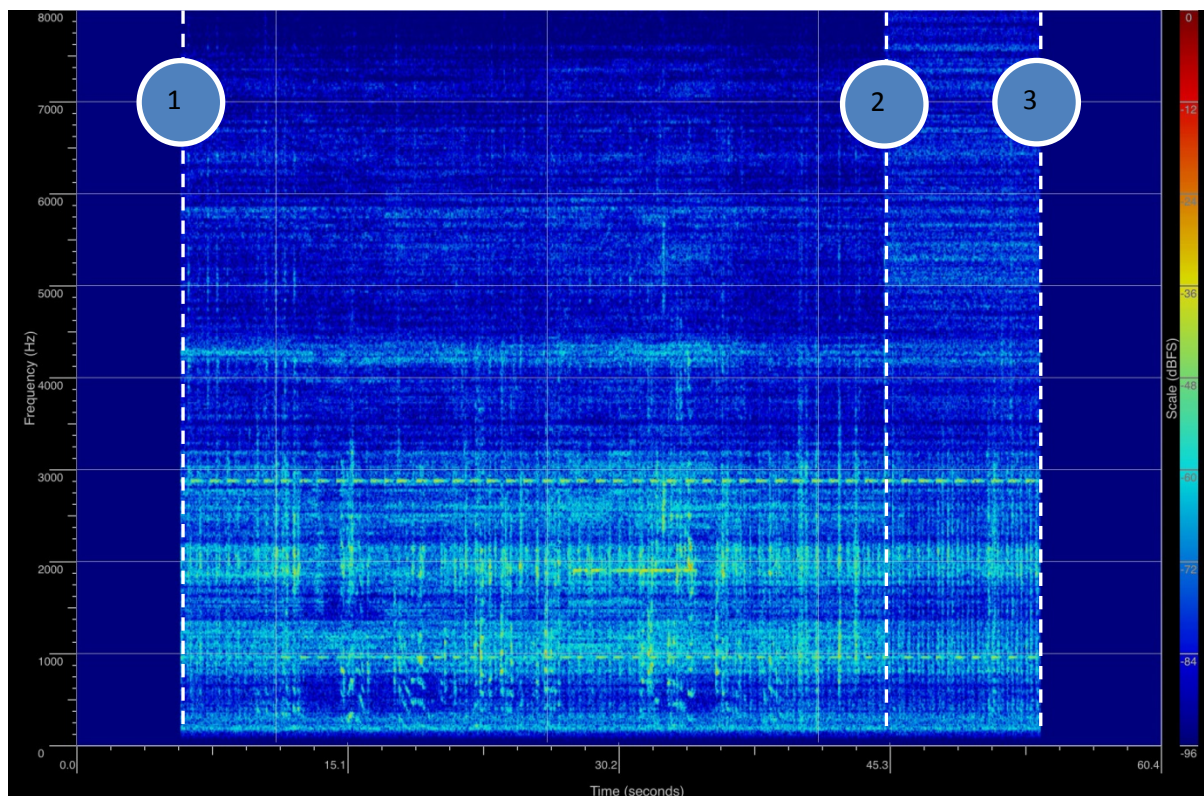


Figure 10-3: Spectrograph recorded during video playback. Frequency is displayed along the y-axis and time along the x-axis. Volume is denoted by colour. Playback begins at (1), with the oxygen disconnection occurring at (2). Playback ends at (3). The dotted line around 2900Hz probably represents recurring alarm tones arising from the patient monitor.

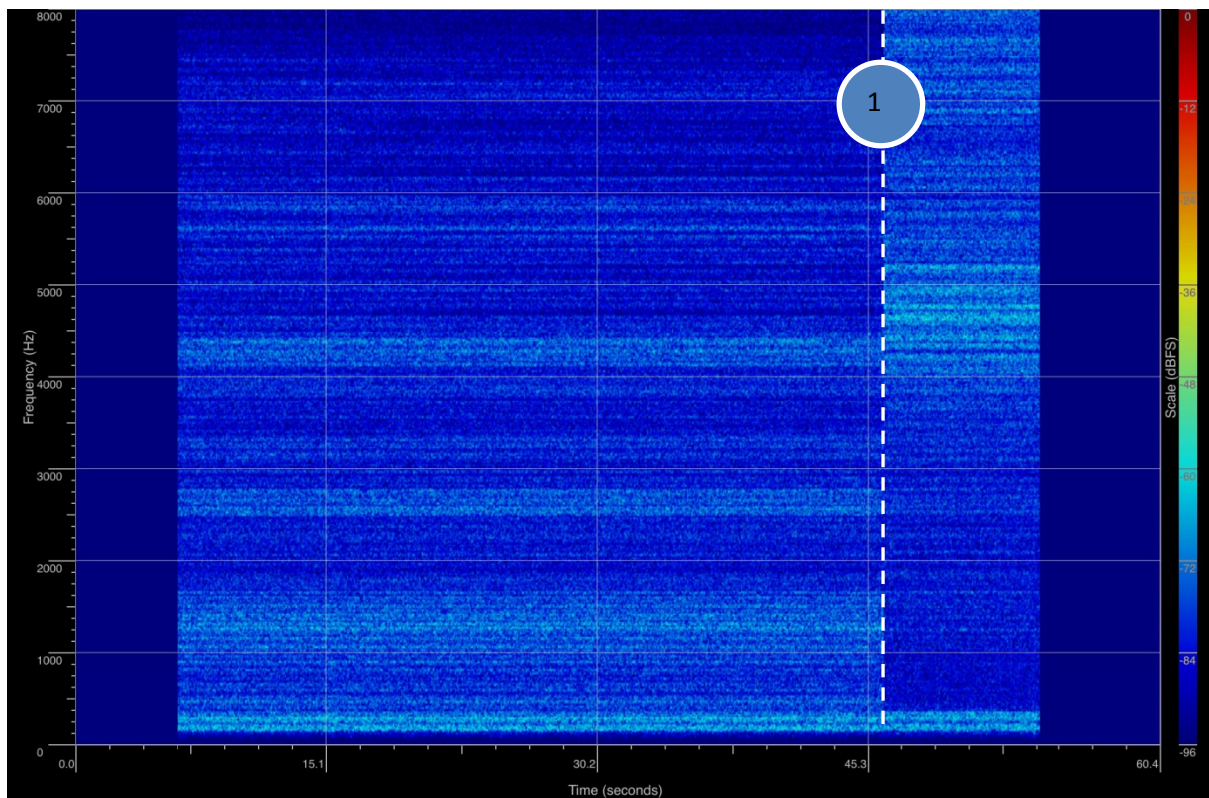


Figure 10-4: Spectrogram recorded in the simulation room no actors present and the alarm system silenced; the frequencies depicted here therefore represent only those generate by the oxygen flow and fixed background noises. The oxygen disconnection occurs at (1).

10.2.6.1 References

1. Boersma P, Weenink D. Praat: doing phonetics by computer 2016 [Available from: <http://www.fon.hum.uva.nl/praat/>].
2. International Organization for Standardization. Acoustics: normal equal-loudness-level contours. Switzerland: ISO; 2003.

10.2.7 Materials used in the online survey described in Chapter 6

On the following pages you will see materials extracted from the online survey used in chapter 6. The data were presented slightly differently when displayed on the website, but this represents a reasonable facsimile of the layout experienced by survey participants. All participants were asked to leave an email address so that they could receive a copy of their consent form, and all participants were offered an opportunity to withdraw from the study at the end of the questionnaire and have their data destroyed.

CPR Evaluation

Welcome

Thank you for considering participating in this research project.

We would like to investigate how well people can concentrate when performing different tasks, and your observations and comments could be very helpful.

If you would like to continue, please click next. On the following page we will give you details about the study, and ask for your consent to take part.

Introduction

You are invited to take part in a voluntary research study. Before you decide if you wish to take part, it is important to know what this might mean for you.

Please read the following information carefully, which explains what you would be asked to do should you wish to take part.

What is the purpose of this research?

This study intends to investigate how well people can concentrate while performing different tasks.

Do I have to take part?

You do not have to take part. You are quite free to decline to participate, and no reasons need to be specified.

What will happen if I do take part?

If you would like to take part, you will be shown a video sequence and then asked to offer your observations on it. Participation shouldn't take longer than about 10 minutes. Unless you ask us to, no member of the study team will ever contact you, and we do not keep a record of your identity or email address.

Expenses and payments

You will not be paid to take part in this study, and no expenses should be expected.

What are the disadvantages of taking part?

Participation will require you to spend a few minutes watching a video whilst concentrating on a specific task. Other than requiring up to 10 minutes of your time, there should be no risk to taking part.

What are the benefits of taking part?

You will not benefit personally from taking part, but it is hoped that your observations and comments will help us to understand how well people concentrate when presented with different tasks. This may help us design training for healthcare professionals in the future.

What will happen to data collected about me?

No identifiable data is collected about you during this study. All data is anonymised, and held on a secure system that complies with the Data Protection Act. What data that is collected is strictly confidential, and accessible only to study staff.

Who is organising and paying for the research?

The study is being organised and funded by the Nuffield Division of Anaesthetics, a part of the University of Oxford.

Who has reviewed the study?

The study has been reviewed by the University of Oxford Central University Research Ethics Committee. They have assigned the following number to the study: MSD-IDREC-C1-2012-139.

How can I get more information?

If you would like more information or if you would like a copy of this information, please email oxstar.research@ndcn.ox.ac.uk.

* 1. Please click the box to indicate you have read the information and are happy to proceed.

- I have read the required information, and would like to proceed
- I do not want to take part at this time

CPR Evaluation

* 2. What is your age?

* 3. What is your gender?

Female

Male

* 4. How would you describe your professional background?

Medical Doctor

Nurse

Resuscitation Officer

Operating Department Practitioner

Medical Student

Other (please specify)

* 5. Which specialty area(s) do you work in? (please click all that apply)

Anaesthetics

Intensive Care

Acute Medicine

Emergency Medicine

General Surgery

Primary Care

Student

Other (please specify)

* 6. Which of the following best describes your level of training in adult resuscitation? (please click all that apply)

- I am a basic life support provider
- I am an intermediate life support provider
- I am an advanced life support provider
- I am an advanced life support instructor
- None of the above
- Other (please specify)

* 7. Approximately how many resuscitations do you attend every year?

Resuscitations per year:

CPR Evaluation

Video Briefing

On the next page you will find a video sequence depicting a simulated cardiac arrest. Please watch the video all the way to the end.

While watching, we'd like you to imagine yourself in the role of the person leading this team - please watch to make sure the CPR is being performed correctly, and if the team deliver a shock that it is given safely.

Consider the instructions that you would give the team, and the drugs you would administer, if you were a part of this emergency.

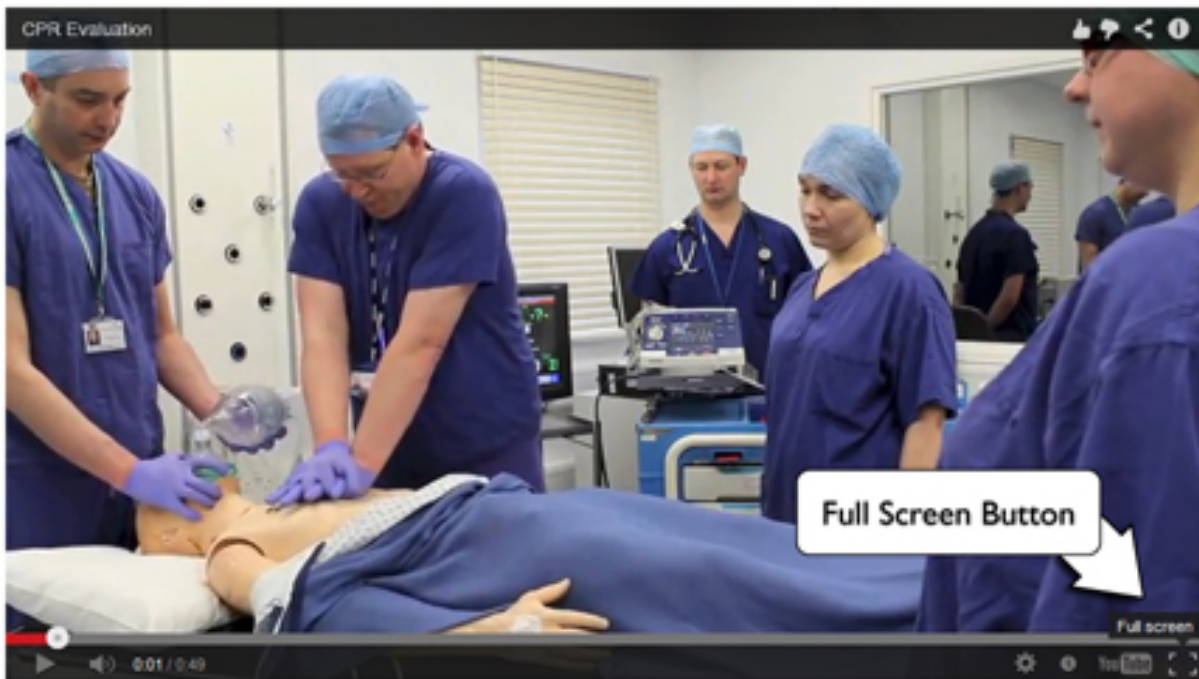
We will be asking you some questions about how well the team performed after the video has finished.

To enable you to listen to the dialogue, you will need to have your speakers or headphones connected, and at a comfortable volume level.

Click next to view the video.

CPR Evaluation

Video Sequence



* 10. Did the video play correctly?

- Yes
- No

CPR Evaluation

Comments on CPR

Please now think about simulated emergency that you saw. We would like to know about your observations on the scene.

* 11. Please make any comments that you wish about the team's performance.

12. Did you notice any unusual or unexpected events?

- Yes
- No

CPR Evaluation

Comments on events

* 13. Please describe the unusual or unexpected event(s) that you noticed.

CPR Evaluation

Comments on events

We would now like to ask some specific questions about events that may, or may not, have occurred during the sequence.

14. Did you notice any of the following events occurring? (Please tick any that apply)

- A change in the number of team members
- The airway person changed
- The CPR person changed
- The cardiac rhythm changed
- The patient's airway changed
- The defibrillator malfunctioned
- The oxygen malfunctioned
- The team at the end were wearing different coloured tops
- The team at the end were wearing different coloured hats
- The airway person had acquired some new equipment by the end of the scene
- The team leader had acquired some new equipment by the end of the scene

Please add any comments that you wish.

CPR Evaluation

Debriefing

Thank you for watching the video and offering your observations. The purpose of this research has been to determine what people look at when they are involved with medical emergencies, and we are interested in how aware individuals are of their environment.

What was this really about?

Some studies conducted by experimental psychologists have found that it is common for people who are required to concentrate on a task to miss important events going on around them, and we have designed this sequence to explore where this might happen in medical practice.

The original experiments were conducted in the late 1990's, and were conducted on undergraduate students, who were asked to watch a video of people passing basketballs¹. If you would like to see this video yourself, you can find it [here](#). Around half of the people who watched this video didn't see the event the experimenters had included.

The same team conducted a follow-up study in 2010² with a similar video. You can view this [here](#). This time they included more than one change, and they found that even when people knew that something was going to happen, it didn't automatically mean that they would notice all of the other events.

No-one has yet looked in to how this might happen to healthcare professionals when they are looking after critically ill patients, so we have devised the video you have just seen using the techniques established by these psychological experiments.

What happened in this video?

In the video you have just seen, there were actually five changes. The first four happened during the cut to the defibrillator screen. The changes included:

1. the person doing CPR and the person standing behind the defibrillator instantly switched places
2. the person managing the airway suddenly had a stethoscope appear around their neck
3. the team leader and the person operating the defibrillator swapped hats
4. the patient was instantly intubated
5. the oxygen supply became detached from the wall while the adrenaline was being given

Before



After



What does this mean?

Not noticing these changes surprises a lot of people - most people would expect to see major changes in a video sequence, but in fact only around half (and sometimes even fewer) actually do see everything.

Whether you notice the changes or not does not seem to be related to intelligence, vision, attention or your ability to concentrate. As far as we can tell, this isn't something that people can control. If we were to show you more videos like these, in some you would notice the changes and in others you probably would not: we are all vulnerable to these omissions, and it is not the case that people who notice the changes will always notice them, and neither to people who missed some changes always miss them. No-one is immune to this phenomenon.

We think that these "perceptual errors" might happen when healthcare professionals are caring for critically ill people, and we would like to explore if anything about how people are trained, or the experience they have, can be used to minimise their effects.

What next?

Thank you for taking part in this experiment. Your anonymised observations and comments will help us to determine how common these perceptual errors actually are.

If you have found it interesting, please forward the link to friends or colleagues - we would like as many people as possible to take part. If you do, please *don't tell them what you have seen, or what to expect*: this will give them the best possible experience.

References

1. Simons DJ, Chabris CF, Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception* 1999; **28**: 1059-74
2. Simons DJ, Monkeying around with the gorillas in our midst: familiarity with an inattention-blindness task does not improve the detection of unexpected events. *i-Perception* 2010; **1**: 3-6

If you have any questions about the experiment, please get in touch with us at oxstar.research@ndcn.ox.ac.uk

CPR Evaluation

And finally...

* 15. Have you heard of Dan Simons' "Invisible Gorilla" experiment before?

Yes

No

* 16. If you missed some of the changes, are you surprised by the things you didn't see?

Totally unsurprised Somewhat unsurprised Neither surprised nor unsurprised Somewhat surprised

Totally surprised I didn't miss any of the changes

10.2.8 Materials used to collect online data for the experiment in Chapter 7

On the following pages you will see materials extracted from the online survey used in chapter 7. The data were presented slightly differently when displayed on the website, but this represents a reasonable facsimile of the layout experienced by survey participants. All participants were asked to leave an email address so that they could receive a copy of their consent form, and all participants were offered an opportunity to withdraw from the study at the end of the questionnaire and have their data destroyed.



NUFFIELD DEPARTMENT OF
CLINICAL NEUROSCIENCES

Understanding Perception during Resuscitation Situations

Welcome

Thank you for considering participating in this research project.

We would like to learn more about the sorts of events that people might expect to notice when they are busy, and your observations and comments could be very helpful.

If you would like to continue, please click next. On the following page we will give you details about the study, and ask for your consent to take part.

Introduction

You are invited to participate in a voluntary research study. Before you decide if you wish to take part, it is important for you to understand what taking part would mean for you.

Take some time to read this information carefully which explains what you would be asked to do, should you choose to take part.

Please feel free to discuss participation with a member of the team. If you'd like more information, or if you have any questions, you can contact the study's organiser at the address above. Take your time to decide if you want to take part in this research.

What is the purpose of the study?

This study aims to explore healthcare-trained volunteers' understanding of perception, specifically in a resuscitation scenario.

Why have I been invited to take part?

We wish to know more about the sorts of things that people of differing levels of experience might expect to reliably notice when involved in an emergency situation. We are approaching volunteers drawn from undergraduates and qualified staff from medical, nursing, and allied professional backgrounds, who are currently practising in the UK. We believe that you are from one of these groups.

Do you have to take part?

You do not have to take part. You are quite free to decline to participate, and no reasons need be specified. You can also withdraw at any time without penalty should you decide to take part, but later change your mind. Again, no reason need be given.

What will happen to you if you take part?

You will be asked to confirm your consent online below and, at your request, you will be sent a countersigned pdf of this consent document.

You will then be asked to complete this questionnaire online. In the questionnaire you will be asked to imagine yourself in either an emergency situation, like a cardiac arrest, or a more routine scenario. You will be presented with a description of a series of events and asked whether you think you would notice them as they happened. There are no tests, and no knowledge or direct experience of the situation is required. There are no 'right' answers to any of the questions – we only want to know what you would expect to notice.

The study should take no longer than 10 minutes of your time.

Expenses and payments

You will not be paid to take part in this study, and no expenses are expected.

What are the disadvantages of taking part?

Participation will require you to spend a few minutes completing an online questionnaire. Other than requiring some of your time, there should be no risk to taking part.

What are the possible benefits of taking part?

You will not receive any personal benefit from taking part in this study, but it is hoped that this research will help us to understand what people expect of their ability to notice events during emergency situations. We think that this could help us run training courses to make teams safer, more effective, and more efficient.

Will your taking part in this study be kept confidential?

Any information collected about you during the course of this study will be kept strictly confidential, treated anonymously, and used only for research purposes or for routine regulatory and audit purposes.

What will happen to data collected about me?

Study staff will ensure that your anonymity is maintained in the trial records. All documents will be stored securely and only accessible by study staff and authorised personnel. Only a participant ID number will identify you in study documents. The study will comply with the Data Protection Act, which requires data to be anonymised as soon as it is practical to do so. Your IP address is not stored, and no identifiable data about you will be stored. Anonymised data from participants will be retained for up to three years after completion of the study.

Who is organising and paying for the research?

This study is funded through the Nuffield Department of Clinical Neurosciences, University of Oxford.

Where is the research being carried out?

The research is being carried out in the Oxford Centre for Simulation, Training, and Research (OxSTaR). This is a part of the Nuffield Department of Clinical Neurosciences, and is based at the John Radcliffe Hospital, Oxford.

Who has reviewed the study?

This study has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee (Ref: **MSD-IDREC-C1-2014-088**).

What if there is a problem?

If you have a concern about any aspect of this project, please contact the lead researcher (Dr Paul Greig via email at: oxstar_research@ndcn.ox.ac.uk or phone: 01865 234528) who will do his best to answer your query. Your concern will be acknowledged within 10 working days and give you an indication of how he intends to deal with it.

If you remain unhappy and wish to make a formal complaint, please contact the chair of the Research Ethics Committee at the University of Oxford (Chair, Medical Sciences Inter-Divisional Research Ethics Committee; email: ethics@medsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford, OX1 2JD). The chair will seek to resolve the matter in a reasonably expeditious manner.

How can I get more information?

If you would like more information about this study please contact Dr Paul Greig via email at: oxstar_research@ndcn.ox.ac.uk or telephone: 01865 234528)

Thank you for reading this information and for considering participation in the study. If you wish to take part, please complete the consent form below.

By clicking the box below you confirm that:

- I have read and understood the information presented above
 - I have had an opportunity to contact the study team to ask questions, and all of my questions (if any) were answered satisfactorily
 - I voluntarily agree to be part of this research study, and to provide necessary information to study personnel as requested
 - I understand that I am free to withdraw from the study at any point, without penalty, and I do not need to give a reason
 - I allow the study team and sponsor to use my information as described in the information pages above. I am aware that only anonymised data will be stored by the University, for a period of up to three years after completion of the study
-

* 1. Please indicate if you wish to take part in the study by clicking below:

- I wish to take part in this study
- I do not wish to take part in this study



Understanding Perception during Resuscitation Situations

Demographic Data

To help us ensure we are surveying a mix of people from different professional backgrounds, please fill out the following details.

* 5. Please enter your age:

* 6. What is your gender?

Female

Male

* 7. How would you describe your professional background?

Medical doctor

Nurse

Resuscitation Officer

Operating Department Practitioner

Medical student

Other (please specify)

* 8. Which of the following best describes your training in adult resuscitation?

I have not been trained in adult resuscitation

I am a basic life support (BLS) provider

I have completed an Immediate Life Support (ILS) course recognised by the Resuscitation Council (UK)

I have completed an Advanced Life Support (ALS) course recognised by the Resuscitation Council (UK)

I instruct on Advanced Life Support (ALS) courses recognised by the Resuscitation Council (UK)

Other (please specify)



Understanding Perception during Resuscitation Situations

Briefing

Thank you. In the following pages we will describe three scenarios and a series of events.

In each case we'd like you to imagine yourself in the situation, and indicate whether you would expect to notice if each event occurred.



Understanding Perception during Resuscitation Situations

Scenario: A real-life emergency situation

Situation: We'd like you to imagine that you are involved in the management of a real-life resuscitation situation. The patient has had a VF cardiac arrest, and the team are providing CPR.

You are in charge of the arrest team, and both giving them instructions and monitoring their performance. The team expect this leadership to be only at a level appropriate to your training.

The picture below represents the sort of scene that you might be involved with.

CPR in progress



During the resuscitation the following events occur:

- One member of the team leaves the area to get results from a blood gas. When they return they are wearing a different coloured hat.
- One member of the team leaves the area to get a piece of equipment. When they return they have been switched with a completely different person.
- You look down to read the patient's notes. When you look up the patient has been intubated.
- You look down to read the patient's notes. As you do, sinus rhythm appears on the patient monitor.

* 9. How certain are you that you would notice each of these changes?

	I am certain I would notice this	I think I would notice this	I think I would not notice this	I am certain I would not notice this
The change in hats	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The switch in the team members	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The patient becoming intubated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The change in the patient's rhythm	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Understanding Perception during Resuscitation Situations

Scenario: A real-life emergency situation (continued)

The scenario continues as described.

CPR in progress



Whilst you are managing the arrest one of two events happens. Either:

1. The oxygen supply becomes disconnected
- or
2. The patient's IV line becomes disconnected

* 10. How certain are you that you would notice each of these changes?

	I am certain I would notice this	I think I would notice this	I think I would not notice this	I am certain I would not notice this
The oxygen disconnection	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The IV disconnection	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Understanding Perception during Resuscitation Situations

Scenario: Watching a video of a resuscitation

Situation: This time we'd like you to imagine that you are watching *a video* of a resuscitation situation. The scene is depicted in the image below.

While you watch the video you have been given instructions to monitor the team's performance. After the video is finished you are going to be asked questions about team's quality of CPR, the safety of defibrillation, and the drugs that were administered. This means you are watching carefully, and concentrating quite hard on how the team work.

CPR in progress



During the video the camera cuts to the defibrillator to show the rhythm. You see VF displayed, but when the camera cuts back the following items have changed:

- The two female team members have swapped hats
- A stethoscope has appeared around the airway provider's neck
- The person delivering CPR has instantaneously swapped with the person at the back
- The patient has become intubated

* 11. How certain are you that you would notice each of these changes?

	I am certain I would notice this	I think I would notice this	I think I would not notice this	I am certain I would not notice this
The change in hats	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The appearance of the stethoscope	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The CPR providers switching places	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The patient becoming intubated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Understanding Perception during Resuscitation Situations

Scenario: Watching a video of a resuscitation (continued)

The scenario continues as described.

CPR in progress



Towards the end of the video the team leader gives the patient a dose of adrenaline. While she is doing this the oxygen supply becomes disconnected from the wall.

The tubing is visible between the shoulders of the people at the head of the bed. Leaking gas flowing from a disconnected supply also sounds different to gas flowing through the system normally, so you might be able to hear a change even if you hadn't seen it.

* 12. How certain are you that you would notice the disconnection of the oxygen supply?

I am certain I would
notice this

I think I would notice this

I think I would not notice
this

I am certain I would not
notice this

The oxygen
disconnection



Understanding Perception during Resuscitation Situations

Scenario: A non-medical situation

Situation: Finally, we'd like to ask you about a situation completely unrelated to healthcare.

Please imagine you are walking in an area that you know well. You are stopped by a stranger and asked for directions to a nearby building. Whilst giving directions two people walk between you carrying a door, disrupting your view of the stranger.

During this interruption, and out of your sight, the stranger is swapped for a new individual. The new individual is the same gender, but of different height and build, and is wearing different coloured clothes.

Look at the image below. Imagine that the person on the left approached you first, and was then swapped with the person on the right.

Identity Switch



Image reproduced from Levin DT, Momen N, Drivdahl S, Visual Cognition 2000; 7: 397-412.

* 13. How likely do you think you are to notice the change in your conversation partner while you are talking to them?

I am certain I would
notice this

I think I would notice this

I think I would not notice
this

I am certain I would not
notice this

A switch in conversation
partner



Understanding Perception during Resuscitation Situations

Are you familiar with the phenomenon of perceptual errors?

* 14. Have you heard of 'perceptual errors', 'change blindness', or 'inattention blindness' before?

Yes

No

* 15. Have you heard of the 'Invisible Gorilla' before?

Yes

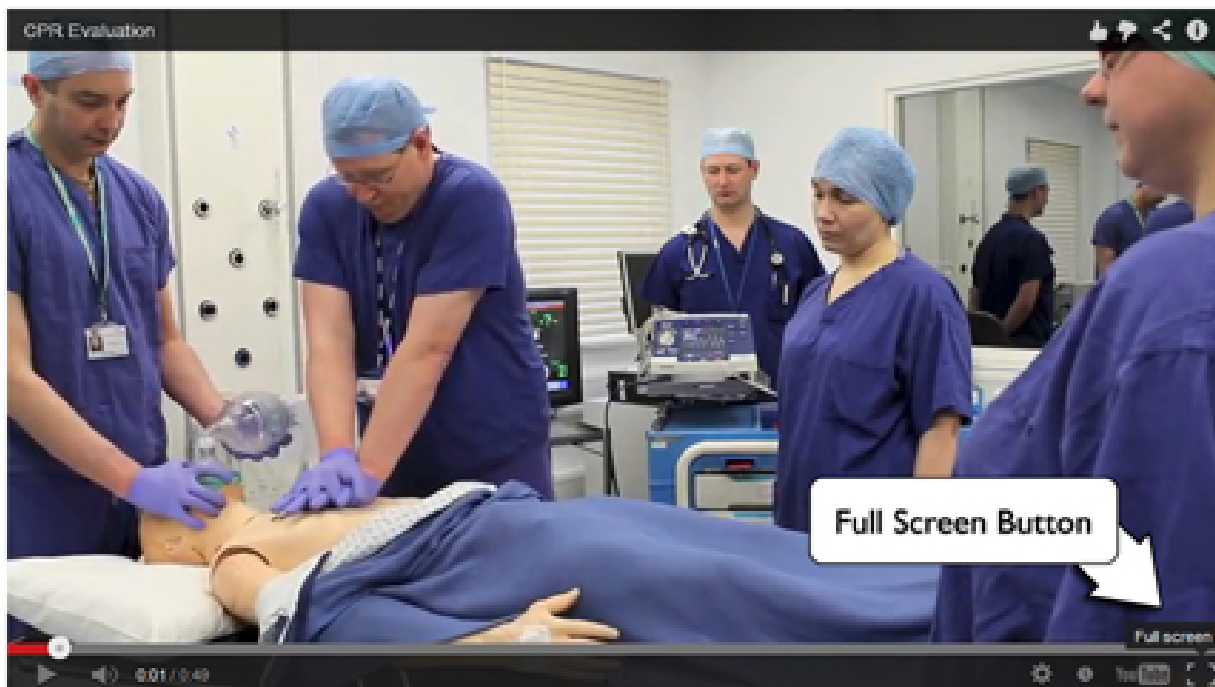
No



Understanding Perception during Resuscitation Situations

The CPR video experiment

Scenario 1 is based on an experiment that we actually conducted. Here is the original video that we used in the study.



* 16. Have you seen this video before, or have you read the paper about the experiment entitled 'Perceptual Limitations: an underrecognised source of error' published in *Resuscitation*?

- I took part in the experiment
- I have seen the video
- I have read the paper
- I have done none of the above
- I'm not sure

Other (please specify)



Understanding Perception during Resuscitation Situations

Debriefing

Thank you for taking part in this study. We'd like to spend a bit of time explaining why we are interested in what you expect to notice.

Psychologists have been investigating a phenomenon called 'perceptual error' for many years. A perceptual error has occurred when an individual misses an event that they should otherwise have been able to see, and occurs because of the way the brain processes sensory information. We are all susceptible to these types of errors – when we are placed under stress, or when we have to concentrate on a situation, the brain can't process all of the information available to it simultaneously. Some information gets 'deleted' before you become aware of it.

One famous experiment by a Harvard psychologist called Dan Simons asked people to watch a video of teams passing basketballs. Simons asked his volunteers to count the number of passes that the team in white shirts made. You can view the video [here](#). If you haven't seen the video before, you may wish to watch it now, before reading any further!

Around halfway through this video a person walks through the scene wearing a gorilla costume; half of the people who watch the video counting passes don't see the gorilla at all, whereas almost everyone sees it if they watch the video without having to count passes¹. The harder the task you have to concentrate on, the more likely you are to miss an event.

Another video by the same group can be seen [here](#). In this video the experimenter (holding a map) walks up to an unsuspecting person on a University campus and asks for directions. Part way through the conversation they are interrupted as two people (who are involved in the experiment) carry a door between them. During the interruption the person with the map is switched with someone else. Again, around half of the people who took part in the experiment didn't notice that they finish their conversation talking to a different person².

This is a really surprising result – most of us think that we would notice if a conversation partner were switched part-way through the interaction, but unless you know the person well the brain doesn't automatically record much information about them, making it difficult to compare before and after the interruption.

This last study particularly interests us, because the experimenters followed it up by asking people who weren't involved in the study to say if they would expect to see that change, much as we have done here. They found that 97% of the people that they questioned expected to notice the switch, even though only around half actually did³.

We conducted our own study similar to Simons' gorilla experiment⁴. In this experiment we showed people a video of a resuscitation situation, but we added some continuity errors and events to the sequence. This is the video that you watched on the previous page. Most people didn't notice the changes, even though we included clinicians with a lot of resuscitation experience and that some of the events would affect how the resuscitation was managed.

We are keen to see if people with healthcare training, like the people tested by the psychologists, are surprised by the sorts of events that can be missed and we'd like to compare the results of these questionnaires with the rates of noticing that we found in our experiment. We hope that the results will help us to better design training courses to make people aware of the risk that significant events can pass unnoticed.

If you'd like more information, or if you'd like to ask any questions about the work, do please get in touch via oxstar_research@ndcn.ox.ac.uk. We'd be delighted to discuss further, and we can arrange telephone contact if you'd prefer. Also, if after reading this you've decided you don't want to participate in the trial any longer, please let us know about that too.

We're really grateful that you've taken the time to help us.

Paul Greig
Sophie Robin
University of Oxford

To ensure that others who participate in the study in the future have the same experience as you, we would be really grateful if you would avoid discussing this background information with anyone before they have taken part.

References

1. Simons DJ, Chabris CF Gorillas in our midst: sustained inattentive blindness for dynamic events. *Perception* 1999; 28: 1059-74.
2. Simons DJ, Levin DT Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review* 1998; 5: 644-9.
3. Levin DT, Momen N, Drivdahl S Change blindness blindness: the metacognitive error of overestimating change-detection ability. *Visual Cognition* 2000; 7: 397-412.
4. Greig P, Higham H, Nobre A Failure to perceive clinical events: an under-recognised source of error. *Resuscitation* 2014; (In press).



NUFFIELD DEPARTMENT OF
CLINICAL NEUROSCIENCES

Understanding Perception during Resuscitation Situations

Thank you

Thank you for taking the time to consider this study.

You can now close this window.

NASA-TLX Rating Scale Definitions

Title	Endpoints	Descriptions
Mental Demand	Low/High	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand	Low/High	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand	Low/High	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Effort	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Performance	Perfect/Failure	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Frustration Level	Low/High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Table 10-5: NASA-TLX scale definitions

The resulting scores can be summated to yield a 'raw' score, although the authors of the Index recommend weighting the scales before summation.

Weighting is achieved by presenting the participant with a series of paired scale titles. They must then indicate which they felt was most important contributor to workload in the task at hand. Each time a scale is chosen it gains one point, and at the end of the weighting exercise the raw scale score is multiplied by this weighting factor.

The maximum score for the highest weighted scale is theoretically 500, if the scale was to be chosen in every paired instance. A scale could have a weighting of zero, if it were never to be chosen during the weighting process. In this case, the participant is indicating the specified scale contributed no workload to their task,

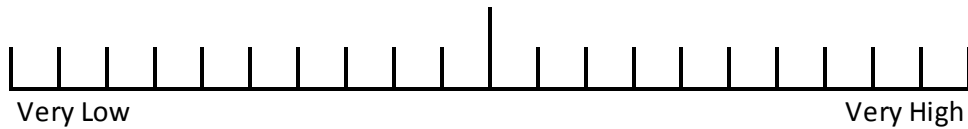
The formula that generates the final NASA-TLX score combines weighted scores to yield a single value, with a maximum value of 100.

Simulation Study: Record Form		
Scenario 1: High Workload Conditions		
PLEASE COMMENT ON THE FOLLOWING ITEMS:		
Can you list the cardiac rhythms that you noticed?		
Please could you list all the drugs and fluids that were given during the arrest?		
Did management follow RC(UK) ALS guidelines? If not, please can you describe where it differed?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Were there any communication issues with the team If so, please describe them here:	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Were there any equipment issues or critical incidents? If so, please describe	<input type="checkbox"/> Yes	<input type="checkbox"/> No
If you would like to make any additional comments about your scenario, please add them below:		

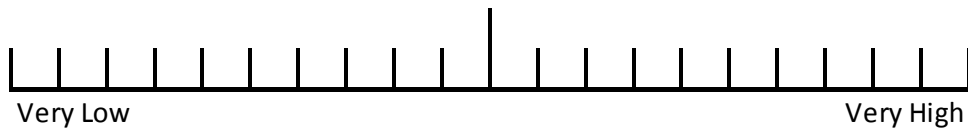
NASA Task Load Index: Scenario 1

Hart and Staveland's NASA Task Load Index (TLX) method assesses workload on five 7-point scales. Increments of high, medium, and low estimates for each point result in 21 gradations on the scales.

Mental Demand How mentally demanding was the task?



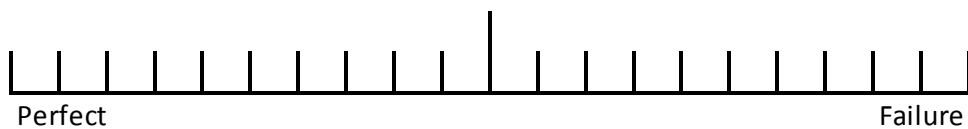
Physical Demand How physically demanding was the task?



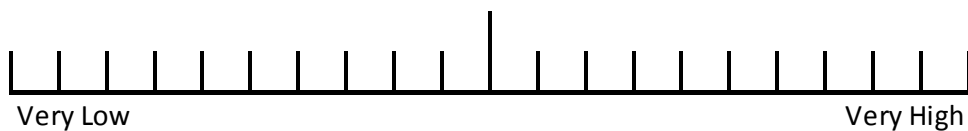
Temporal Demand How hurried or rushed was the pace of the task?



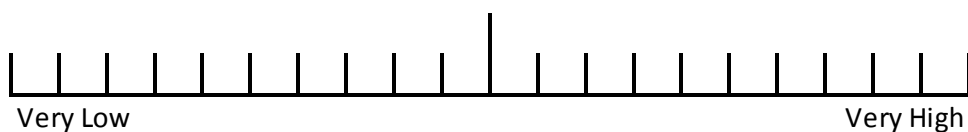
Performance How successful were you in accomplishing what you were asked to do?



Effort How hard did you have to work to accomplish your level of performance?



Frustration How insecure, discouraged, irritated, stressed or annoyed were you?



Simulation Study: Record Form

Scenario 2: Low Workload Conditions

PLEASE COMMENT ON THE FOLLOWING ITEMS:

Can you list the cardiac rhythms that you noticed?

Please could you list all the drugs and fluids that were given during the arrest?

Did management follow RC(UK) ALS guidelines?
If not, please can you describe where
it differed?

Yes

No

Were there any communication issues with the team
If so, please describe them here:

Yes

No

Were there any equipment issues or critical incidents?
If so, please describe

Yes

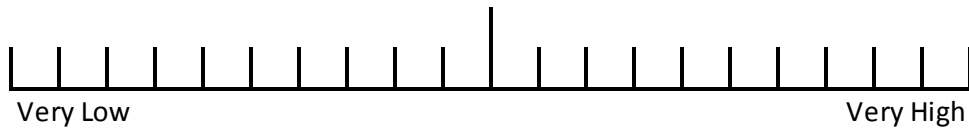
No

If you would like to make any additional comments about your scenario, please add them below:

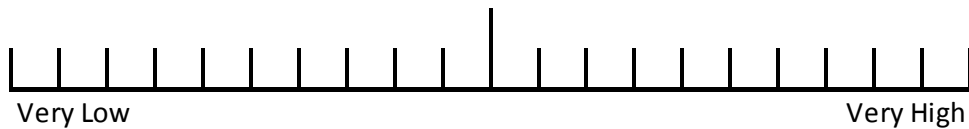
NASA Task Load Index: Scenario 2

Hart and Staveland's NASA Task Load Index (TLX) method assesses workload on five 7-point scales. Increments of high, medium, and low estimates for each point result in 21 gradations on the scales.

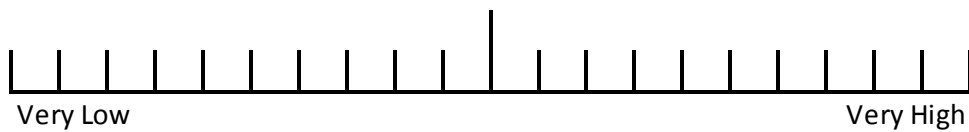
Mental Demand How mentally demanding was the task?



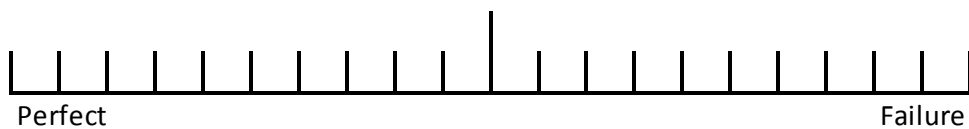
Physical Demand How physically demanding was the task?



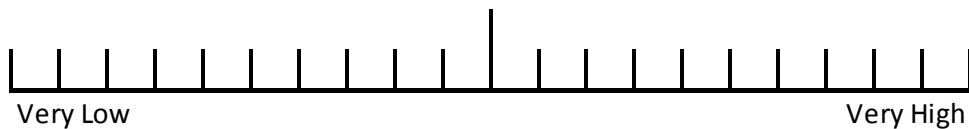
Temporal Demand How hurried or rushed was the pace of the task?



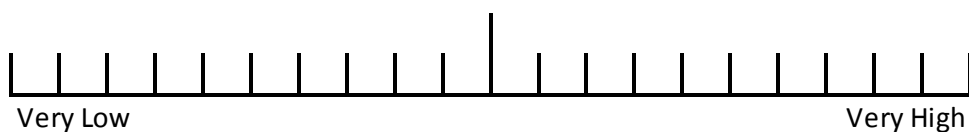
Performance How successful were you in accomplishing what you were asked to do?



Effort How hard did you have to work to accomplish your level of performance?



Frustration How insecure, discouraged, irritated, stressed or annoyed were you?



In addition to how the focus of people's attention varies between the two scenarios, we also included some extra events to see if people notice things happening around them. We'd like you to finish by looking at the list and recording if you noticed any of them, and how certain you are for each.

Not all of the events listed below will actually have happened in your scenarios.

Scenario 1	I'm certain I did not see this	I think I did not see this	I think I did see this	I'm certain I did see this	If you noticed this, could you provide details:
I had a different number of people on my arrest team at the end	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Members of my arrest team were different at the end	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Members of my arrest team were dressed differently at the end	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems with the oxygen supply	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems with the defibrillator	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems with the IV line	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems giving medications	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Items of equipment went missing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Scenario 2	I'm certain I did not see this	I think I did not see this	I think I did see this	I'm certain I did see this	If you noticed this, could you provide details:
I had a different number of people on my arrest team at the end	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Members of my arrest team were different at the end	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Members of my arrest team were dressed differently at the end	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems with the oxygen supply	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems with the defibrillator	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems with the IV line	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
There were problems giving medications	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Items of equipment went missing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

10.3 Oxford University Hospitals Anaesthetic Monitor Survey

10.3.1 Introduction

The primary intention of this data-gathering exercise was to generate an accurate library of clinical monitoring devices used across Oxford University Hospitals NHS Foundation Trust. From this library displays could then be selected for the experiment described in chapter 4, examining the role of familiarity and professional background on visual search strategies.

10.3.2 Methods

A survey was conducted across all sites of Oxford University Hospitals NHS Foundation Trust (OUH) to take a snapshot of all devices available across the Trust. The monitors provided in each anaesthetic room, operating theatre, and recovery unit were examined and photographed, as were the equipment by each general intensive care unit (ICU) bed-space. Specialist areas, such as cardiothoracic critical care or the neurosciences ICU, were not examined as part of this survey because their monitoring needs are highly specific.

The make, model, and screen size of each monitor was noted, along with an image of the default screen layout displayed after powering up the device. These images were then examined for parameter position, colour scheme, typeface, and iconography.

10.3.3 Results

Critical care areas by hospital site (excluding non-theatre interventional areas)						
Hospital		Area Type				Total
		Anaesthetic Rooms	Operating Theatres	Recovery Beds	Intensive Care Beds	
John Radcliffe	Women's Centre	4	4	3	0	11
	Main Building	10	10	11	17	48
	West Wing	14	14	16	12	56
	All sites	28	28	30	29	115
Churchill Hospital	All sites	10	10	9	8	37
Nuffield Orthopaedic	All sites	6	6	18	0	30
Horton Hospital	All sites	5	5	6	6	22

Table 10-6: Higher-care areas at each hospital site

OUH comprises four hospitals spread across central and northern Oxfordshire. Between these sites there are eight geographically distinct theatre suites and associated recovery facilities. Two theatre areas, in the delivery suites of the John Radcliffe and Horton General Hospitals, do not use a theatre-based recovery area. Instead once patients can leave theatre safely they are discharged directly to a ward area under one-to-one midwifery care. There are general intensive care units located at the John Radcliffe, Churchill Hospital, and Horton Hospital.

There are 49 operating theatres (OT) and 49 anaesthetic rooms (AR) in total, along with 63 recovery spaces and 43 ICU beds. These data are summarised in Table 10-6.

There were 92 anaesthetic machines; 45 located in AR, and 47 in theatre. Neither of the two AR in the John Radcliffe delivery-suite had machines in use. Two theatres (with their associated anaesthetic rooms) had been taken out of service for refurbishment at the time of the survey. There was an additional anaesthetic machine present in three of the six recovery areas. The machines comprised 12 different models from four different manufacturers. The most commonplace machine was the Datex Aespire, present in 34 locations.

There was similar variability in monitor provision. A monitor was attached to every AR and OT anaesthetic machine, although this was not true of all recovery area machines. There were 14 different models of monitor supplied by one of two manufacturers, either Philips or GE/Datex Healthcare. The John Radcliffe site uses a mixture of manufacturers, with 48 GE/Datex monitors and 30 Philips devices. The Churchill site uses exclusively GE/Datex equipment, while both the Nuffield Orthopaedic Centre and Horton Hospital use exclusively Philips monitors. ICU beds on all sites were monitored by devices manufactured by Philips.

There were 43 unique monitor displays recorded. Many of these monitors were variations of each other – for example differing only in whether ancillary values such as temperature were displayed or not. For analysis purposes displays were grouped based on the colour and screen position of the

fundamental parameters (ECG, NIBP, SpO₂, and airway gas measurements). The only common feature across all displays was the colour and position of the ECG values. Except where the monitor was monochrome this was always green and was always topmost on each display. There was more heterogeneity in the displays of the other values. Blood pressure could be red, purple, or green. Pulse oximetry varied between cyan, yellow, or blue. Screen positions for these parameters also varied widely.

Airway gas values were more consistent in their colouring – end-tidal carbon dioxide (E_tCO₂) values and waveforms were most frequently white, and inhalational anaesthetic measurements were coloured according to the agent detected (purple for isoflurane, yellow for sevoflurane, or blue for desflurane). These differed in location and whether a waveform was displayed alongside numerals.

In 39 cases monitors from the same manufacturer were present in both AR and OT, but in no cases were these identical models. In only seven cases did anaesthetic room screen layout match that found in theatre.

10.3.4 Discussion

The AAGBI observes that *'having a variety of models for the same purpose can increase the risk of operator confusion'*, and despite an expectation that all anaesthetists should be *'trained in the use of, and familiar with, all equipment which they use regularly'* (1), monitor provision across Oxford University Hospitals NHS Foundation Trust is poor.

In discussion with colleagues it was suggested that this may have been at least in part an active decision, with the belief that choosing from a range of manufacturers would be commercially attractive. There are justifications for this position – a variety of devices gives flexibility in function, and reduces vulnerabilities to faults in specific devices (1).

It appears that little thought was given to the ergonomic considerations of using a wide variety of devices, and it has been highlighted nationally that a problem exists in healthcare in that purchasing decisions are often made remotely from the primary users of the device (2).

This is not an isolated phenomenon, as highlighted by the current standards for equipment in critical care published by the Intensive Care Society (ICS) (3). This highlights that equipment is frequently 'not designed to be easy to use'. The ICS highlight that non-standardised equipment presents a risk, particularly to staff moving between clinical areas, that Trusts should take note. In this context, they recommend that Trusts should provide training and assessment for staff encountering new devices.

The ICS also highlight that '*unnecessary variation in the provision of health care tends to increase patient harm*'. Standardisation would confer benefits by streamlining training, and improve staff flexibility by enabling them to work confidently in other environments.

This survey successfully collected sufficient data to enable design of a study exploring the links between familiarity and professional background on visual search, ensuring the accuracy and validity of the stimuli used in the experiment described in Chapter 4.

10.3.5 References

1. Association of Anaesthetists of Great Britain and Ireland. Safe Management of Anaesthetic Related Equipment. London: Association of Anaesthetists of Great Britain and Ireland; 2009.
2. Martin JL, Norris BJ, Murphy E, Crowe JA. Medical device development: The challenge for ergonomics. *Appl Ergon.* 2008;39(3):271-83.
3. Thomas AN, Roberts JC. Standard for Equipment in Critical Care. London: Intensive Care Society.

10.4 Funding statement

This study was supported by the Laerdal Fund for Acute Medicine and the University of Oxford. The eye-tracking equipment used in this study was jointly funded by the Laerdal Fund for Acute Medicine, the University of Oxford Medical Research Fund, and the Nuffield Department of Clinical Neurosciences, University of Oxford.

The author's salary was met by the University of Oxford. The Nuffield Department of Clinical Neurosciences and Pembroke College, University of Oxford contributed to University and College fees respectively.

10.5 Acknowledgements

Although this document is the work of the author, it could not have been produced without the support of many people that I am grateful to acknowledge here.

My supervisors, Dr Helen Higham and Professor Anna Nobre, have been invaluable. Their mentorship and guidance, comments and criticism have been essential in making this document what it is. They have both given generously of their time and expertise, and no question was too trivial for them to consider. Their clarity and dedication to their fields is inspiring, and I am fortunate to have enjoyed their overwhelming support at every stage of this work.

My wife, Julie, has been a source of unending patience and tea for the last four years. She has put up with my stresses and occasional disappointments with fortitude and good humour at (almost) all times. Her experience in publishing, and expertise with proof-reading, has trapped many typographical errors that evaded my eye. No blame for the remaining errors lies with her.

Many colleagues have also given time and effort to the studies described here. Dr Sophie Robin and Dr Jonathan Dickerson both contributed their time to assist recruitment to the studies described in Chapters 6 (SR), 7 (SR), and 8 (JD). Their input was incredibly helpful, and I am certain that without their efforts these experiments would not have been as successful as they were.

The actors in the video described in Chapter 6 were Dr Martin Birch, Ms Julie Darbyshire, Mr Alan Inglis, Mr Alex Rawlings, and Ms Rosemary Warren. Dr Murooj Abdi, Dr Tom Cassar, Dr Frances Dixon, Dr Georgie Eltenton, Dr Emma Fisher, Mrs Catriona Fleming, Dr Matt Luney, Dr Victoria Millar, Dr Amy Puttick, Ms Anne Thompson, and Ms Rosie Warren led the cast of assistants in the simulation experiment described in chapter 8. Not only did the team play these roles perfectly, but their thoughts and insights about how to reliably simulate some of the device failures was incredibly helpful in planning these scenarios. I am still in awe (and slightly concerned) by the deftness with which some of the sleight of hand was performed.

Mr Charlie McDermott, Mr Alan Inglis, and Mr Russel Sinclair, the OxSTaR technical team, made my visions of the simulation a reality. Their technical and organisational skill made a potentially challenging job straightforward, and I am enormously grateful for their assistance. They are a constant, and occasionally underappreciated, source of expertise that are a benefit to everyone who works in the simulator. Their grace under fire, and ability to swiftly work around technical niggles, saved the day more than once.

I must also say a heartfelt thank you to the volunteers who took part in these studies. For professionals and students alike taking time out of busy schedules to give up time for experiments is supremely selfless act. Without their generosity, none of this work could have been possible.

On a final note, my time working with the OxSTaR office and the Nuffield Department of Clinical Neurosciences has been some of the best of my working life. It has been a privilege to work with such a group of clever, humorous, and inspiring people for the last 4 years, and I shall miss them all.

26th September 2016

PG

10.7 Publications and Presentations Arising from this Thesis

The following publications and presentations have arisen from material generated by this work.

Where available, copies of these manuscripts are provided in the following pages.

Journal Publications:

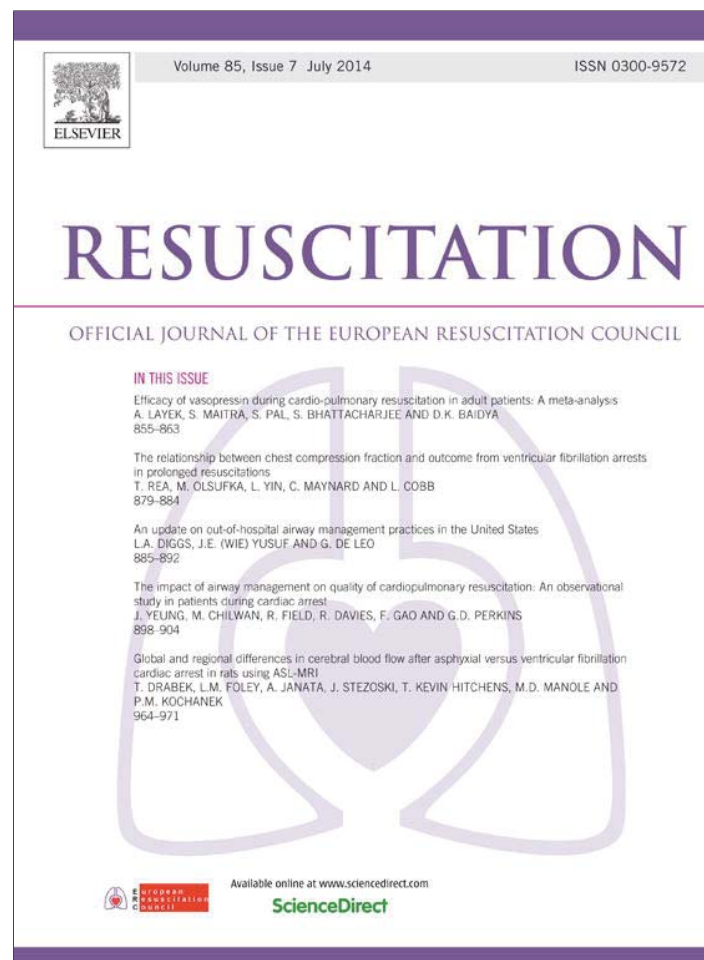
- Greig PR, Higham H, Nobre A Failure to perceive clinical events: an under-recognised source of error. *Resuscitation* 2014; **85**: 952-6.
- Greig PR, Higham H, Vaux E Lack of standardisation between specialties for human factors content in postgraduate training: an analysis of specialty curricula in the UK. *BMJ Quality & Safety* 2015; **24**: 558-60.
- Greig PR, Higham H, Vaux E Response to letter from Youngson et al. *BMJ Quality & Safety* 2016; **25**: 132-3.

Oral and Poster Presentations:

- Greig PR, Higham H, Nobre AC Perceptual Error in clinical practice: when seeing isn't believing. Proceedings of the Anaesthetic Research Society Meeting: Merton College, Oxford, UK, April 11–12, 2013. *British Journal of Anaesthesia* 2013; **111**: 318P-9P.
- Greig PR, Higham H, Nobre A. The use of eye-tracking with simulated monitors: a case for standardisation? *Changing Behaviours: Association for Simulated Practice in Healthcare Annual Meeting*. Nottingham, 2014.
- Greig PR, Higham H, Nobre A. Perceptual failure: an under-recognised source of error *Changing Behaviours: Association for Simulated Practice in Healthcare Annual Meeting*. Nottingham, 2014.

10.8 Manuscripts of published work

The following pages contain copies of various papers and posters arising from this thesis at time of submission.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

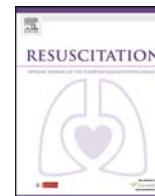
In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at ScienceDirect

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Simulation and education

Failure to perceive clinical events: An under-recognised source of error[☆]Paul R. Greig^{a,*}, Helen Higham^{a,b}, Anna C. Nobre^{c,d}^a Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, United Kingdom^b Nuffield Division of Anaesthetics, Oxford University Hospitals NHS Trust, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, United Kingdom^c Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, United Kingdom^d Oxford Centre for Human Brain Activity, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, United Kingdom

ARTICLE INFO

Article history:

Received 14 February 2014

Received in revised form 4 March 2014

Accepted 24 March 2014

Keywords:

Adult resuscitation

Situational awareness

Non-technical skill

Perceptual error

Inattention blindness

Experimental psychology

ABSTRACT

Introduction: Attentional focus narrows as individuals concentrate on tasks. Missing an event that would otherwise appear obvious is termed a perceptual error. These forms of perceptual failure are well-recognised in psychological literature, but little attention has been paid to them in medicine. Cognitive workload and expertise modulate risk, although how these factors interplay in practice is unclear. This video-based experiment was designed to explore the hypothesis that perceptual errors affect clinicians. **Methods:** 142 volunteers with varying levels of experience of adult resuscitation were shown a short video depicting a simulated cardiac arrest. This video included a series of change-events designed to elicit perceptual errors. The experiment was conducted on-line, with participants watching the video and then responding via combinations of open-ended free-text and directed questioning.

Results: 141 people experienced at least a single perceptual error. Even the most clinically significant event (disconnection of the patient's oxygen supply) was missed by three in four viewers. Although expertise was associated with increased likelihood of detecting an occurrence, even highly significant events were missed by up to two thirds of the most experienced observers.

Discussion: This study demonstrates, for the first time, that perceptual errors occur during healthcare-relevant scenarios at significant levels. Events such as an oxygen malfunction would meaningfully affect patient outcome and, although expertise conferred some advantages, events were still missed more often than not. Data acquisition is fundamental to good-quality situational awareness. These results suggest perceptual error may be a contributor to adverse events in practice.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Situational awareness, a term that describes the ability of an individual to process information about the environment in which they are functioning, is considered to be a “safety-critical skill”,¹ particularly in emergency-care type settings. The development of situational awareness can be modelled as a three-step process:² acquisition of relevant information, integration of that information into a coherent mental model, and then use of the mental model to make decisions. A failure in any one of these steps limits

effective care delivery. Data from both healthcare and industry already demonstrate that loss of situational awareness contributes to a high proportion of critical incidents.^{3–5}

Many factors can cause a loss of situational awareness. Poor training or judgement may limit the synthesis of mental models or the decisions made using them, and these may be the focus of investigations after serious untoward incidents (SUI). Comparatively little attention has been paid to human performance limitation, and specifically the role of perception, in loss of situational awareness.⁶ Introspection, based on day-to-day experience, leads us erroneously to conclude that our ability to perceive objects in our environment is unlimited, richly detailed, and automatic.⁷ There is a general assumption that educated, trained, experienced practitioners are skilled in noticing events in their environment.

This assumption is ill-founded, and it is surprisingly easy to demonstrate that perception is fallible. A number of studies in experimental psychology have demonstrated that individuals frequently miss conspicuous events when placed under relatively

[☆] A Spanish translated version of the summary of this article appears as Appendix in the final online version at <http://dx.doi.org/10.1016/j.resuscitation.2014.03.316>.

* Corresponding author at: OxSTaR Centre, Nuffield Division of Anaesthetics, Modular Building Level 2, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, United Kingdom.

E-mail address: paul.greig@ndcn.ox.ac.uk (P.R. Greig).

trivial stress, and that insight into these failings is universally poor. Consider this situation: you are stopped in the street by a stranger asking for directions. During the conversation you are distracted, and during this distraction the stranger is swapped with a different person. When you return to the conversation you are talking to a completely new individual. How likely are you to notice the change? When asked, 98% of people questioned expected to notice the switch.⁷ This experiment was carried out, but fewer than 50% of people tested were in fact aware that their conversation partner had changed part-way through their interaction.⁸ This phenomenon is termed 'change-blindness', and the poor insight into the effect 'change-blindness-blindness'.

A related phenomenon, 'inattention blindness' occurs when the brain edits out unexpected events before the individual ever becomes aware. A classic demonstration of inattention blindness involves watching a video of teams passing basketballs.⁹ Part-way through the sequence an unexpected event occurs, as an experimenter crosses the scene dressed in a gorilla costume. If participants simply watch, without a primary task on which to concentrate, all report witnessing the costumed individual. If observers are given a concentration task to perform, for example counting the number of passes made by one team, around half will remain completely unaware of the gorilla's presence. The more taxing the concentration task, the more likely observers are to miss the unexpected event. It is possible that perceptual errors follow the 'inverted-u' pattern described by the Yerkes–Dodson Law,¹⁰ which suggests that task performance generally suffers at both high- and low-extremes of arousal. There is some evidence to suggest an optimal level of distraction may exist. Rates of inattention blindness may paradoxically fall in the presence of low-level distractors in another sensory modality.¹¹

The role of memory in perceptual error has been the subject of debate in the psychological literature.¹² It was thought that the results of the original perceptual error studies might represent a failure to encode the unexpected stimulus into memory – in essence the object was perceived but swiftly forgotten. More recent work has demonstrated that memory traces for such objects do in fact persist even in the event of a perceptual failure,^{13,14} suggesting that these phenomena are not primarily the result of memory lapse.

Change-blindness and inattention blindness are two aspects of 'visual awareness failure', an umbrella term that encompasses a range of well-recognised limitations in perceptual ability, including also 'repetition blindness' and 'attentional blink'.¹⁵ Taken together, these phenomena highlight significant gaps in human perceptual abilities. It seems plausible to expect that errors of perception occur in clinical practice, where staff are routinely placed under far higher degrees of stress than described in these studies. Clinical personnel operate in high-stakes, high-pressure environments where accurate decisions must be made, often under significant time constraints.

There are reasons to expect susceptibility to perceptual errors decreases with increasing expertise. When performing a familiar task it is known that experts can derive more relevant information from single glances than do novices,¹⁶ and that experts report lower workload for comparable activities.¹⁷ In the context of perceptual errors however, the protective effects of expertise appear to be highly task-specific – when the gorilla video described above is shown to professional basketball players, they are far more likely to notice the gorilla's presence,¹⁸ but professional athletes from other disciplines perform at levels comparable to the general public. It is unclear how expertise and workload effects interact in clinical practice. If perceptual errors occur at the frequencies suggested by psychologists' data they may be significant, but largely unrecognised, contributors to loss-of-situational-awareness incidents in healthcare.

Our group, and others, have previously demonstrated that clinical training does not appear to confer any generalisable protection against perceptual error.^{19,20} These previous studies have used variations of the Simons "invisible gorilla" paradigm. There are good reasons, however, to expect people not to notice items (such as gorillas) if they are not a feature of their normal working environment,²¹ making it difficult to assess the applicability of these results to normal practice. This study was designed to redress that gap.

2. Methods

The experiments were approved by the University of Oxford Central University Research Ethics Committee (MSD-IDREC-C1-2012-139). All participants gave informed consent.

Volunteers for this experiment were largely drawn from the University of Oxford and staff of NHS South Central region. Participants were approached by email invitation, and were also free to pass the invitation on to colleagues at their own discretion. Participants were stratified by their level of resuscitation training. Individuals certified as basic life support (BLS) providers or with no formal resuscitation training were placed in the basically trained group. Accredited advanced life support (ALS) providers formed the advanced group, while ALS instructors were categorised as experts. The experiment was conducted online, using materials created locally. Participants completed the survey in their own time and on their own computers. They were instructed to ensure that their computer volume was audible and set to a comfortable level, and instructions were presented textually. Questions were placed on-screen individually, and progress through them was self-paced.

Participants were shown a 50 s video, created specifically for this experiment, depicting a simulated adult resuscitation in progress. Participants initiated playback by clicking a button on-screen after reading their instructions, and once started the sequence could not be paused or rewound. Participants were instructed to observe the team and to be prepared to comment on the appropriateness of CPR and defibrillation technique. After watching the sequence free-text comments were collected on the team's performance.

The video sequence shown to the participants incorporated a number of change blindness probes and a single inattention blindness probe. After 12 s has elapsed the team pause to evaluate the patient's rhythm. While the camera cuts to the defibrillator a number of instantaneous changes occur. In increasing clinical importance, the changes are: two team members swap hats, one team member acquires a stethoscope around his neck, the provider delivering CPR is swapped with a bystander, and the patient becomes instantaneously intubated. These are illustrated in still images taken from the sequence (Fig. 1).

After the rhythm check the patient is defibrillated and a dose of adrenaline given. During administration of the adrenaline (40 s elapsed) the oxygen supply becomes disconnected from the wall. The release of the oxygen tubing was not masked by an edit, thus was both visible and associated with a change clearly audible in the video soundtrack. This event tested for inattention blindness.

After commenting on the team's performance participants were asked, by open-ended questions, to describe any unexpected or unusual events that they had witnessed. Comments were again collected in free-text. Following these responses, all observers were presented with a list of potential events and asked to select any that they witnessed (Table 1).

After answering all questions participants were asked a final question to determine familiarity with the phenomenon of perceptual error, and whether they had previously seen the Simons 'Invisible Gorilla' video.

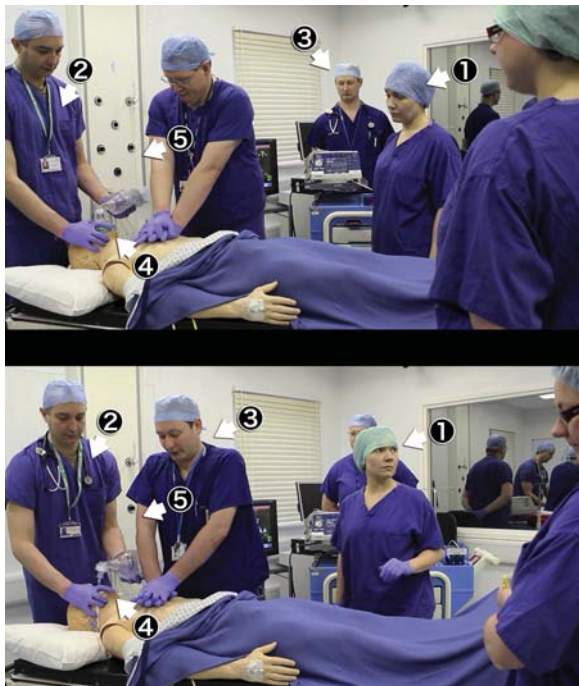


Fig. 1. Still images taken from the experimental video sequence. The two images represent before (top) and after the cut to the defibrillator (bottom). Five change events have occurred: the two female team members have swapped hats (1), the airway provider has acquired a stethoscope (2), the CPR provider has switched with the individual in the background (3), the patient has been intubated (4), and by this point the oxygen has become disconnected (5).

Table 1
List of events from which participants were invited to select any they had witnessed.

“Did you notice any of the following events occurring? (Please tick any that apply)”

- A change in the number of team members
- The airway person changed
- The CPR person changed
- The cardiac rhythm changed
- The patient’s airway changed
- The defibrillator malfunctioned
- The oxygen malfunctioned
- The team at the end were wearing different coloured tops
- The team at the end were wearing different coloured hats
- The airway person had acquired some new equipment by the end of the scene
- The team leader had acquired some new equipment by the end of the scene

Sample size was calculated to detect a difference of 30% in ‘noticing rate’ between expert and novice groups, a level comparable to the expertise-effect demonstrated amongst athletes.¹⁸ The continuous variables (age and number of resuscitations attended annually) were not normally distributed so non-parametric comparisons were made using the Kruskal–Wallis H test, and pairwise comparisons performed with Dunn’s procedure and a Bonferonni correction. Chi-squared analysis (or Fisher’s exact test where indicated) was employed to compare proportional data. Two-sided *p*-values are quoted for each test. Binary logistic regression was planned to explore the relationship between predictors of ‘noticing rate’.

3. Results

The study recruited 142 participants. The professional background of volunteers and their resuscitation background is summarised in Table 2, and group demographics in Table 3.

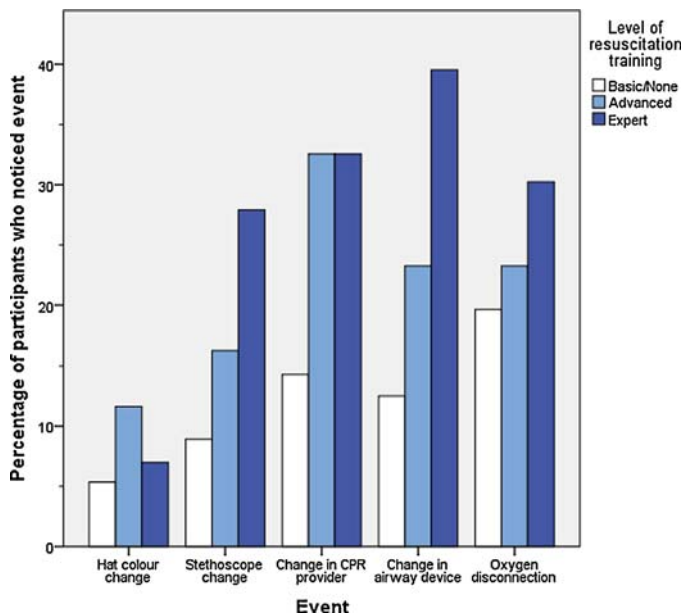


Fig. 2. Summary of participants who correctly reported noticing each event, by level of resuscitation experience.

The basically trained group was both younger ($p < 0.01$) and attended fewer resuscitations ($p < 0.01$) than the other two groups. Advanced and expert providers did not differ significantly in age ($p = 0.59$), but experts attended significantly more resuscitations annually than both basic- and advanced-level providers ($p < 0.01$).

Eight participants failed to offer comments on the quality of the resuscitation team’s performance, leaving the box blank. It is difficult to confirm their engagement with the primary concentration task. All of these, however, experienced perceptual errors.

Responses to the prompted list of changes are summarised in Fig. 2 and Table 4. Only a single participant, an advanced provider, noticed every event. If the event with the least clinical significance were to be discounted (the hat change), a further 3 participants (all expert providers) noticed everything. Overall only 34 of 142 participants (23.9%) noticed the most clinically important event, the oxygen disconnection.

There was a significant association between experience and noticing the change in the stethoscope, the change in CPR provider, and the change in the patient’s airway. There was no significant association between experience and reporting the change in hat colour, or the oxygen disconnection. Multifactorial analysis by logistic regression was planned, but not conducted, as level of resuscitation training was the only association detected by unifactorial analyses.

More than half of the participants ($n = 77$) had previously seen the ‘Invisible Gorilla’ video sequence. There was no association between being aware of the inattentional blindness effect and noticing rate for any of the events.

4. Discussion

While it might be easy to dismiss some perceptual errors as camera trickery or stage-magic techniques, these data demonstrate for the first time that perceptual errors do affect healthcare staff in dynamic situations, occurring during events directly relevant to clinical practice. Furthermore, they support the expectation that perceptual errors are prevalent even amongst expert care providers. The high levels of perceptual errors witnessed here, were they to occur in a genuine emergency, would meaningfully affect care delivery and patient outcome.

Table 2
Summary of participants by professional background and resuscitation experience.

		Professional background (n)					Total
		Doctor	Medical student	Nurse	Resuscitation officer	Non-clinical	
Level of resuscitation training	Basic/none	5	48	0	0	3	56
	Advanced	25	16	2	0	0	43
	Expert	22	0	6	15	0	43

Table 3
Participant demographics.

		Males	Females	Median and interquartile range			
				Age	Resuscitations annually		
Level of Resuscitation Training	Basic/none	26	30	22.5	20–24	0	0–1
	Advanced	28	15	31	24–44	2	0–6
	Expert	18	25	40	34.5–45	25	10–40

Table 4
Summary of noticing rates for each group, by level of resuscitation training.

		Noticing rates				p-Value
		Overall	Level of resuscitation training			
			Basic/none (n = 56)	Advanced (n = 43)	Expert (n = 43)	
Event	Hat change	11 (7.7%)	3 (5.4%)	5 (11.6%)	3 (7.0%)	0.499
	Stethoscope change	24 (16.9%)	5 (8.9%)	7 (16.3%)	12 (27.9%)	0.044
	CPR provider switch	36 (25.4%)	8 (14.3%)	14 (32.6%)	14 (32.6%)	0.05
	Airway device switch	34 (23.9%)	7 (12.5%)	10 (23.3%)	17 (39.5%)	0.008
	Oxygen disconnection	34 (23.9%)	11 (19.6%)	10 (23.3%)	13 (30.2%)	0.469

A video-based study such as this does not fully replicate the rich environment in which staff routinely work. Although being present at a resuscitation situation offers greater opportunities to gain situational awareness, it also offers greater stress and additional distractors. It is difficult to state with certainty, but echoes of perceptual errors can be seen in reports of critical incidents. The unfortunate death of Elaine Bromiley,²² occurring during a “can’t intubate-can’t ventilate” situation in anaesthesia, may be one such example. It seems likely that faulty perception explains part of the process underpinning fixation errors, whereby relevant clinical information (in this case critically low oxygen saturations) can apparently pass unnoticed at times of extreme stress.

The mechanisms by which perceptual errors occur are varied, but can generally be conceptualised as an overloading of cognitive systems that process sensory information. While there exists in the environment a potentially infinite amount of data available to be perceived, the capacity of the human brain to process that data is limited. Most of the available information is irrelevant, thus for the brain to process it would waste capacity. The brain has instead evolved to automatically record only sufficient task-relevant data to gain the gist of a scene. Fine detail is not recorded automatically, and instead is acquired as needed. This has been likened to using the environment as an “external memory bank”,²³ reducing cognitive workload and obviating the need to process richly detailed constructs of non-relevant objects. While these mechanisms function effectively in a predictable environment they may be sub-optimal in a rapidly changing situation.

The various filters that exist to minimise processing load are not under conscious control, therefore the individual is prone to ignoring unexpected but potentially important events. Vulnerability to inattentive blindness is likely an ubiquitous trait, so detecting some level of perceptual error in healthcare professionals is not unexpected. It is disappointing, although perhaps not surprising, to have detected high levels of perceptual errors across all groups in this study. If one considers the two events that would affect patient

management most significantly, the change in airway device and the malfunction of the oxygen supply, more than three out of every four observers did not notice these events. The more highly-trained groups performed better than the novice group in each case, but that only one in three experts noticed the oxygen disconnection suggests that perceptual error could be a significant contributor to loss-of-situational-awareness events in practice.

The expert group did not significantly out-perform the novice group in all cases. Experienced clinicians are better able to focus on only salient information sources during emergency situations, freeing up cognitive resources for other tasks.²⁴ The colour of colleagues’ attire is a clinical irrelevance with no bearing on care-delivery, thus the lack of association between expertise and noticing of the hat change comes as no surprise. Individuals are also much more likely to notice changes when they affect the interpretation of a scene, rather than changes which do not.^{25,26} This may go some way to explaining why the airway change showed the strongest expertise effect: intubation mandates a change to CPR delivery (moving from a 30:2 ratio to asynchronous compressions). Swapping the airway in this manner changed the meaning of the scene for the experts, but not necessarily for novices unaware of this requirement.

More puzzling is the weak effect of expertise when considering the oxygen disconnection. It was associated with both a visual event at the moment of disconnection, and an audible change which persisted to the end of the scene. A functioning oxygen supply is a critical piece of equipment during adult cardiac arrest situations (indeed the application of high-flow oxygen is a pass/fail point during summative assessments at the conclusion of all adult resuscitation courses, on which the experts are trained to examine). It is a surprise that so many did not notice this malfunction.

It is possible expertise in this medical context is less clearly definable than in sports, where environments and conditions are more strictly standardised. It could also be speculated that, as the final event in the sequence, the rates of noticing might have been

depressed by search satisfaction. If this is the case, it seems to have affected the novice group more than the expert group, as only three of the 11 inexperienced observers who noticed the oxygen disconnection also reported seeing another event, compared to seven of 13 in the expert group.

Although not as strong an effect as expected expertise remains the only predictor of inattentive blindness detected in this dataset. No association could be detected between noticing each event and age, gender, professional background, or knowledge of the inattentive blindness paradigm.

This final point may be important. More than half of the observers were aware of inattentive blindness through prior exposure to the 'Invisible Gorilla' sequence. That this group did not perform any better than the non-exposed group suggests that simply raising awareness of the phenomena will not be sufficient to protect staff against their effects.

It seems likely that perceptual errors are innate and universal features of human cognition. No data yet exist about specific training interventions that can mitigate against their effects, however there is industrial evidence that teams which communicate effectively are less vulnerable to perceptual errors than teams that do not:²⁷ while some members of the team might miss an event, so long as one person notices and can successfully share that information, the team as a whole can be protected. This validates the case for structured team-training, with a focus on effective communication skills, handover, and leadership skills.

5. Conclusions

This experiment demonstrates that trained clinicians are vulnerable to perceptual errors even when viewing scenes with which they have specific expertise. This experiment establishes as plausible the hypothesis that perceptual failures are a cause of loss of situational awareness in practice.

The participants in this study were conventionally trained to a standard appropriate to their clinical role, and there is nothing here or in other literature, to suggest that similar results would not be found were the study to be replicated elsewhere. Very few people are aware they may be vulnerable to these effects until they are exposed to materials such as this. Although these videos do not replicate the experience of clinical practice, they do depict a realistic scenario with which many healthcare personnel are faced. We believe it likely that perceptual errors are a widespread and unrecognised source of error in healthcare.

In order to determine the most effective strategies to mitigate the effects of these phenomena, future research must be directed towards establishing how widespread perceptual errors are during the management of emergencies in live clinical settings, which practitioners are at greatest risk, and under what circumstances these errors are enabled.

Conflicts of interest statement

The authors of this paper have no conflicts of interest to declare.

This study did not receive any specific funding. Use of the simulation facilities of the OxSTAR Centre for one afternoon was generously donated by the Nuffield Department of Clinical Neurosciences, University of Oxford, to permit filming of the video sequence.

Acknowledgements

The authors wish to thank (in alphabetical order) Dr M Birch, Miss J Darbyshire, Mr A Inglis, Mr A Rawlings, and Ms R Warren for assisting in the production of the video sequence. The authors are also grateful to the Nuffield Department of Clinical Neurosciences, University of Oxford, for the use of the Oxford Centre for Simulation, Training and Research for the purposes of making the recording. Recruitment was assisted by our colleagues in the Resuscitation Departments around our region, and in particular by Mrs S Hampshire, Dr M Ince, Miss S Robin, Mrs A Thompson, and Ms J Wright.

References

- Vincent C. Patient safety. 2nd ed. Oxford: Blackwell; 2010.
- Endsley M, Garland D. Theoretical underpinnings of situation awareness: a critical review. *Situational awareness: analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 1–24.
- Singh H, Petersen LA, Thomas EJ. Understanding diagnostic errors in medicine: a lesson from aviation. *Qual Saf Health Care* 2006;15:159–64.
- O'Brien KS, O'Hare D. Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics* 2007;50:1064–91.
- Walker GH, Stanton NA, Kazi TA, Salmon PM, Jenkins DP. Does advanced driver training improve situational awareness. *Appl Ergon* 2009;40:678–87.
- Chabris CF, Weinberger A, Fontaine M, Simons DJ. You do not talk about Fight Club if you do not notice Fight Club: inattentive blindness for a simulated real-world assault. *iPerception* 2011;2:150–3.
- Levin DT, Momen N, Drivdahl S. Change blindness blindness: the metacognitive error of overestimating change-detection ability. *Vis Cogn* 2000;7:397–412.
- Simons DJ, Levin DT. Failure to detect changes to people during a real-world interaction. *Psychon Bull Rev* 1998;5:644–9.
- Simons DJ, Chabris CF. Gorillas in our midst: sustained inattentive blindness for dynamic events. *Perception* 1999;28:1059–74.
- Coleman AM. A dictionary of psychology. 3rd ed. Oxford University Press: Oxford; 2009.
- Beanland V, Allen RA, Pammer K. Attending to music decreases inattentive blindness. *Conscious Cogn* 2011;20:1282–92.
- Simons DJ. Attentional capture and inattentive blindness. *Trends Cogn Sci* 2000;4:147–55.
- Simons DJ, Chabris CF, Schnur T, Levin DT. Evidence for preserved representations in change blindness. *Conscious Cogn* 2002;11:78–97.
- Jensen MS, Yao R, Street WN, Simons DJ. Change blindness and inattentive blindness. *Wiley Interdiscip Rev: Cogn Sci* 2011;2:529–46.
- Levin DT, Varakin DA. No pause for a brief disruption: failures of visual awareness during ongoing events. *Conscious Cogn* 2004;13:363–72.
- Drew T, Evans K, Vo ML, Jacobson FL, Wolfe JM. What can you see in a single glance and how might this guide visual search in medical images. *Radiographics* 2013;33:263–74.
- Weinger MB, Herndon OW, Zornow MH, Paulus MP, Gaba DM, Dallen LT. An objective methodology for task analysis and workload assessment in anesthesia providers. *Anesthesiology* 1994;80:77–92.
- Memmert D. The effects of eye movements, age, and expertise on inattentive blindness. *Conscious Cogn* 2006;15:620–7.
- Greig PR, Higham H, Nobre AC. Perceptual error in clinical practice: when seeing isn't believing. Proceedings of the anaesthetic research society meeting. *Br J Anaesth* 2013;111:318P–9P.
- Drew T, Vö ML-H, Wolfe JM. The Invisible Gorilla strikes again: sustained inattentive blindness in expert observers. *Psychol Sci* 2013;24:1848–53.
- Most SB, Scholl BJ, Clifford ER, Simons DJ. What you see is what you set: sustained inattentive blindness and the capture of awareness. *Psychol Rev* 2005;112:217–42.
- Harmer M. Independent review on the care given to Mrs Elaine Bromiley on 29 March 2005. London: Clinical Human Factors Group; 2005.
- O'Regan JK. Solving the real mysteries of visual perception: the world as an outside memory. *Can J Psychol* 1992;46:461–88.
- Schulz CM, Schneider E, Fritz L, et al. Visual attention of anaesthetists during simulated critical incidents. *Br J Anaesth* 2011;106:807–13.
- Werner S, Thies B. Is 'change blindness' attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Vis Cogn* 2000;7:163–73.
- Feil A, Mestre JP. Change blindness as a means of studying expertise in physics. *J Learn Sci* 2010;19:480–505.
- Tollner-Burgasser A, Riley MA, Nelson WT. Individual and team susceptibility to change blindness. *Aviat Space Environ Med* 2010;81:935–43.

**BMJ Quality
& Safety**

Lack of standardisation between specialties for human factors content in postgraduate training: an analysis of specialty curricula in the UK

Paul R Greig, Helen Higham and Emma Vaux

BMJ Qual Saf published online June 3, 2015

Updated information and services can be found at:

<http://qualitysafety.bmj.com/content/early/2015/06/03/bmjqs-2014-003684>

These include:

**Supplementary
Material**

Supplementary material can be found at:

<http://qualitysafety.bmj.com/content/suppl/2015/06/03/bmjqs-2014-003684.DC1.html>

References

This article cites 9 articles, 3 of which you can access for free at:

<http://qualitysafety.bmj.com/content/early/2015/06/03/bmjqs-2014-003684#BIBL>

**Email alerting
service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>

Lack of standardisation between specialties for human factors content in postgraduate training: an analysis of specialty curricula in the UK

Paul R Greig,¹ Helen Higham,^{1,2} Emma Vaux^{3,4}

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2014-003684>).

¹Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

²Nuffield Department of Anaesthetics, Oxford University Hospitals NHS Trust, Oxford, UK

³Department of Renal Medicine, Royal Berkshire Hospital NHS Foundation Trust, Reading, UK

⁴Royal College of Physicians, London, UK

Correspondence to

Dr Paul R Greig, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford OX3 9DU, UK; paul.greig@ndcn.ox.ac.uk

Received 14 October 2014

Revised 24 April 2015

Accepted 15 May 2015

INTRODUCTION

It is well recognised that a significant proportion of errors involving trainee doctors result from failures of non-technical skills (NTS),¹ which occur at least as frequently as knowledge and technical errors.² Regardless of background, all trainees need generic skills of leadership, decision-making, team-working and resource management.³ It might, therefore, be expected that curricula for different specialties would use similar definitions and teaching methods to specify NTS standards. We have performed an analysis of medical training curricula to determine the extent to which different medical specialties set training objectives in NTS, and to seek trends in the prominence with which these skills feature.

METHODS

All hospital-based medical, surgical and critical-care specialties were obtained in mid-2013, along with each curriculum's immediate predecessor (where available). The curricula were initially searched for the core keywords 'non-technical skills', 'situational awareness' and 'human factors', as well as a list of secondary keywords (generated by a modified Delphi process) grouped under headings 'task management', 'team working', 'situational awareness' and 'decision making'. The list was refined over two generations before consensus was reached. Each curriculum was analysed using NVivo V10 (QSR International, Warrington, UK).

RESULTS

Curricula of 31 medical, 3 critical-care and 12 surgical specialties were reviewed comprising approximately two million words of text over 88 documents (see online supplementary appendix tables A1

and A2). NTS terms occurred infrequently across most of the examined texts, with most occurrences in anaesthesia, emergency medicine (EM) or intensive care medicine (table 1).

Only these critical-care specialties specify requirements for formal training in NTS, using methods including self-study, tutorials and simulation. NTS are generally described as both knowledge and behaviours requiring observation and assessment, including at professional examinations.⁴ The availability of a validated assessment tool in anaesthesia⁵ facilitates the delivery of such training.

No comparable training or assessment requirements exist in the curricula of medical and surgical specialties. Although well placed to train and assess NTS, simulation is seldom mentioned outside critical-care specialties, and almost exclusively recommended only for procedural skills training or resuscitation. NTS-related objectives, where specified, are generally considered knowledge skills and assessed accordingly.

Several medical curricula contained no relevant instances of any keywords. These curricula touched on aspects of NTS, but lacked in specific detail on learning objectives or assessment recommendations. For example, communication was considered only in the context of doctor-patient relationships, language skills or record-keeping. Safety was considered in terms of adherence to guidelines, evidence-based practice and continuing professional development. Understanding components of team-working is expected by many of these curricula, but it was generally unclear what the curriculum authors consider these components to be, or how they should be taught.

To cite: Greig PR, Higham H, Vaux E. *BMJ Qual Saf* Published Online First: [please include Day Month Year] doi:10.1136/bmjqs-2014-003684

Short report

Table 1 Instances and examples of core non-technical-skills keywords

Specialty	HF	NTS	SA	Sample text from selected curricula	Assessment tool
Anaesthetics	14	28	12	(Demonstrate) 'awareness of human factors ... and the importance of non-technical skills in achieving consistently high performance such as: effective communication, team-working, leadership, decision-making and maintenance of high situation awareness ' (A27 critical incidents/attitudes and behaviours) (Describe) 'the central role human factors plays in developing a culture of safe practice and how collaboration and team-working enhances safety' (advanced training (domain 1: clinical practice)/knowledge and skills)	CbD, CEX, examination, simulation ALMAT, MSF
Emergency medicine	3	4	4	(The trainee) 'ensures the primacy of patient safety in all aspects of communication and cooperation and is able to utilise cognitive strategies, human factors and CRM to maximise this'—CC15 communication with colleagues and cooperation/behaviour in emergency department context (Demonstrate) 'awareness of human factors concepts and terminology and the importance of non-technical skills in achieving consistently high performance such as: effective communication, team-working, leadership, decision-making and maintenance of high situation awareness '—(C critical incidents/attitudes and behaviours) 'Demonstrate good non-technical skills such as: effective communication, team working, leadership, decision making and maintenance of high situational awareness ' (C critical incidents/skills)	ACAT, CbD, CEX, MSF ACAT, CbD, examinations ACAT, CbD, DOPS, simulation
Intensive care medicine	3	4	2	'Outline human factors theory and understand its impact on safety' (2.7 prioritisation of patient safety in clinical practice/knowledge and 11.3 identifies hazards and promotes safety for patients and staff/knowledge) (Demonstrate) 'awareness of human factors concepts and terminology and the importance of non-technical skills in achieving consistently high performance such as: effective communication, team-working, leadership, decision-making and maintenance of high situation awareness ' (3.10 critical incidents/attitudes and behaviours)	CbD CEX, CbD, examinations, simulation
Allergy	1	0	0	'Outlines human factors theory and understands its impact on safety' (1.7 prioritisation of patient safety in clinical practice/knowledge)	CbD
Immunology	1	0	0	'Outlines human factors theory and understands its impact on safety' (1.7 prioritisation of patient safety in clinical practice/knowledge)	CbD
Clinical pharmacology	1	0	0	'Be able to describe ... the human factors which lead to drug use errors' (36 drug errors/knowledge)	PbD
Surgical specialties*	0	0	1	(ST3 person specification) ' situation awareness : capacity to monitor and anticipate situations that may change rapidly' (selection criteria/personal skills)	Application, interview, references
ACAT	Acute care assessment tool			Direct observation of a trainee over an extended period, commenting on management, reasoning, interpretation of signs, team-working and resource management	
ALMAT	Anaesthesia list management assessment tool			Direct observation of a trainee supervising trainee's conduct over a whole session, including management, record keeping, interpretation of signs/investigations, team-working, leadership and communication	
CbD	Case-based discussion			A post-hoc assessment of case management intended to demonstrate clinical reasoning, decision-making and medical knowledge	
CEX	Clinical evaluation exercise			Direct observation of a trainee over shorter periods. Feedback covers assessment, diagnostic skill, clinical reasoning and decision-making, safety, communication and professionalism	
DOPS	Direct observation of procedural skills			An assessment of technical skill performance; in some specialties (notably emergency medicine) reference is also made to NTS components such as SA	
MSF	Multi-source feedback			Collection of feedback from a trainee's peers from various professional backgrounds, covering knowledge, skills, performance, contribution to safety and quality, communication and team-working and probity	
PbD	Problem-based discussion			A post-hoc assessment of knowledge demonstrating reasoning, decision-making and application of knowledge in relation to drug treatment usually at a population level	

Summary of the occurrences of 'human factors', 'non-technical skill' and 'situational awareness'; unlisted specialties contained no relevant instances of the search terms in 2010.

*Some surgical specialties (cardiothoracic surgery, neurosurgery, oromaxillofacial surgery, otolaryngeal surgery, plastic surgery, paediatric surgery, trauma and orthopaedics and urology) each contained this single instance of 'situational awareness'.

CRM, crisis resources management; HF, human factors; NTS, non-technical skills; SA, situational awareness; ST3, specialty trainee 3.

DISCUSSION

There is evidence that NTS training impacts patient safety in a wide range of clinical areas,^{6–8} but NTS learning objectives feature rarely outside critical-care specialties. Although, at first glance, human factors seem primarily of interest to practitioners dealing with the most acute situations, the principles are

widely applicable, and have value in less urgent settings. Industrial data clearly show that human errors contribute to critical incidents in a breadth of circumstances,⁹ and although error frequency is higher when time and workload pressures increase,¹⁰ significant incidents still occur at lower operating tempos.¹¹

It seems likely that most curricula are drafted assuming NTS can be acquired implicitly, for example, through the observation of suitable role models, rather than recognising a need for formalised training and assessment. This may be less desirable as self-assessment of NTS quality is unreliable,¹² and the way objectives and assessments are set and measured has a motivational impact on learners.¹³ We suggest a need to move to explicit NTS teaching and assessment in all curricula.

There is a need for assessment tools to enable both trainee and assessor to objectively define acceptable standards. Validated systems exist in anaesthesia for NTS assessment, and comparable tools exist in surgery and EM, but there is a need to develop such taxonomies in other specialties. NTS are a behavioural system, and assessments must include an observational component: although retrospective assessments such as case-based discussions are often cited as assessment tools, it is unlikely that instruments designed to assess knowledge are appropriate for measuring NTS.

Many specialty programmes are recognising these needs at a local level; however, teaching and assessment can vary within and between institutions. A fundamental feature of multidisciplinary teams working in high-reliability organisations is that team members observe each other's practice and offer feedback on it. This requires each member of the team, regardless of professional background, to have a shared understanding of NTS.⁴ With such variability in training methods, the body of literature and educational materials generated by one specialty would at best be opaque, and at worst inaccessible, to another. This presents significant barriers to delivering multidisciplinary team-training.

We hope that this paper will spark a debate about the role perceived for NTS training within individual specialties, and how we can find commonalities around which to build quality training to enhance patient safety.

Contributors PRG was responsible for the design of the experiment, data collection and drafting of the manuscript.

HH and EV have made important contributions and amendments to the drafting of the manuscript.

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

Data sharing statement No additional unpublished data is available from this study.

REFERENCES

- 1 Singh H, Thomas EJ, Petersen E, *et al.* Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Arch Intern Med* 2007;167:2030–6.
- 2 Cooper JB, Newbower RS, Long CD, *et al.* Preventable anesthesia mishaps: a study of human factors. *Qual Saf Health Care* 2002;11:277–82.
- 3 General Medical Council. *Good medical practice*. London: General Medical Council, 2013.
- 4 Curriculum Development Working Group. *Curriculum for a CCT in anaesthetics*. 2nd edn. London: The Royal College of Anaesthetists, 2010.
- 5 Flin R, Maran N, Glavin R, *et al.* *Framework for observing and rating anaesthetists' non-technical skills*. Aberdeen: University of Aberdeen, 2003.
- 6 Starmer AJ, Spector ND, Srivastava R, *et al.* Changes in medical errors after implementation of a handoff program. *N Engl J Med* 2014;371:1803–12.
- 7 Neily J, Mills PD, Young-Xu Y, *et al.* Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010;304:1693–700.
- 8 Arriaga AF, Bader AM, Wong JM, *et al.* Simulation-based trial of surgical-crisis checklists. *N Engl J Med* 2013;368:246–53.
- 9 Flin R, Maran N. Identifying and training non-technical skills for teams in acute medicine. *Qual Saf Health Care* 2004;13 (Suppl 1):i80–4.
- 10 van Beuzekom M, Boer F, Akerboom S, *et al.* Patient safety: latent risk factors. *Br J Anaesth* 2010;105:52–9.
- 11 Harmer M. Independent review on the care given to Mrs Elaine Bromiley on 29 March 2005. Secondary independent review on the care given to Mrs Elaine Bromiley on 29 March 2005. 2005. <http://www.chfg.org/articles-films-guides/elaine-bromiley-report>
- 12 Arora S, Miskovic D, Hull L, *et al.* Self vs expert assessment of technical and non-technical skills in high fidelity simulation. *Am J Surg* 2011;202:500–6.
- 13 Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356:387–96.

Response to letter from Youngson *et al*

Paul Greig, Helen Higham and Emma Vaux

BMJ Qual Saf 2016 25: 132-133 originally published online October 29, 2015

doi: 10.1136/bmjqs-2015-004936

Updated information and services can be found at:
<http://qualitysafety.bmj.com/content/25/2/132>

These include:

References

This article cites 3 articles, 2 of which you can access for free at:
<http://qualitysafety.bmj.com/content/25/2/132#BIBL>

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

[BMJQS Noteworthy articles \(39\)](#)

Notes

To request permissions go to:
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:
<http://group.bmj.com/subscribe/>

Although much excellent work has been done, we still maintain that more might be done to integrate these developments into the formal training curricula across all medical specialties. We note that, when published in 2011, the paper by Crossley *et al*, which considered the validation of the NOTSS system, concluded that ‘the broad implication is that the ISCP and other surgical training programmes worldwide should consider including non-technical skills in their curriculum and assessment framework, and could have confidence in NOTSS as an assessment instrument’.

This recommendation that NTS training should be adopted into surgical curricula was consistent with the paucity of NTS-learning objectives we found in the documents analysed (which were the then-current versions for 2012–2013). This set of surgical curricula contained only a single mention of the system (found in the 2007 and 2010 editions of the trauma and orthopaedics curriculum) that stated, ‘A pilot version of such a form is in development using the validated Non Operative (sic) Technical Skills in Surgery (NOTSS) system from the RCSEd/University of Aberdeen project’.

The current RCSEd website notes that NOTSS has been successfully trialled in formative assessment and ‘could be used as a summative assessment tool for progression through surgical training and beyond’.⁴ But, it is not clear to what extent it is currently employed. We have returned to the ISCP website to review how NOTSS features in 2015 curricula. Although the PDF versions of all curricula were not publically available, we were able to conduct a very brief review of the current editions for core surgical training, general surgery, oral and maxillofacial surgery, plastic surgery and trauma and orthopaedics. It does appear that, although clearly NTS receives increased prominence compared with previously, NOTSS is not specifically mentioned in any of the curriculum documents available to us. NOTSS is mentioned in the Surgical Tutor’s Handbook⁵ as

Response to letter from Youngson *et al*

We thank Professor Youngson *et al*¹ for their interest in our work² and their comments calling for greater standardisation in medical non-technical skills (NTS) teaching.

We are pleased that Professor Youngson agrees with our conclusion about the need for better cooperation and communication between specialties on how such training and assessment should be carried out. We are also delighted to acknowledge the important work being carried out by the Royal College of Surgeons (Edinburgh) (RCSEd) and the University of Aberdeen on the development of Non-Technical Skills for Surgeons (NOTSS), and are of course aware of the paper by Crossley *et al*.³

another assessment tool alongside existing workplace-based assessment (WPBA) measures. We would ideally prefer to see a consideration of NTS be part of all WPBA observations, rather than one that guidance documents imply the trainee is required to specifically request.⁶

The work of teams developing, training and using tools such as NOTSS is crucial. We applaud the efforts expended by groups to create and validate these systems. We stand by our original assertion that there is work to be done to integrate these systems into the curriculum documents that set the standards of postgraduate training and assessment. The existence of high quality and well-validated tools alone is not enough to ensure widespread adoption in all areas, and we think that inclusion into curriculum documents along with appropriate assessment tools could drive this uptake. We certainly agree with Youngson *et al* that making such progress will undoubtedly require the support of all specialty associations, Royal Colleges and the General Medical Council.

Paul Greig,¹ Helen Higham,² Emma Vaux^{3,4}

¹Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, Oxfordshire, UK

²Nuffield Department of Anaesthetics, Oxford University Hospitals NHS Trust, Oxford, UK

³Department of Renal Medicine, Royal Berkshire Hospital NHS Foundation Trust, Reading, UK

⁴Royal College of Physicians, London, UK

Correspondence to Dr Paul Greig, Nuffield Department of Clinical Neurosciences, University of Oxford, OxSTaR Centre, Modular Building Level 2, John Radcliffe Hospital, Headley Way, Oxford, Oxfordshire OX3 9DU, UK; paul.greig@ndcn.ox.ac.uk

Twitter Follow Emma Vaux at @VauxEmma

Contributors All authors contributed equally.

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Greig P, Higham H, Vaux E. *BMJ Qual Saf* 2016;**25**:132–133.

Accepted 13 October 2015

Published Online First 29 October 2015



CrossMark

- ▶ <http://dx.doi.org/10.1136/bmjqs-2014-003684>
- ▶ <http://dx.doi.org/10.1136/bmjqs-2015-004935>

BMJ Qual Saf 2016;**25**:132–133.
doi:10.1136/bmjqs-2015-004936

REFERENCES

- 1 Youngson G, Paterson-Brown S, Russ S. Response to: 'Lack of standardisation between specialties for human factors content in postgraduate surgical training: an analysis of specialty curricula in the UK' by Greig *et al*. *BMJ Qual Saf* 2016;**25**:131–2.
- 2 Greig PR, Higham H, Vaux E. Lack of standardisation between specialties for human factors content in postgraduate training: an analysis of specialty curricula in the UK. *BMJ Qual Saf* 2015;**24**:558–60.
- 3 Crossley J, Marriott J, Purdie H, *et al*. Prospective observational study to evaluate NOTSS (Non-Technical Skills for Surgeons) for assessing trainees' non-technical performance in the operating theatre. *Br J Surg* 2011;**98**:1010–20.
- 4 The Royal College of Surgeons of Edinburgh. <http://www.rcsed.ac.uk/education/patient-safety-and-notss/notss.aspx> (accessed 24 Sep 2015).
- 5 Surgical Tutor's Handbook London, Royal College of Surgeons of England, 2015.
- 6 Intercollegiate Surgical Curriculum Programme. https://www.iscp.ac.uk/static/help/NOTSS_for_the_Uninitiated.pdf (accessed 24 Sep 2015).

A total of 36 patients undergoing ENT surgery underwent three laryngoscopies, first with Macintosh, then two of the three indirect laryngoscopes as determined by block randomization. At the point of maximal laryngeal exposure, standardized lateral photographs were taken and the POGO scores were noted. Equivalent photographic images for each of the indirect laryngoscopes were then overlaid onto the lateral photographs using vector imaging software (CorelDraw, 13). The resultant images were marked for the upper incisors (I), anterior airway point (T, at the level of the crico-thyroid membrane), and effective viewing 'eye' position (E). The area of the tongue in front of the blade was analysed as in the earlier study.

The POGO results for Airtraq were significantly better than for each of the other two indirect laryngoscopes ($P < 0.001$). The mean area of the tongue between the blade and the IT line was larger for Airtraq (19.34 cm²) than Glidescope (11.85 cm²) or Truview (9.4 cm²) (Fig. 7).

Whereas for Macintosh, difficulty increases when the 'residual volume of the tongue' is not accommodated and the blade tip is pushed back from the anterior airway line (IT), with Airtraq, larger tongue volumes can be accommodated with little or no effect on the laryngeal view. In addition, its viewing 'eye' position is beyond where the main bulk of the tongue is accommodated.

The Airtraq blade has important functional and optical advantages over the others.

Reference

- 1 Charters S, Charters P. ARS London Meeting, December 13, 2012

Perceptual error in clinical practice: when seeing is not believing

P. R. Greig¹, H. Higham¹ and A. C. Nobre²

¹ Nuffield Division of Anaesthetics, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK and ² Brain & Cognition Laboratory, Department of Experimental Psychology, University of Oxford, Tinbergen Building, South Parks Road, Oxford OX1 3UD, UK

Medical error is a significant cause of avoidable morbidity and mortality: clinical errors affect one in 10 in-patients, with a fatality rate of 1 in 300.¹ Failures in non-technical skills such as situational awareness contribute to 60–80% of medical errors,² and while situational awareness depends on perception, it is clear that perception is more than simply 'seeing' or 'hearing'. Psychological research indicates that for information to reach consciousness, it must pass a series of unconscious processes that filter out data before the individual becomes aware of it. These filters are 'hard-wired' but can easily become overloaded, making the individual susceptible to perceptual failings like change- or inattentional-blindness.

Change blindness occurs when seemingly substantial or conspicuous changes in a scene pass unnoticed. It has been demonstrated using reading tasks, still images, computer graphics, and real-world interactions.³ Inattentional blindness is a similar, although distinct, effect whereby individuals focused on a task fail to notice otherwise significant events occurring around them. It has also been demonstrated using the computer, video, and real-life paradigms.⁴

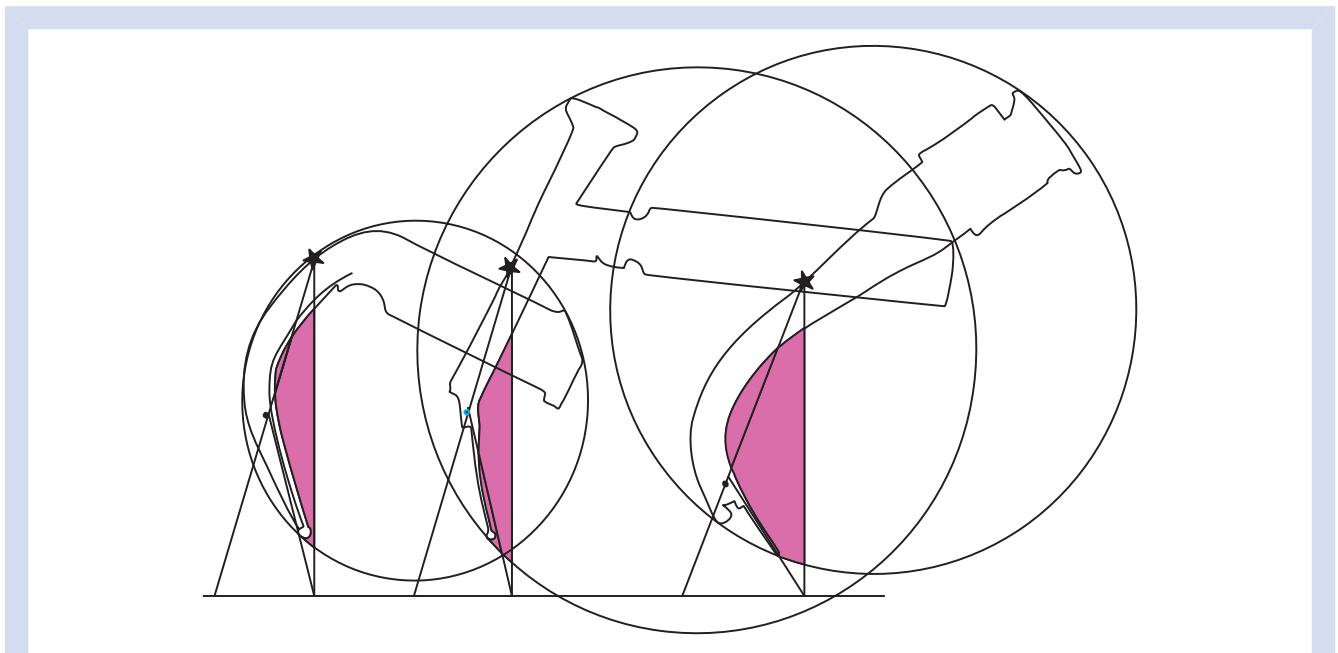


Fig 7 Three indirect laryngoscopes are shown (Glidescope, Truview, and Airtraq) relative to their vertical IT lines with T points at the same level. (The star denotes point I and the black circle point E.)

A total of 47 volunteers were recruited at the 2012 ASPiH Conference (16 doctors, 15 nurses, and 16 non-clinical staff). All were shown Dan Simons' 'Invisible Gorilla' video (video 1; www.youtube.com/watch?v=vJG698U2Mvo). Participants who had previously seen this were also shown 'Monkey Business' (video 2; www.youtube.com/watch?v=IGQmdoK_ZfY). Participants were questioned about change-events incorporated into each sequence.

Half of the participants (23 of 47) were unaware of the experiment and watched only video 1, with 13 of the 23 (56.5%) failing to notice the gorilla. The remaining participants were familiar with the experimental paradigm, and therefore watched both video 1 and 2. Only 2 of these 24 (8.3%) failed to see the gorilla in sequence one, and 3 (12.5%) failed in sequence 2. None of the participants who failed in sequence 2 also failed in sequence one. Only 7 of the 24 (29.2%) noticed all of the sequence 2 events.

These effects have been associated with real-world consequences, including faulty eye-witness testimony, road-traffic accidents, and aircraft crashes, but little attention has been paid to their impact in hospital practice. Future work will investigate the impact of perceptual error using clinically relevant stimuli.

Funding

None.

References

- 1 Vincent C. *Patient Safety*. Oxford: Blackwell, 2010
- 2 Singh H, Thomas EJ, Petersen E, Studdert D. *Arch Intern Med* 2007; **167**: 2030–6
- 3 Simons DJ, Levin DT. *Psychon Bull Rev* 1998; **5**: 644–9
- 4 Simons DJ, Chabris CF. *Perception* 1999; **28**: 1059–74

Characterizing medical students' learning during immersive simulations

I. Davies¹, H. Higham¹ and I. Lunt²

¹Nuffield Department of Anaesthetics, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK and ²University of Oxford Department of Education, 15 Norham Gardens, Oxford OX2 6PY, UK

When students qualify, they need to manage the uncertainties of their new role¹ and using simulation as a safe place for them to develop this aspect of professional practice could help them with this process. Simulation as a teaching technique is nearly ubiquitous in medical education. While feedback and good debriefing maximize learning from simulation, not much is known about the process of learning within an immersive simulation or which theory of learning is the most appropriate to use. There has been a recent consensus² on research areas in simulation which included a call to better understand simulation learning. This study characterizes the process of learning by medical students during simulation scenarios.

Groups of eight final-year medical students spent 1 day in the simulator centre as part of their anaesthesia module. They each took part in an immersive scenario using a simulation mannikin (Laerdal SimMan) of a different acutely ill medical patient, with a faculty member taking the role of a nurse to assist them. Each scenario was based on an acute medical emergency that would require urgent intervention by the student including them calling for help. In six of the eight scenarios, the attendance of an anaesthetist was appropriate, for example, for a compromised airway or breathing or for urgent surgical intervention. Each whole group of students was randomly assigned to all perform their scenario either paired with a 'buddy' of their choosing from the group or solo. Each scenario was video recorded. Afterwards, participating students were invited to attend for a 30 min semi-structured interview.

A total of 76 students were recruited voluntarily over 5 months from 12 anaesthesia modules. Sixty-eight students attended the simulation day and 65 students were interviewed. Sixty-two interviews were transcribed. The transcripts were analysed using inductive and deductive analysis for themes and the videos were used for triangulation. One theme that emerged was how the students changed their understanding of uncertainty during the simulation scenarios a change which corresponds with certain theories^{3,4} of personal epistemology development.

References

- 1 Knight LV, Mattick K. *Soc Sci Med* 2006; **63**: 1084–96
- 2 Dieckmann P, Phero JC, Issenberg SB, Kardong-Edgren S, Østergaard D, Ringsted C. *Simul Healthc* 2011; **6**: 155–67
- 3 Baxter Magolda MB. *Rev High Educ* 1996; **19**: 283–304
- 4 Perry WG. *Forms of Ethical and Intellectual Development in the College Years*. New York: Holt, Rinehart, and Winston, 1970

Does ethnicity affect acute postoperative pain and nurses attitudes to treating it?

M. Al-Hashimi*, N. Griffin-Teall*, S. Bowrey* and J. P. Thompson

Division of Anaesthesia, Department of Cardiovascular Sciences, Leicester Royal Infirmary, Leicester LE1 5WW, UK

Previous data suggest that there are significant differences in the perception, reporting, and treatment of acute and chronic pain between patients from different racial and ethnic groups.¹ Some studies have shown that patients from minority ethnic groups report greater pain, but receive fewer analgesics than Caucasian patients, although other studies have found no differences. Some of the discrepancies in the literature may be because pain and ethnicities are multidimensional and can be shaped by culture.² We performed an observational study to evaluate the racial and ethnic differences in pain perception and treatment after total abdominal hysterectomy, and whether any differences were related to the

USING EYE-TRACKING WITH CLINICAL DEVICES: A CASE FOR STANDARDISATION?

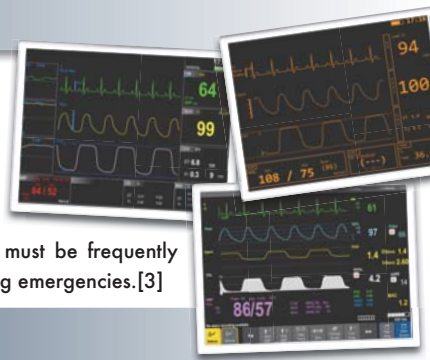


PAUL GREIG, HELEN HIGHAM, ANNA C NOBRE

Introduction

Situational awareness is a concept that embodies individuals acquiring relevant information about events occurring around them, and then using that information to make decisions. In healthcare good situational awareness is a safety-critical skill,[1] particularly relevant to acute-care specialties such as anaesthesia. One of the most frequently referred-to sources of information in such specialties is the clinical monitor attached to the patient.

The Association of Anaesthetists of Great Britain and Ireland (AAGBI) specifies monitoring standards for anaesthetic cases,[2] stipulating the nature and frequency of observations required. Monitoring is crucial to patient safety, and must be frequently referred to during rapidly changing scenarios; up to a third of visual attention may be devoted to a patient monitor during emergencies.[3]



Methods

Critical care monitors across Oxford University Hospitals (OUH) NHS Trust sites were surveyed noting make, model and default screen layout. From these data a representative sample of 11 layouts was generated, to which an additional fictional display was added. This shared no common features with the real displays.

In task 1 (T1) the monitors were presented in random order while participants recorded five values: heart rate, blood pressure, oxygen saturation, and end-tidal carbon dioxide/anaesthetic concentrations. After recording values participants rated their familiarity with the layout using a visual analogue scale (VAS) calibrated from 0 ('I have never seen this display before') to 100 ('This is the only display I ever work with').

In task two (T2) participants had to locate the respiratory rate as quickly as possible, using the same monitors presented in a new random order. Participants were not required to record any values.



Data analysis in D-Lab

Results

Figure 1: Scatterplot of familiarity score against task duration

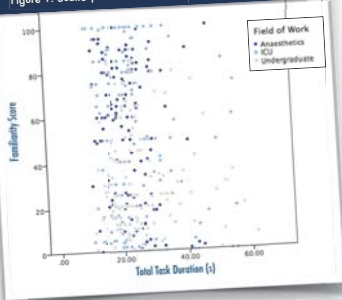
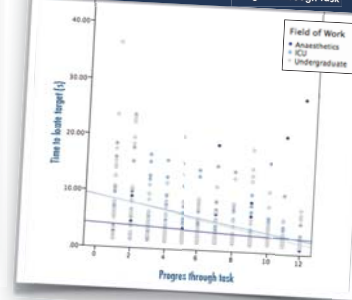


Table 1a: Equipment Summary		Devices	Manufacturer	Models	Displays
Monitor Provision	Anaesthetic Rooms	48	3	7	20
	Operating Rooms	48	3	7	19
	Recovery & ICU	88	2	8	13

Table 1b: Participants Summary		Males	Females	Mean Age	Mean VAS
Professional Background	Undergraduate	8	7	23.0	21.7
	Intensive Care	5	10	33.9	37.6
	Anaesthetics	7	8	40.3	46.5

Figure 2: Task time as a function of progress through task



Across the four OUH sites there were 40 unique display patterns, presented on 184 different monitors supplied by three manufacturers. There were only seven instances across 48 theatres where an operating theatre monitor shared identical layout with its anaesthetic room. Equipment and participants are summarised in table 1.

Excluding the fictional monitor, anaesthetists gave higher familiarity scores than both ICU staff and undergraduates ($p < 0.01$). Nine participants failed to rate the monitor in their workplace highest and seven failed to give the fictional monitor their lowest score. One participant rated the fictional monitor's familiarity as 100.

Figure 1 demonstrates very poor correlation between self-rated familiarity and overall task time ($\rho = -0.093$) in both tasks. Similarly weak correlations were observed for breadth of eye motion ($\rho = -0.05$), number of fixations ($\rho = -0.048$), or total fixation duration ($\rho = -0.129$). The mean duration of each task was 24.0 ± 0.8 s (T1) and 4.8 ± 0.4 s (T2). A significant effect was noted for professional background: anaesthetists completed T1 3.4s ($p < 0.01$) and T2 2.2s faster ($p < 0.01$) than the other groups.

ICU nurses did not significantly outperform undergraduates, but both they and the undergraduates showed improvement across T2 from first to last display. Although anaesthetists were generally faster than both other professional groups at each stage, they did not improve to the same extent. These data are summarised in figure 2.

Discussion

Equipment provision across OUH is poorly standardised. This might be expected to contribute to error. Although not explicitly tested in this experiment, existing data suggests unfamiliarity adds to mental workload. Critical-care situations make high demands on cognitive resources even before accounting for any equipment-effects.

It is surprising that correlations between reported display-familiarity and eye-tracking metrics were so poor, however self-rating may in this case be unreliable. Many participants appear to have struggled to recognise devices with which they work on a daily basis. It is known that people have difficulties replicating highly familiar items like coins, however fine details on the faces of coins are not critically important to their recognition and use (shape, size and colour are probably more important). It is hard to argue this case for clinical monitors, in which using the colour-coding and screen-placement of parameters is an intrinsic part of their use.

The performance advantage seen amongst anaesthetists, coupled with the learning effects seen amongst the other groups in task two, suggests that there is a familiarity effect at play. Although it is difficult to quantify the magnitude of the familiarity effect on the basis of these data, it would be appropriate to recommend that Trusts make more effort to standardise their equipment provision in critical-care environments.

Subscribe to the OxStaR Aurasma™ channel, and you can scan tagged items using the Aurasma™ app to display extra material using your smartphone or tablet.

The authors are grateful to the Laerdal Foundation for Acute Medicine, the Nuffield Department of Clinical Neurosciences, and the University of Oxford Medical Research Fund for their generous support of this work

References

- Vincent C. Patient Safety. 2nd ed. Oxford: Blackwell; 2010.
- Birks R, Gemmell L, O'Sullivan E et al. Recommendations for standards of monitoring during anaesthesia and recovery. 4th ed. London: AAGBI; 2007.
- Schulz C, Schneider E, Fritz L et al. Visual attention of anaesthetists during simulated critical incidents. Br J Anaesth. 2011 June 1, 2011;106(6):807-13.

PERCEPTUAL LIMITATIONS: AN UNDER-RECOGNISED SOURCE OF ERROR



PAUL GREIG, HELEN HIGHAM, ANNA C NOBRE



Introduction

Situational awareness is a 'safety-critical' skill in medicine [1], but attentional focus narrows as individuals concentrate on tasks. Missing an event that would otherwise appear obvious is termed a perceptual error. Individuals greatly overestimate their perceptual reliability [2].

Perceptual failures are well-recognised in psychological literature [3,4], but little attention has been paid to them in medicine. Cognitive workload and expertise modulate risk, although how these factors interplay in practice is unclear. This video-based experiment was designed to explore the hypothesis that perceptual errors affect clinicians.

Methods

142 volunteers, with varying levels of experience in adult resuscitation were recruited. Participants with no resuscitation qualifications, or who were BLS trained, were considered to be least experienced. Qualified ALS providers were considered to possess advanced experience. Instructors on accredited RC(UK) ALS courses were classed as experts. All were shown a short video depicting a simulated cardiac arrest. This video included a series of events designed to elicit perceptual errors. The experiment was conducted on-line, with participants watching the video and then offering open-ended free-text comments and responses to directed questions.

The events included: (1) two team members switching hats, (2) the stethoscope moving from one team member to another, (3) the CPR provider and the observer switching position, (4) the patient's airway changing, and (5) the oxygen supply disconnecting.



Figure 1: frame taken from the experimental video

Results

142 volunteers (70 females; 72 males) took part. Participants were grouped by resuscitation experience as detailed in table 1. The basically trained group was both younger ($p < 0.01$) and attended fewer resuscitations annually ($p < 0.01$) than the other groups. Advanced and expert providers did not differ significantly in age ($p = 0.59$), but the experts attended significantly more resuscitations than both the basic- and advanced-level providers ($p < 0.01$)

		Professional Background (n)					Total
		Doctor	Medical Student	Nurse	Resuscitation Officer	Non-clinical	
Level of Resuscitation Training	Basic/None	5	48	0	0	3	56
	Advanced	25	16	2	0	0	43
	Expert	22	0	6	15	0	43

Table 1: Participants' demographics

Figure 2 illustrates noticing rates for each event, grouped by participant experience. The associations between experience and noticing were significant for the stethoscope change, the switch in CPR provider, and the change in airway. There was a non-significant trend in favour of expertise in the case of the oxygen disconnection.

Discussion

This study demonstrates, for the first time, that perceptual errors occur during healthcare-relevant scenarios at significant levels. Events such as an oxygen malfunction would meaningfully affect patient outcome and, although expertise conferred some advantages, events were still missed more often than not. At best, fewer than four in ten experts noticed events that should mandate a change in treatment.

The mechanisms by which perceptual errors occur can generally be considered as the overloading of cognitive systems that process sensory information [5,6]. The brain lacks the capacity to automatically process all of the information potentially available to it, so systems exist to 'filter' data before it reaches awareness. These filters are largely sub-conscious, and therefore outside voluntary control.

Expertise would be expected to reduce vulnerability to perceptual error. Experts experience lower workload for the same tasks when compared with novices, and can derive more information from single glances. It is somewhat surprising that, while experts did outperform the least experienced, the effect was not as strong as suggested by existing literature [7].

Acquisition of accurate data is fundamental to good-quality situational awareness. These results suggest perceptual error may be a contributor to adverse events in practice.

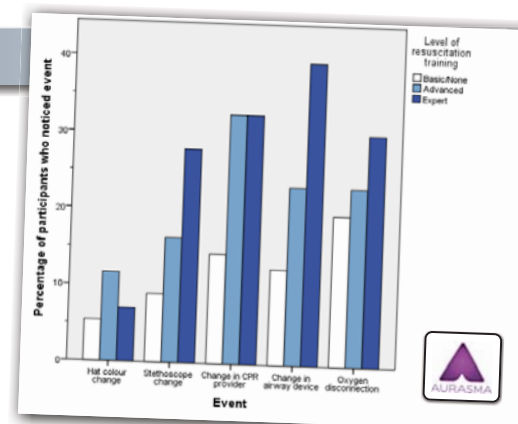


Figure 2: noticing rates for each event, by level of experience

References

- Vincent C. *Patient Safety*, 2nd edn. Oxford, Blackwell, 2010.
- Levin DT, Momen N, Drivdahl S Change blindness: the metacognitive error of overestimating change-detection ability. *Visual Cognition* 2000; 7: 397-412.
- Simons DJ, Levin DT Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review* 1998; 5: 644-9.
- Simons DJ, Chabris CF Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception* 1999; 28: 1059-74.
- Most SB, Scholl BJ, Clifford ER, Simons DJ What you see is what you set: sustained inattention blindness and the capture of awareness. *Psychological Review* 2005; 112: 217-42.
- O'Regan JK Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology* 1992; 46: 461-88.
- Memmert D The effects of eye movements, age, and expertise on inattention blindness. *Consciousness and Cognition* 2006; 15: 620-7.

This poster incorporates additional augmented reality (AR) features. Scan tagged items using the Aurasma™ app to display extra material on your smartphone or tablet.

