Latent preconditions of medication administration errors: Development of a proactive error-management tool

Samantha Jane Carruthers

Submitted in accordance with the requirements for the degree of PhD
University of Leeds
Institute of Psychological Sciences
September 2008

This work is dedicated to the memory of Jim Carruthers
The very best of men.

Acknowledgements

I would like to gratefully acknowledge Rebecca Lawton, Peter Gardner and John Wright for their enthusiastic supervision and continued support and advice over the last four years and allowing me the room to work in my own way.

Thanks also to all the fantastic nurses at Bradford Royal Infirmary and St Luke's Hospital who generously gave their time and expertise during interviews and observations in the first two years of this project and without whom this research would not have been possible.

I would not have been able to write this thesis without the help and support of my family and friends. Many thanks to Chris and Jim and the rest of the family for always being there. Special thanks to Claire for being a great friend and colleague and for her unwavering encouragement every time I reached a brick wall but most importantly for making me laugh! Thank you to Ellen for being the best and most thoughtful friend and for spending what must have been the longest and most boring weekend of her life helping me to stuff two thousand envelopes for the final study! Thank you to Shirley for simply being the first person in my life to ever inspire me and for always pointing out the silver lining!

Last but definitely not least, I am eternally grateful for the unconditional support from my amazing husband who encouraged me to pursue a life I thought was beyond my reach. His indomitable spirit inspires me to work harder every day and fulfill a potential I didn't think I had. His faith in me takes my breath away. I hope I make him as proud of me as I am of him every single day.

Abstract

Iatrogenic injury has been found to occur in around 10% of UK hospital admissions, equating to the harm of approximately 850,000 patients each year. The Department of Health has made repeated calls for NHS research to learn from proactive error management techniques (EMTs) employed within other 'safety-critical' organisations (DOH 2000, 2001). The aim of this research was to develop a valid and reliable proactive measure of latent organisational failures (EMT) for use in secondary care using a psychological theory of organisational accidents (Reason, 1990, 1997). This theory purports that errors occur as a result of a complex interaction between unsafe acts and systemic organisational weaknesses known as latent failures. This tool will be used to measure and monitor organisational safety in health care and predict the likelihood of medication administration errors (MAEs).

Twenty semi-structured interviews were conducted in study 1 with qualified nurses from several general medical wards and senior managers from Bradford Teaching Hospitals NHS Foundation Trust. Using error vignettes, participants were asked to discuss their perceptions of error causation. Additional qualitative data was collected using clinical observations and incident report review. Using thematic content analysis, ten latent workplace and organisational causes of MAEs were identified, consistent with psychological error theory and error causes evidenced within other safety-critical industries (Reason, 1997; Groeneweg, 1992; Helmreich, 2000; Colla *et al.,* 2005), including team functioning, human resources, culture and training. In terms of Reason's organisational accident model, combining three pools of independent qualitative data afforded an in-depth exploration of latent error causes at an individual (e.g. unsafe practices), workplace (e.g. team functioning) and organisational level (e.g. use of policies and protocols).

Study 2 was conducted to conceptualize identified latent preconditions of MAE within a proactive questionnaire measure; the Organisational Safety Questionnaire (OSQ). Revisiting qualitative data collected in Study 1, this study explored the ways in which each latent organisational failure would manifest at a hospital ward level. One hundred and forty-five safety indicators were generated based on these manifestations of poor safety. Pilot studies to test the face validity of indicators and content analysis to remove less commonly endorsed items led to refinement of the tool to 82 items.

Given several notable drawbacks to using NHS formal incident reporting systems as an outcome measure, study 3 was conducted to develop an independent measure of MAEs against which to test

the predictive validity of the OSQ (the Drug Round Behaviour questionnaire; DRBQ). This study explored the types of MAEs which can arise in secondary care as a direct or indirect result of the ten latent preconditions. Using the qualitative data obtained in study 1, a 27-item measure of 10 types of MAE (NCC MERP, 1995) was developed which was not reliant upon adverse patient outcomes and intended to also capture near misses. After a pilot study was conducted to improve the construct and face validity of the tool, 13 items which reflected 7 types of MAE had good face validity and were retained for study 4.

The final study was conducted to measure the validity and reliability of the OSQ. The 82-item OSQ was administered to qualified and unqualified nurses working in 54 clinical areas across 2 two Bradford hospitals. Analysis revealed that the OSQ was relevant for all *qualified* nurses working in 34 of these clinical areas. Although developed as 10 subscales representing 10 latent preconditions of MAE, factor analysis yielded only one overall construct from 28 items named 'organisational safety'. However, these items reflected 8 of the 10 proposed predictors of MAE which supports their role in the occurrence of MAE. The 28-item OSQ had good internal consistency and concurrent validity (with an independent 9-item measure of local safety culture; Vogus & Sutcliffe, 2007). While the OSQ was significantly predictive of MAEs measured by the DRBQ, it did not significantly predict formally reported incidents. However, this may have been an artefact of low statistical power which may have been improved with a larger sample. Finally, high safety risk wards said they were less likely to *formally* report their errors than lower risk wards, yet all wards reported a similar number of incidents. It is proposed that high risk wards report a comparatively smaller percentage of the errors which *actually occur* compared to lower risk wards due to poorer safety cultures. Interestingly, high safety risk wards admitted making significantly more MAEs on the DRBQ than 'safer' wards suggesting the DRBQ was a more sensitive measure of the *actual* number of drug administration errors occurring on wards. The Organisational Safety Questionnaire represents a novel, valid and reliable proactive measure of safety which is not currently available in health care which would be useful in measuring the effects on systems interventions and other organisational changes.

This thesis has explored and identified latent organisational causes of medication administration errors in secondary care and used methodological techniques used in other safety-critical industries to develop a valid and reliable measure of organisational safety which was successful in predicting medication administration errors. Findings are discussed in terms of the benefit of rigorous qualitative methods in this type of research and the direction of future research which could examine the generaliseability of the tool to other health care professionals or fields of medicine.

TABLE OF CONTENTS

CHAPTER 3. SYSTEMS INTERVENTIONS TO REDUCE MEDICATION ERRORS: A SYSTEMATIC REVIEW

CHAPTER 6. DEVELOPMENT OF THE DRUG ROUND BEHAVIOUR QUESTIONNAIRE

CHAPTER 7. VALIDATION OF THE ORGANISATIONAL SAFETY QUESTIONNAIRE

CHAPTER 8. CONCLUSIONS

FIGURES

TABLES

CHAPTER 1
INTRODUCTION AND OVERVIEW

1.1. Introduction

> Errors are to be expected, even in the best organisations. When errors occur, the question should not be "who is at fault?" but "why did our defences fail?" Focussing exclusively on individuals misses an essential part of the error story and blocks the path to effective remediation. We cannot change the human condition, but we can change the conditions under which humans work. (Reason, 2000, pp. 769)

Human error is routinely blamed for accidents in many high risk industries, particularly in health care. The public are often faced with newspaper headlines such as 'Negligent doctor orders killer jab!' (Daily Mail, July 2003), or 'Nurses hygiene to blame for superbug' (Evening Times, April 2004). To believe such headlines would be to assume that when health care professionals enter the hospital grounds they become impervious to the cognitive constraints particular to all human beings and in some ways are 'superhuman'.

Over the last decade, research and national policies have highlighted that the investigation of only the immediate causes of error has had a minimal effect on error management strategies in health care. Many researchers have indicated and driven government policy towards more detailed exploration of medical error causation which may present as an early warning sign much earlier in the chain of events preceding human error (Coxon *et al.,* 2003). This thesis will present evidence towards a clearer understanding of such 'early warning signs' in the field of secondary health care. This chapter outlines the overall research aims and objectives of the thesis and briefly outlines each of the four studies undertaken. This chapter also includes a summary of the most current international and national statistics on medical error and a brief overview of a key UK government report, *An Organisation with a memory* (Department of Health; DoH, 2000) to give context to the origins of this research.

1.2. Research sponsorship

This study was funded by the Economic and Social Research Council (ESRC), Bradford Teaching Hospitals NHS Foundation Trust (as part of an ESRC-CASE collaboration) and the University of Leeds. The overall aim of the research was to employ a mixed methods approach to explore organisational factors and other local working conditions which contribute to medication errors in secondary care. While it was the initial intention of the research to study *all* types of medical errors it quickly became evident after reviewing a number of clinical incident reports that this would be considerably beyond the scope of the project. After liaising with the collaborating institution and reviewing the evidence on the prevalence of different types of error, the decision was made to focus on medication errors, in particular those involving the administration of drugs in view of their importance in health care. Being the collaborating industrial sponsor, all field work was conducted in two of Bradford NHS Trust secondary care hospitals; Bradford Royal Infirmary and St Luke's Hospital.

1.3. Medical error statistics

In the UK, adverse incidents leading to patient harm (iatrogenic injury) have been found to occur in around 10 per cent of hospital admissions, equating to the injury of an estimated 850,000 patients per year. More than 6,600 adverse incidents involving medical devices were reported to the Medical Devices Agency in 1999; 87 of which resulted in death and 345 serious injuries (DoH, 2000). Of particular relevance to this thesis which focuses on errors made in the administration of medication, is the finding that approximately 10,000 patients in the UK each year are involved in some form of adverse drug reaction[1]. Apart from the obvious costs to the health of patients, there are also financial implications. The NHS spends around £400 million each year in settlement of clinical negligence claims and in 1998/99, had a potential liability of approximately £2.4 billion for

[1] An adverse drug reaction refers to the 'unwelcome, negative consequences associated with the use of administered medications' (Nebeker *et al.*, 2004).

expected and existing claims, many of which litigation analysis found to be avoidable. Additionally around £2 billion every year is spent in prolonged hospital stays and additional treatment costs associated with iatrogenic injury (DoH, 2000). Comparable statistics can be found internationally.

USA

The US publication *To Err is human: Building a safer health system* (Institute of Medicine; IoM, 1999), reported that between 44,000 to 98,000 hospitalised patients die each year as a result of iatrogenic injury. Even using the lowest of these figures means medical error related deaths are the eighth leading cause of mortality in the USA, higher than the number of deaths attributable to AIDS, breast cancer or car accidents (McFadden *et al.*, 2004). The IoM report estimated that medical errors cost the US around $37.6 billion each year. The Harvard Medical Practice Study (Brennan *et al.,* 1991; Leape *et al.,* 1991) constituted landmark research in this field, reviewing patient records of 30,121 hospitalizations from 51 acute care hospitals in New York State in 1984. Results revealed that Adverse Events (AE) occurred in almost 4 per cent of hospitalizations; the most common AEs were medication-related (19 per cent) and approximately 70 per cent of these events were preventable.

Australia

In the Quality of Australian Health Care Study in 1995, 14,179 hospital admission records from 1992 from 28 hospitals were reviewed to identify adverse events and adverse drug events which had resulted in 'unintended injury to patients caused by health care management' (Wilson *et al.*, 1995). Results revealed almost 17 per cent of reviewed admissions were associated with adverse events and 1.6 per cent of these were associated with adverse drug events (ADEs); almost half of which were deemed preventable. Unfortunately, this study (and many others conducted in Australia – Larmour *et al.,* 1991; Easton *et al.,* 1998; Runciman *et al.,* 2002, 2003) focused predominantly on

the number of adverse events which have *resulted in admission to hospital*[2] but made little distinction between these and the number of hospital admissions which had *resulted in adverse events* or drug events as is the case for most research in this area.

New Zealand

In the New Zealand Adverse Events Study, Davis *et al.* reported that 'of the top 20 risk factors that account for nearly three quarters of all [New Zealand] deaths annually, adverse in-hospital events come in at number 11 above air pollution, alcohol and drugs, violence and road traffic injury' (Davis *et al.*, 2002, 2003).

Canada

The Canadian Adverse Events study (Baker *et al.,* 2004) estimated the incidence of Adverse Events (AEs) in 2000 ('...unintended injuries or complications resulting in death, disability or prolonged hospital stay that arise from health care management... ') by reviewing patient records from five Canadian acute care hospitals. AEs were identified in 255 of the charts, equating to 7.5 per cent of hospital admissions in 2000; approximately 37 per cent of which were deemed by physician reviewers as preventable. It was calculated that patients spent an additional 1521 days in hospital as a result of these 255 AEs alone. Baker *et al.* extrapolate these findings to suggest that of the 2.5 million hospital admissions in Canada each year, around 185,000 are associated with an AE and approximately 70,000 of these are potentially preventable.

Iran

In a study of 370 patient records in a hospital at the Tehran Medical Sciences University, Gholami and Shalviri (1999, as cited in Kanjanarat *et al.*, 2003) found approximately 17 per cent of these patients had at least one Adverse Drug Reaction (ADR) during their hospital stay. Gholami and

---

[2] Wilson *et al.*, estimate that 49 per cent of detected adverse events were the reason for hospital admission.

Shalviri classified 10 per cent of the ADRs as mild, 86 per cent as moderate, 1 per cent as severe, and 3 per cent as lethal. Of these ADRs, approximately 60 per cent were identified as preventable and 96 per cent as predictable reactions.

1.4. An organisation with a memory

In 2000, an expert group on learning from adverse events in the NHS published the government report: *An Organization with a Memory,* which drew attention to the scale of medical error and proposed ways in which future patient safety research should be directed to effect positive changes. One particular research objective included gaining a clearer understanding of the root causes of adverse incidents in NHS care (DoH, 2000). The report acknowledged that in order to reduce the number of adverse events to patients, healthcare organisations must learn the lessons about the underlying causes of these events. The report further acknowledged that current incident analysis systems in the NHS focused largely upon 'localised causes' of adverse incidents, often blaming human actions which immediately precede events. Recommendations emphasized the need for a new strategy of incident analysis which recognizes the presence of both *active failures* (unsafe acts) and *latent conditions* (Reason, 1990) in most adverse events. Latent (dormant) conditions are organisational or other systems weaknesses (e.g. training issues, lack of resources etc.) which develop over time but which can go unnoticed or disregarded until they result in an unsafe act or combine with other latent conditions to compromise patient safety. The distinction between latent conditions and active failures will be discussed in more detail in chapters 2 and 6.

Of particular interest to this thesis, the DoH report emphasizes that future patient safety researchers should focus on ways to proactively identify these latent systemic causes of poor patient safety which are entrenched within the organisational culture to enable more effective error management. The report further cautions that;

'Any attempt at risk management that focuses primarily upon the supposed mental processes underlying error (forgetfulness, inattention, carelessness, negligence, and the like) and does not seek out and remove these situational 'error-traps' is sure to fail. The local human errors are the last and probably the least manageable part of the causal sequence leading up to some adverse event.' (DoH, 2000, pp. 21)

The research presented in this thesis aimed to explore one of the recommendations of *An Organization with a Memory*; learning safety management lessons from other non-health care high risk industries such as aviation and nuclear power generation. The report makes particular reference to James Reason's *Swiss cheese* model of accident causation and the ways in which this model has been used as a proactive method of error management in other safety-critical industries. Although this model will be described in more detail in the following chapter, in short the model proposes that 'holes' in consecutive layers of organisational defence (e.g. barriers and safeguards) created by active failures and latent conditions, can line up and create 'windows of opportunity' for accidents and errors (Reason, 1997, pp. 11). In Reason's later framework of organisational accidents he proposes that once latent conditions have been identified within an organisation, proactive measures can be taken to reduce the likelihood of errors and accidents before they arise. The DoH report acknowledges that while such models have been developed with other more technologically-driven industries in mind, much can and should be learned from these safety management techniques. Although some research has been conducted using Reason's framework of organisational accidents to retrospectively investigate the latent causes of adverse drug events (e.g. Vincent, Taylor-Adams & Stanhope, 1998), there has been little, if any, conducted with a proactive approach to error management.

1.5. Rationale and research objectives

The research conducted here is well-timed in that it responds to a call from various research councils (e.g. MRC[3], ESRC[4] and EPSRC[5]) to increase the numbers of skilled researchers in the area of patient safety working in the UK. As already mentioned, the Department of Health has repeatedly called for NHS research to learn from error management techniques employed in industries which have attempted to proactively identify latent conditions before they have the opportunity to result in error. Reason (1997) describes such proactive methods of error management in detail within a range of organisations including Shell Expro and British Airways in his book *Managing the Risks of Organisational Accidents*. However, despite the obvious appeal and face validity of these methods there is little empirical evidence available to support their adoption beyond the field in which they were developed. The aim of this research was to employ a range of qualitative methods to explore the latent conditions which predict medication administration errors in secondary care. Medication administration errors are defined as errors committed by nurses in the preparation and/or administration of medication to patients (e.g. incorrect drug/dosage/patient). As such these errors are distinguishable from errors made during any other part of the medication process such as prescribing, transcribing or dispensing. The second aim of the project was to conceptualize this knowledge within an instrument which could effectively measure the presence (and hence absence) of safety in a secondary care environment. In order to be useful, the instrument was developed with the following criteria in mind;

1. **Objectivity**. Achieved by making items tangible and reducing subjective responding.
2. **Predictive validity**. Ability to proactively predict future medication errors.
3. **Generaliseability**. Ability to monitor the safety of multiple secondary care environments.
4. **Analysis at ward / unit level.** As with several industrial instruments of this nature, this tool should generate data at the team/ward/unit level rather than at an individual level.

---

[3] Medical Research Council
[4] Economic and Social Research Council
[5] Engineering and Physical Sciences Research Council

It was intended that the tool would take the form of a type of safety audit which could be administered at regular intervals (e.g. once every 6-12 months) and ideally to any qualified nurse on any ward within secondary care. In this way, ongoing organisational safety could be effectively monitored over time. Although similar instruments are available within other high-risk industries (e.g. Tripod Delta; Hudson *et al.,* 1994), there is to date no evidence that such instruments are generalisable beyond the industrial environments in which they were developed, to the management of safety in health care. Furthermore, there is little empirical evidence supporting their ability to predict future errors. For the proposed health care safety management tool to be considered a *proactive* method of predicting future medication administration errors, it should show predictive (criterion) validity with an objective measure of errors. However, due to problems associated with the current NHS incident reporting system (e.g. underreporting and reporting bias; Henriksen & Kaplan, 2003; Lilford, Mohammed, Braunholtz & Hofer, 2003), a third aim of this research was to develop an independent measure of drug administration error with which to validate the organisational safety tool. In consultation with the Director of Research at the collaborating institute (Bradford Teaching Hospitals NHS Foundation Trust), it was agreed the research should be conducted within general medicine to keep the research as generic as possible with the potential for the tool to be used in multiple clinical fields. It was hypothesized that conducting the research in a more specialised area of health care (e.g. obstetrics/paediatrics) may have resulted in the development of a tool which was not as generalisable to other areas of hospital care.

1.6. Thesis plan

This thesis will address four main research questions;

1. What are the latent causes of drug administration errors in secondary care? (study 1)
2. How do these latent causes manifest themselves on a day-to-day basis in the workplace? (study 2)
3. What types of drug administration errors occur in secondary care? (study 3)
4. Can the latent causes identified be empirically shown to predict drug administration errors? (study 4)

The thesis is arranged as follows;

Chapter 2 – Review: Organisational perspectives of error: The systems approach

This chapter discusses organisational error theory; in particular Reason's Swiss cheese model and framework of understanding organisational accidents (Reason, 1990, 1997) with a view to exploring identified latent causes of human error. Research conducted in a variety of research fields is also presented to lend support to Reason's framework.

Chapter 3 – Review: Organisational and workplace interventions to reduce medication error

This chapter presents a systematic review of research which has implemented systems interventions in secondary care to reduce medication errors in an attempt to explore the role of these latent variables in the occurrence of error. Findings are discussed in terms of theoretical approach, methodological designs, interventions and outcome measures.

Chapter 4 – Study 1: An exploratory study to identify latent causes of drug administration errors

This chapter describes study 1 which comprised interviews, observations and a retrospective review of clinical incidents to explore the latent causes of medication administration errors (MAEs). Methodological disparities and concerns of each of the three comparative sources of data are discussed with a view to combining the data in a meaningful way. Findings are presented and one of the ten extracted causal 'themes' is discussed in full to give a general idea of how *all* themes were identified as latent preconditions of MAEs. Finally, findings from a test of the inter-rater reliability of the ten themes are also presented in this chapter.

Chapter 5 – Study 2: Development of the Organisational Safety Questionnaire (OSQ)

Using the qualitative data obtained in study 1, this study explored the ways in which each of the ten latent failures manifest on a ward level. For example, how do poor communication or training

issues look like on a particularly 'bad' working day. This chapter describes thematic content analysis of qualitative data collected in study 1 to generate 'safety indicators' based on these manifestations of poor safety. Findings from a pilot study to test the face validity of these indicators are also presented.

Chapter 6 – Study 3: Development of the Drug Round Behaviour Questionnaire (DRBQ)

This chapter describes several disadvantages to using NHS formally reported incidents as an outcome measure against which to test the predictive validity of the OSQ, followed by a discussion of other ways in which medical errors have been categorized or classified into comprehensive taxonomies. Study 3 is then described in which an independent measure of medication administration errors was developed (DRBQ). Using the qualitative data obtained in study 1, this study explored the types of MAEs occurring in secondary care using a well-known taxonomy of medication errors as a template for analysis. Finally, this chapter will describe a pilot study conducted to improve the structure and face validity of the tool.

Chapter 7 – Study 4: Validation of the OSQ

This chapter presents the final research study which involved hospital-wide administration of the OSQ, along with the DRBQ and an independent behavioural measure of safety climate at the ward level (Vogus & Sutcliffe, 2007) against which to test the internal consistency, predictive and concurrent validity of the tool.

Chapter 8 – Conclusions

The final chapter of this thesis draws research findings together in order to discuss overall findings, methodological limitations of this type of applied research, practical recommendations for future research and overall conclusions.

CHAPTER 2
ORGANISATIONAL PERSPECTIVES OF ERROR: THE SYSTEMS APPROACH

2.1. Introduction

Large scale accidents have often involved investigation into the causes of human error. However, over the last 50 years research - particularly in the field of cognitive psychology, has attempted to study error in less catastrophic events with the aim of differentiating 'correct performance' from 'human fallibility' (Reason, 1990). The result has been a broad spectrum of error explanations from individual perspectives which account for cognitive and personality characteristics leading people to commit errors (Croskerry, 2002) to engineering perspectives which recommend humans should be 'engineered out' of a system by increasing automation (Vincent, 2006). Psychological theories of human error in particular have evolved over time from traditional cognitive approaches (Bartlett, 1932; Knowles, 1963) to more recent organisational models of explanation (Reason, 1997). This chapter will focus predominantly on one particular error causation framework; the *Swiss cheese* model, proposed by James Reason (Reason, 1990, 1997). This conceptual model is one of the most widely accepted explanations and taxonomies of human error and Reason's research on organisational accidents provides the main foundation of the current research. Reason's earlier psychological error taxonomy; Generic Error Modelling System (GEMS; Reason, 1990) will be discussed in chapter 6 with a view to informing the development of a second tool to measure drug administration errors.

2.2. The 'Swiss cheese' model

Reason hypothesized that errors are best understood as an interaction between unsafe acts known as *active errors* (e.g. slips/lapses, mistakes and violations) and systemic organisational weaknesses

known as *latent failures* (Reason, 1990). Reason argues that the majority of adverse events in safety critical industries will involve a combination of these two sets of variables. Using Swiss cheese as a useful metaphor, Reason argues that all organisations have layers of defensive barriers and safeguards (in the form of alarms, physical barriers, people or rules and procedures) which protect it from hazards. Within each defensive layer are 'holes' (similar to those found in Swiss cheese) which are continuously opening, closing and moving location. Figure 2.2 illustrates this accident trajectory.

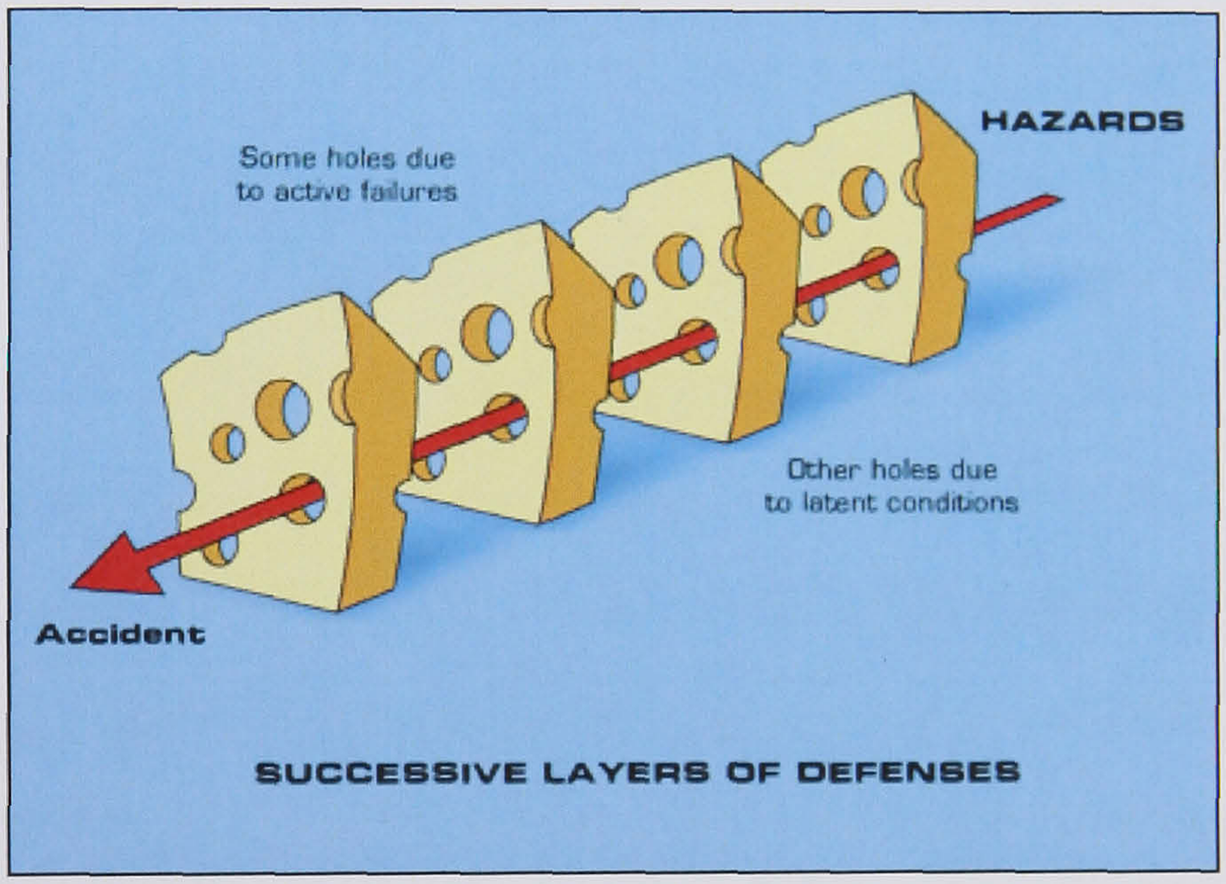


Figure 2.2: Reason's *Swiss cheese* model of accidents

Reason argues that holes in defensive layers can occur as a result of active errors and others as a result of latent conditions. Opportunities for adverse events arise when holes in multiple layers of defence line up and create an accident trajectory.

2.2.1. Active errors and error management

Reason proposes that active errors occur at the 'sharp-end' of the workforce and often temporally or proximally precede a negative outcome. These errors would involve individuals in direct contact with the system such as nurses, pilots, engineers and police officers for example (Reason, 1997).

For example, a nurse who administers the wrong medication to a patient or a train driver who mistakenly drives through a stop signal would be considered to have committed active errors at the sharp-end of the organisation.

In terms of error management, organisations that focus on specific unsafe acts often look no further for potential causes of an adverse event once they have identified those individuals closest (in time or space) to the outcome. Corresponding error management countermeasures, particularly in health care tend to focus on blame, retraining, discipline or other responses targeted at specific individuals ('blame-and-retrain-cycle' - Reason, 1997; Walton, 2004; Battles, 2006). Such solutions are often employed in the hope this will lead to improved vigilance in the person deemed 'responsible' for the error and act as an error deterrent to others. This individualistic approach to error management has long since been the predominating stance taken in health care and is likely to be driven in part by a psychological phenomenon known as Fundamental Attribution Error (FAE; Fiske & Taylor, 1978). This theory asserts that individuals commonly attribute internal personality factors to the actions of others yet offer external situation factors to explain their own behaviour. In this way, individuals are more likely to explain errors they make as being due to situational variables (e.g. being placed under extreme pressure) and errors made by someone else as being due to negative aspects of their personality (e.g. recklessness) or cognitive ability (e.g. lack of vigilance). This attribution tendency is frequent in the media which often implies that patients have been harmed by carelessness or intent (Wieman & Wieman, 2004). Headlines such as '[k]illed by doctors' blunder…', (The Sun, Friday 20th April, 2000); '[d]eath jab doc behind bars…' (The Sun, Tuesday 23rd September, 2003) are relatively commonplace within the tabloid press.

2.2.2. Latent conditions and error management

Reason suggests that '[a]ctive failures are like mosquitoes. They can be swatted one by one, but they will still keep coming. The best remedies are to create more effective defences and to drain the

swamps in which they breed. The swamps, in this case, are the ever present latent conditions' (Reason, 2000, pp.769). These latent conditions are defined as inevitable 'resident pathogens' within an organisational system which arise from decisions made by top-level management, designers and procedure writers. Such failures include poor management or supervision, lack of adequate job training or poorly designed tools and equipment for example. These latent failures lay dormant as 'gaps' in the organisation's defences and promote local error-producing conditions due to the system's weakened state. When these latent failures combine with inevitable human tendencies to make active errors at the sharp-end of the system, the end result is often an adverse event. This process is referred to as the *accident trajectory* (Reason, 1997, pp. 12).

In terms of error management, Reason argues that whilst active errors are inevitable and unpredictable, latent conditions on the other hand can be identified before an adverse event occurs. If holes due to latent failures can be proactively identified and effectively remedied or alternatively systems defence layers can be added, the problem could be averted before a negative outcome ensues. Figure 2.2.2 illustrates this concept. Reason argues that when errors occur '…the question should not be "who is at fault?" but "why did our defences fail?"'.

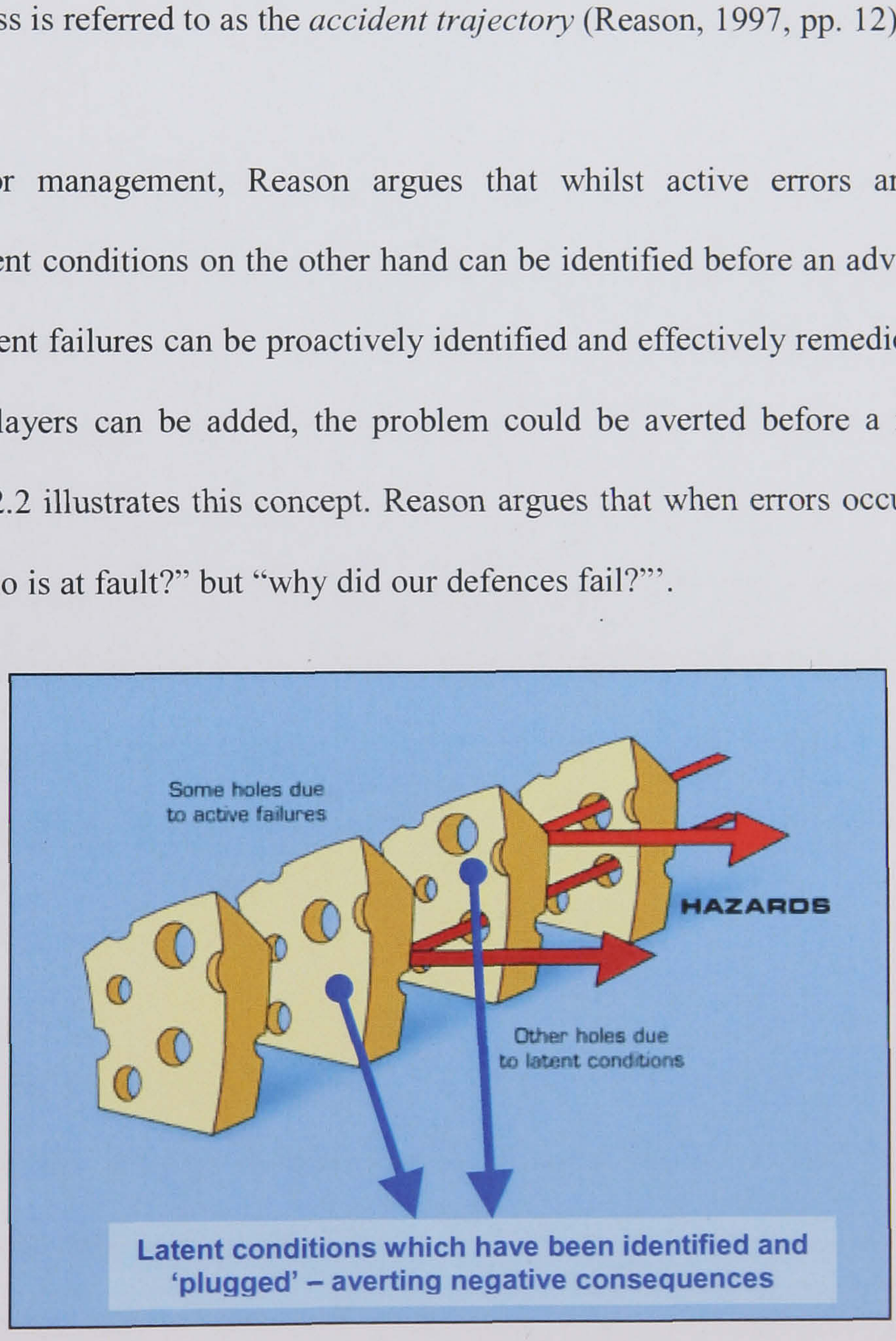


Figure 2.2.2: Identifying latent conditions and averting adverse consequences

2.3. The 'Organisational Accident'

As an extension of his *Swiss cheese* model, Reason formulated a framework of understanding the relative contribution of various latent failures in the occurrence of accidents in high-risk organisations (Reason, 1997). This *organisational accident theory* attempts to explain the development of accidents from unsafe acts, through local workplace factors to organisational factors at the 'blunt-end' of the system. For instance, after incidents or accidents have occurred, retrospective incident investigation can be conducted in a bottom-up approach from unsafe acts upwards through error-provoking local working conditions and on to the organisational determinants of the accident, thus establishing top-down causation of future accidents.

It is important to emphasise that Reason's organisational accident model was formulated within his extensive experience of industry and was developed particularly from the examination of three case studies to identify common features. These case studies were the *Barings Bank Collapse* (the loss of £869 million due to *rogue trader* Nick Leeson between 1992 & 1995); the *Nakina Derailment* (1992 train crash in Ontario, Canada) and finally the analysis of a near-miss crash of a Boeing 737 in 1995. These in-depth analyses of large-scale industrial accidents (and one potentially fatal near-miss) lead to the development of general principles of accident trajectories from latent conditions to active failures common to each accident. Figure 2.3 summarises and illustrates Reason's theory. Green arrows indicate the direction of error causality and the blue and red arrows indicate potential investigative steps.

Reason suggests that latent conditions can be considered along a causal sequence starting with organisational processes such as forecasting, budgeting, allocation of resources and managing which operate at the blunt-end of the system. The effects of these decisions are filtered down throughout the organisation to individual workplaces (e.g. hospital wards, flight decks, control rooms etc.) where they manifest as defective local working conditions. Inadequate or faulty

equipment, poor worker-supervisor relationships, insufficient training, excessive workloads and poor team communication are some of the local working conditions resulting from defective organisational processes. These variables combine with inevitable human tendencies to commit active errors at the sharp-end of the system and increase the potential for adverse events.

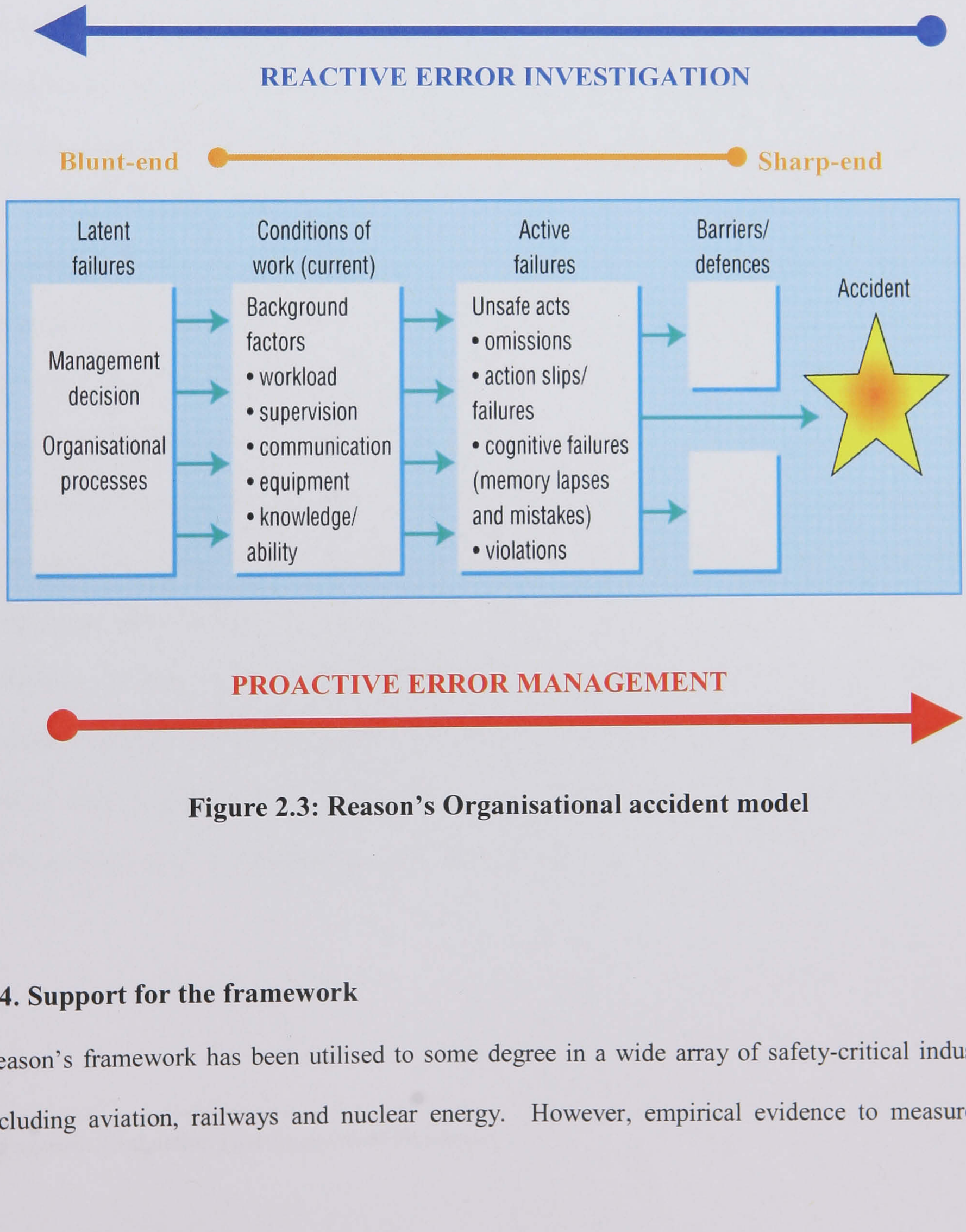


Figure 2.3: Reason's Organisational accident model

2.4. Support for the framework

Reason's framework has been utilised to some degree in a wide array of safety-critical industries including aviation, railways and nuclear energy. However, empirical evidence to measure the

validity or reliability of the framework (i.e. whether latent failures consistently *predict* future safety) is distinctly lacking. Nevertheless, in view of the three main aims of this research: exploring latent failures specific to health care, developing ways to measure these failures and developing a tool to measure error behaviour with which to validate these failures, particularly relevant research will be outlined here.

2.4.1. Oil and Gas production

Of particular importance to the current research was a study of latent failures specific to the oil industry conducted by a team of researchers from yniversities in Manchester and Leiden (Reason *et al.,* 1988; Groeneweg, 1992; Hudson *et al.,* 1994). The study was commissioned by the Royal Dutch/Shell Group to provide them with a 'global indication' of their company's 'safety health' since millions of dollars were being invested in behaviour modification techniques with little effect on the rate of accidents (Hudson *et al.,* 1994). Since earlier error management attempts in the Shell group had consisted largely of training and disciplinary measures aimed at unsafe acts, the aim of this research was to examine another dimension - latent conditions. To this end, the research group developed the *Tripod* model; derived from the relationship between the three components of accident causation identified by Reason; latent conditions, unsafe acts (active errors) and accidents, illustrated in Figure 2.4.1.a. To measure and control latent failures, the team developed two error management tools; one to *proactively* identify latent failures in the organisation (Tripod-DELTA[6]) and the other (Tripod-BETA) to *retrospectively* investigate accidents. This section will focus on the tool most relevant to the research proposed - Tripod-DELTA.

---

[6] Diagnostic EvaLuation Tool for Accident Prevention

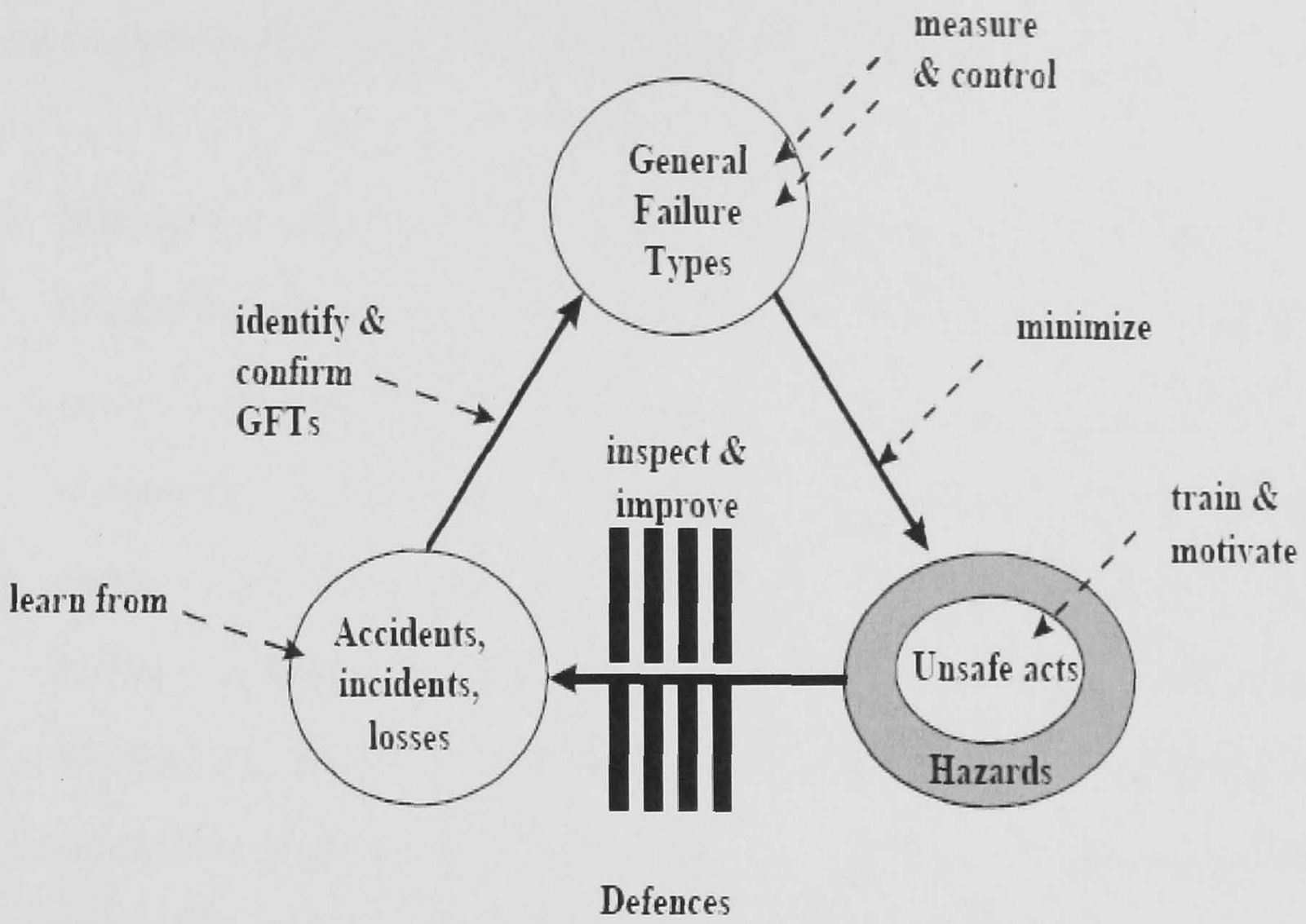


Figure 2.4.1.a: The three 'feet' of Tripod Delta: General failure types, unsafe acts and accidents

Eleven distinct latent failures were identified from extensive analysis of accident scenarios, audit reports and one study conducted on several offshore oil rig platforms, as reflecting high-level underlying organisational failures which would lead to an infinite number of unsafe acts (Groeneweg, 1992)[7]. These failures were termed *General Failure Types* [8] (GFTs; Reason *et al.,* 1989; Hudson *et al.,* 1994) and are outlined overleaf (adapted from Reason, 1997, pp. 134-135 & Gordon *et al.,* 2000). The research group proposed that the greater extent to which one or more

[7] Research later showed these constructs to be generalisable to a wide range of oil exploration-related fields, including seismic activity, offshore drilling and maintenance and gas production (Groeneweg, 1992).
[8] GFT's are now referred to by the Tripod Group as Basic Risk Factors (BRF's) due to crossover with the abbreviation of GFT in Holland which refers to biologically degradable waste disposal.

GFTs are present in any industrial organisation, the less stable their corresponding safety will be and the more likely accidents will occur (Groeneweg, 1992; Hudson *et al.,* 1994).

1. **Hardware** – Failures due to the inadequate quality of materials, tools or equipment, non-availability of hardware and failures due to ageing. This also includes policies and responsibilities for hardware purchase, quality of stock system, quality of supply, theft or loss of equipment, short-term renting, compliance to spec and non-standard use of equipment.
2. **Design** – Deficiencies in the layout or design of facilities, equipment or tools that leads to misuse or unsafe acts, increasing the likelihood of particular types of errors and violations. For example, the failure of designers to provide external guidance (knowledge gulf), objects designed as opaque in their inner workings or range of safe actions (execution gulf) and a failure of items to provide feedback to user (evaluation gulf).
3. **Maintenance management** – Failures in the systems for ensuring technical integrity of facilities, plant equipment and tools. This GFT relates to the management of maintenance NOT the execution of maintenance activities. For example, was required maintenance work carried out in timely fashion?
4. **Procedures** – Failures related to the quality, accuracy, relevance, availability and workability of procedures.
5. **Error enforcing conditions** – Factors such as time pressures, changes in work patterns, physical working conditions which act on the individual or in the workplace and encourage the performance of unsafe acts. Either 'error-producing conditions' or 'violation-promoting conditions'.
6. **Housekeeping** – Organisational tolerance of deficiencies in conditions of untidiness and cleanliness of facilities and work spaces or in the provision of adequate resources for cleaning and waste removal which increase the chances of unsafe acts.

7. **Incompatible goals** – Failure to manage conflict: between organisational goals (safety and production); between formal rules (company written procedures and the rules generated informally by a work group); between the demands of individuals, tasks and their personal preoccupation or distractions.
8. **Organisation** – Deficiencies in either the structure of a company or the way it conducts its business that allow safety responsibilities to become ill-defined and warning signs to be overlooked.
9. **Communication** – Failures in transmitting information that is necessary for the safe and effective functioning of the organisation to the appropriate recipients in a clear and unambiguous or intelligible form. Transmission failures indicate that the necessary communication channels do not exist (system failures), the necessary information is not transmitted (message failures) or the necessary information is misinterpreted or delayed (reception failures).
10. **Training** – Deficiencies in the system for providing the necessary awareness (of the hazardous conditions of the workplace), knowledge or skill to an individual or individuals in the organisation. This GFT includes provision of on-the-job coaching mentors and supervisors as well as formal courses, provision of training relative to operations and appropriate task analyses.
11. **Defences** – Failures in the systems, facilities and equipment for control or containment of hazards or for the mitigation of the consequences of either human or component failures. These failures comprise: detection/alarm; control and interim recovery; personnel protection/containment, escape and rescue.

Of particular relevance to the current research aims; researchers developed a comprehensive checklist of safety indicators which could proactively measure both the presence of each GFT and the cumulative holistic safety of the localised workplaces and the organisation as a whole.

Indicators were generated during several studies which included detailed observations of employees during high-risk activities followed by focus group sessions to discuss observed practices (for full description of these studies see Groeneweg, 1992, pp. 147-205). Safety indicators were developed for each GFT to represent objective (not subject to interpretation) directly observable (tangible) indicators of failure to which the answer would be 'yes' or 'no'. Pools of indicators for each GFT was then compiled and subject to various face validity tests with senior and second-in-command managers (Groeneweg, 1992). Table 2.4.1 provides examples of generated indicators.

Table 2.4.1: Examples of safety indicators used in Tripod-DELTA proactive tool

General Failure Type	Item
Training	Are there any requests for training been made the last month that have not been granted?
Hardware	Is a high pressure water cleaning system available?
Defences	Is the evacuation time to a hospital in cases of serious injury shorter than 2 hours?
Procedures	Have people who were involved in making one of the last three procedures you received visited your rig?
Communication	Have you been invited to the Coast in the last two months to give an update of the situation on the rig?

The safety indicators (usually 20 for each GFT) deemed to be the best representations of each GFT were then used to compile a final safety management checklist, and administered to relevant employees (those responsible for execution of that activity). A GFT *safety profile* was generated by summing responses within each GFT and developing an appropriate action plan for remediation of

the three most problematic GFTs. Figure 2.4.1.b depicts a systems safety profile comparing the safety of two organisations on the same activity.

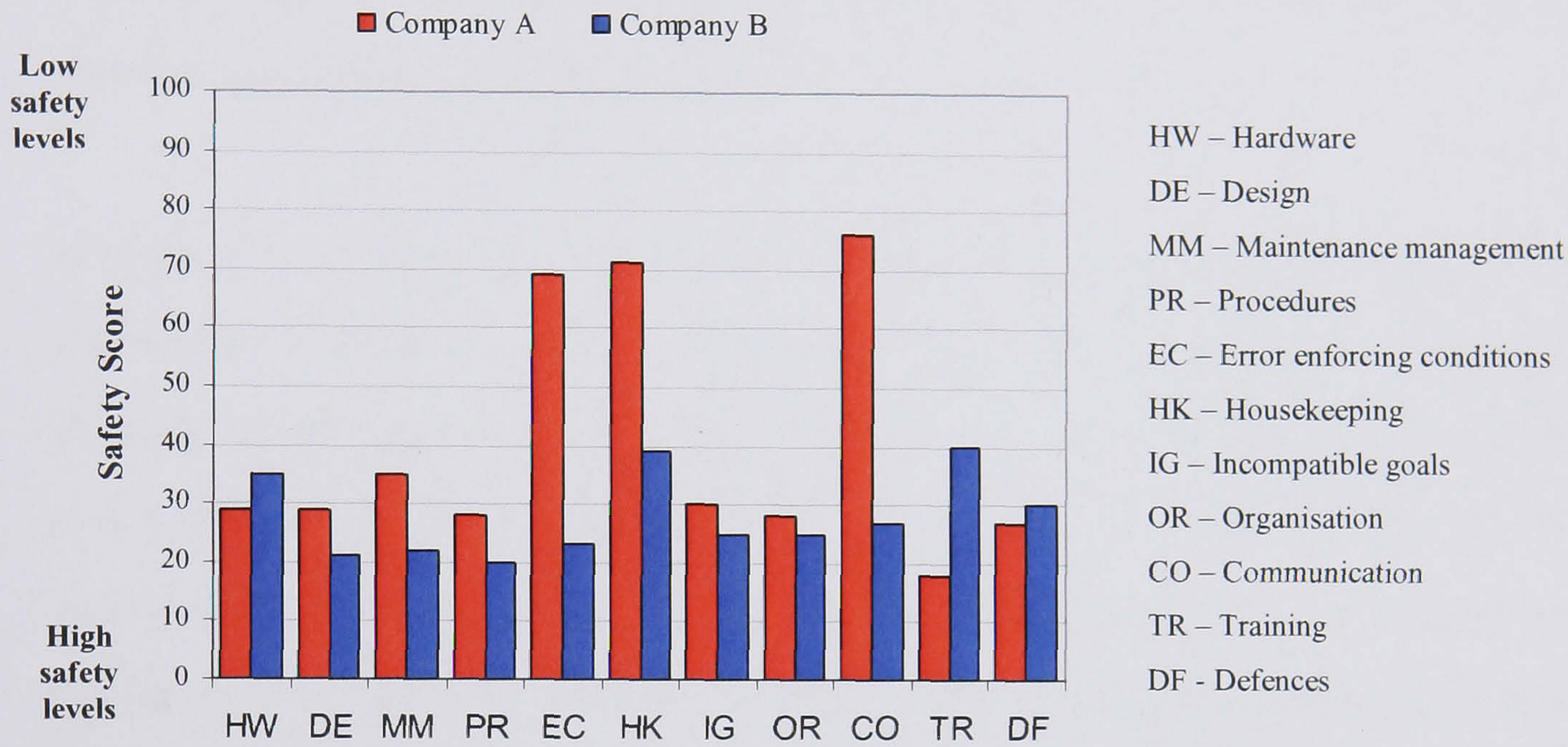


Figure 2.4.1.b: An example of the system safety profile comparing two companies

As this figure illustrates, the 11 GFTs appear along the horizontal axis, whilst the safety score for each GFT is plotted against the vertical; highest scores indicating those GFTs most in need of remediation (since the safety score is the total number of indicators scored in the 'concern' direction). In company A the scores for GFTs overall are generally higher and more erratic compared with the safety levels of company B which are reasonably constant across all 11 GFTs. A clear pattern emerges from company As data which shows the three GFTs with the poorest safety levels and thus the most cause for concern are housekeeping, communication and error enforcing conditions. By comparison, scores for each GFT within company B are more consistent. The three

GFTs with the highest scores for company B are housekeeping, training and hardware. However, these scores are not significantly higher than scores for other 8 GFTs measured, nor are they higher than GFT scores measured in company A. For this reason, the Tripod project has since developed benchmarking tools which aim to measure safety thresholds for particular industries in order to provide individual companies with a benchmark profile with which to make industry comparisons (see www.tripodsolutions.net for further details).

In terms of the concurrent validity of the Tripod-DELTA tool, the Tripod research team administered an independent questionnaire which asked employees (who had not previously completed the checklist but were involved in the same task) to rate on a five point scale from *very poor* to *very good* to what extent each GFT had an impact upon their work (Groeneweg, 1992). For instance, one question asked employees to rate the quality of the design of the platform they were working on. Statistical analysis revealed very high significant correlations between scores on the safety checklist and the validation questionnaires. In a study of the predictive validity of the 11 GFTs, two researchers blind to checklist items or the outcomes of the validating questionnaire reviewed 10 reported incidents for the presence of each GFT in their causation. Their results revealed high and significant correlations between the researchers recorded GFT frequencies in reported incidents and both the safety checklist and validating questionnaire.

Since its conception, the Tripod-DELTA safety management tool has been employed in a number of similar industrial settings as a proactive identification of systems failings likely to result in future errors, including chemical engineering (AKZO Nobel, DSM and Shell Chemicals), the Nuclear Energy Authority in the U.K. and oil and gas production on the North Sea (Groeneweg & Roggeveen, 1998). However, it is worth mentioning that Tripod Delta is rarely (if ever) administered with outcome measures such as error frequency or other quality measures against which to test its predictive value as a safety management tool. One possible explanation for this is

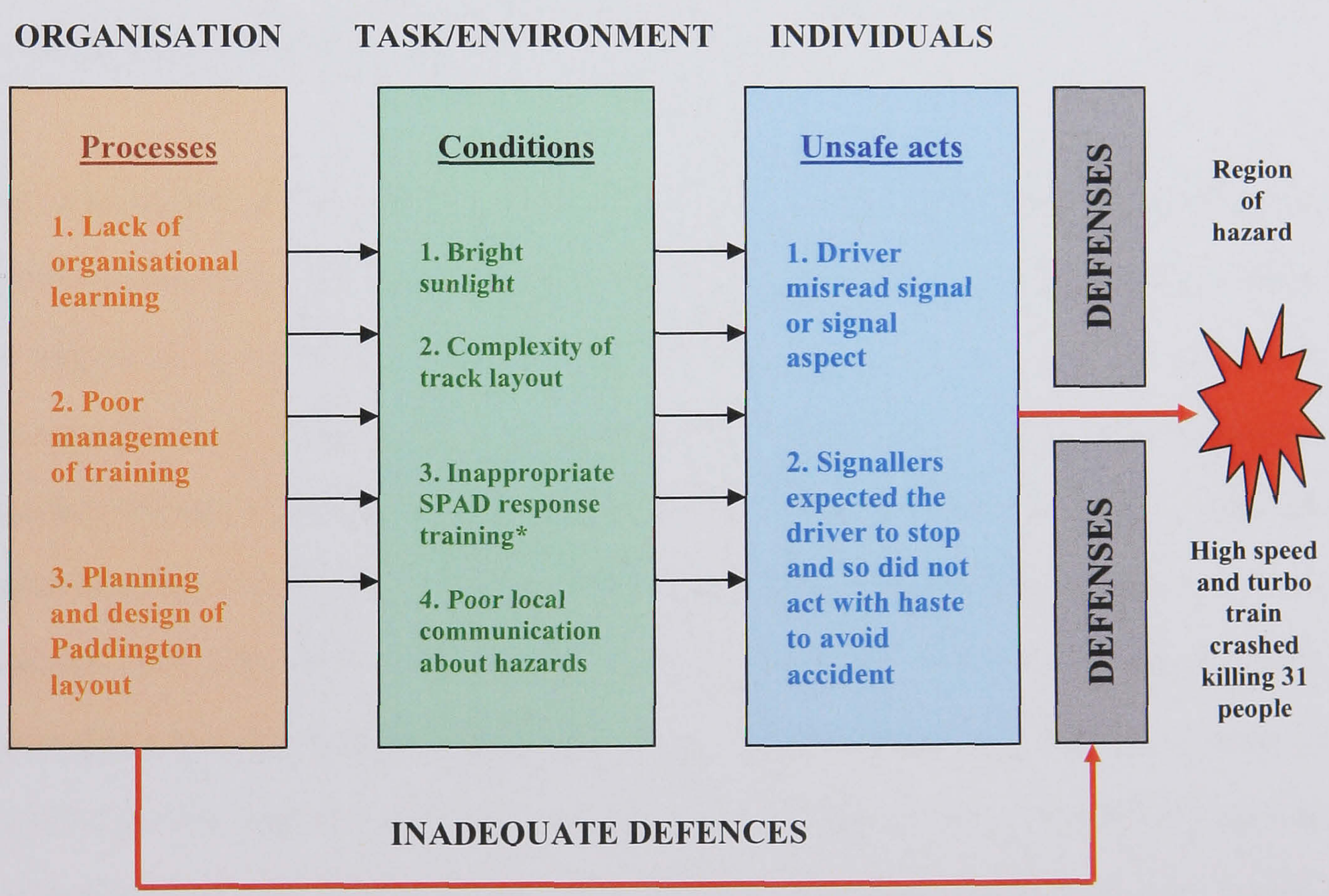
the extremely low occurrence of errors in these industries making predictive validity difficult. Companies such as Motorola and General Electric have popularized the concept of achieving a *Six Sigma* level of safety which equates to only 3 errors occurring in every one million activities (Spencer, 2000). Perrow argues that the reason such industries report such low rates of error is due to the existence of 'correctional defences' built into the system over a number of years which avert the consequences of unsafe acts (Perrow, 1984). Without the existence of reported errors therefore, it is difficult to state with any confidence that Tripod-DELTA tool and the GFTs subsumed within it *predicts* errors. However, the *absence* of rising numbers of adverse incidents goes a little way towards promoting the tools efficacy.

2.4.2. Railways

On October 5th 1999, two trains collided on a main line near Ladbroke grove in the vicinity of London Paddington Station. This collision provoked an extensive multi-disciplinary investigation to identify the contributory factors leading up to the fatal accident, conducted by the UK Health and Safety Executive (HSE, 2002), public enquiries (Cullen, 2000a, 2000b) and human factors experts (Lawton & Ward, 2004). Focussing on the latter investigation, Lawton and Ward were invited to retrospectively review witness interviews and transcripts with a view to conducting systemic crash analysis to identify active errors, local working conditions, situational and task factors, inadequate defences and latent organisational failures. Using the organisational accident model proposed by Reason as a basis for understanding the interplay between these variables, the authors identified multiple contributory factors, summarised in Figure 2.4.2.

The main intention of acknowledging Lawton and Ward's study is to emphasise the usefulness of Reason's model as a framework for understanding this relationship between active errors, local and latent conditions operating beyond the industry in which the model was originally developed. This

human factors approach to retrospective accident investigation heralded a move away from the more traditional view of individual responsibility (i.e. that of the train driver alone). Lawton and Ward acknowledge that error management strategies focussing on the role of one individual act of passing a stop signal at danger would represent an ignorance of 'system pathogens' which would likely result in future catastrophes. In terms of remediation, authors note in this article that at the time of writing a significant number of recommendations made by this analysis, the public inquiry and a previous crash analysis of the Southall crash in 1997 had been implemented.



* Signals passed at danger

Figure 2.4.2: A summary of organisational, local and individual failures in Ladbroke Grove accident (adapted from Lawton & Ward, 2004)

Although Lawton and Ward's analysis applied Reason's organisational accident framework to the retrospective analysis of a railway incident, the latent failures they proposed had preceded the accident were not those identified as most likely causes (e.g. their 11 GFTs) by the Tripod research group. Nevertheless, more than 20 years after its conception, the Tripod group still maintain that the 11 GFTs they specify as the most likely to result in human error can be found in *any* type of organisation (www.tripodsolutions.net). The Tripod Group are now an international risk management company which offer worldwide consultancy services and products which include their proactive (Tripod-DELTA) and reactive (Tripod-BETA) instruments. They advocate on their website that the 11 GFTs which were originally found in the 1992 Shell study, are the 'chief [latent] causes leading to all preconditions which generate unsafe acts…in a variety of industries including the health sector, transport services, IT and banking services'[9].

This is particularly pertinent for the current research project which aims to identify latent failures specific to secondary health care. Groeneweg (2002) suggests that in order to proactively measure the presence of each GFT, an organisation needs only to generate the safety indicators relevant for that particular field or activity. However, research would suggest that the latent failures which are associated with errors in the oil industry may not be as prevalent (or present at all) in other high-risk industries which may rely less on highly automated systems of work and more on hands-on person-to-person work (IoM, 2004). This all depends on whether health care can be considered directly comparable with other high-risk industries. Vincent (2006) suggests that while lessons can be learned from other types of organisation on the causes of human error, research should exercise caution before drawing too many parallels between these environments and health care. Reason concurs and suggests that within non-health care industries there is a large emphasis on highly routine tasks and monitoring automated activities (Reason, 1997). By comparison, the field of health care is considerably less technologically driven and involves a workforce who are

[9] This author could find no published evidence that the Tripod-DELTA tool had been tested in the health sector or any other sector outside the oil and gas production industry.

predominantly involved with 'abnormal person-specific situations' which are inherently more 'error-provocative' and 'risk-laden' (IoM, 2004, pp. 61). Reason also refers to the uncertainty of knowledge faced by health care staff, methods of incident investigation (often very public in health care leading to less opportunities for organisational learning), and the mode of delivering service ('few-to-many' vs. 'few-to-one') as major discrepancies between non health care and health care environments (Reason, 2003, as cited in IOM, 2004, pp. 61-63).

To account for this environmental disparity, Reason has developed two further instruments which proactively measure both local and latent failures in railway (*Review*; Reason, 1993) and airline organisations (*MESH*[10]; Reason, 1995). Although based on the same framework of organisational accidents (Reason, 1997), these tools measure qualitatively different failures relevant to their respective organisations, such as planning, staffing and rostering (*Review*), organisational structure, people management and commercial pressures (*MESH*). Although the premise of accident causation is the same (e.g. the presence of latent organisational failures predict the likelihood of future errors), the failures themselves are specific to that particular environment. Therefore, for the purposes of the current research, it is important to establish which latent failures are the most prevalent and necessarily predictive of errors in the field of medicine.

2.4.3. Medicine

While several proactive and reactive safety measures exist in other high-risk industries (e.g. *Tripod-DELTA* and *Tripod-BETA* in oil production, *Review* in railway operations and *MESH* and *MEDA*[11] used in aircraft maintenance); there is currently no comparable measure (proactive or reactive) of latent failures available in health care. The majority of what little medical literature there is which has applied Reason's organisational accident model has largely involved retrospective analysis of clinical incidents (Vincent *et al.,* 2000; Vincent, 2004; Meurier, 2000; Molloy & O' Boyle, 2005).

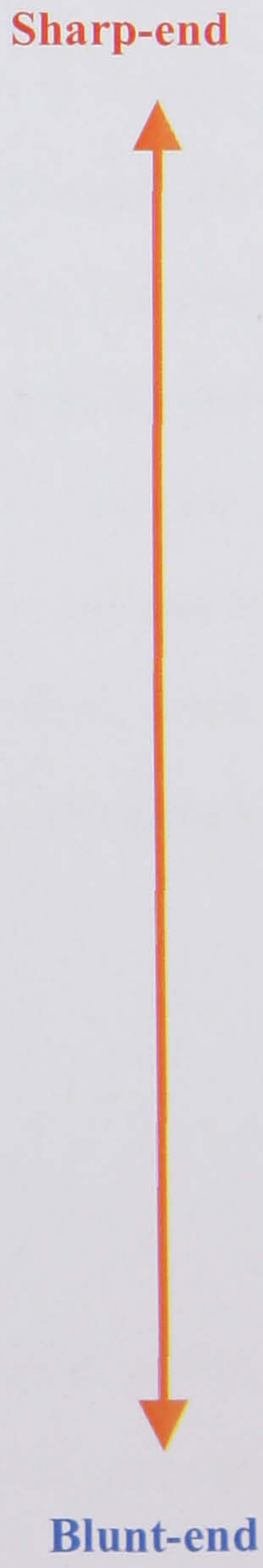
---

[10] Managing Engineering Safety Health
[11] Maintenance Error Decision Aid

One particularly prominent example of this research is the collaborative work of Charles Vincent and colleagues at the Imperial College, London. This collaborative research group comprises members of the Clinical Risk Unit (UCL, London) and members of the Association of Litigation and Risk Management (ALARM). As an extension of Reason's organisational accident model, the group proposed an adapted conceptual framework for the retrospective investigation and analysis of clinical incidents in health care (Vincent, Taylor-Adams & Stanhope, 1998). A summary of this framework is presented in Table 2.4.3.

Table 2.4.3: Summary of Vincent's seven-level framework

Factor types	Contributory influencing factors
Patient factors	Condition (complexity and seriousness)
	Language and communication
	Personality and social factors
Task and technology factors	Task design and clarity of structure
	Availability and use of protocols
	Availability and accuracy of test results
	Decision-making aids
Individual (staff) factors	Knowledge and skills
	Competence
	Physical and mental health
Team factors	Verbal communication
	Written communication
	Supervision and seeking help
	Team leadership
Work environment factors	Staffing levels and skill mix
	Workload and shift patterns
	Design, availability and maintenance of equipment
	Administrative and managerial support
	Physical environment
Organisational and management factors	Financial resources and constraints
	Organisational structure
	Policy, standards and goals
	Safety culture and priorities
Institutional context factors	Economic and regulatory context
	National health service executive
	Links with external organisations



This hierarchical framework comprises levels of safety operating across seven levels of the organisation from 'sharp-end' patient variables, through task and other workplace variables to 'blunt-end' management and regulatory body factors. These variables were identified using a combination of patient record review, staff interviews and a review of human factors variables (psychological and organisational) identified by other major frameworks outside health care (e.g. socio-technical pyramid; Hurst & Radcliffe, 1994). Vincent suggests that in any given clinical situation, the condition of the patient will have the most significant influence on outcome; a feature of health care which distinguishes it from other high-risk industries (Vincent, 2006).

Although Vincent's framework constitutes a broad and comprehensive conceptual taxonomy for investigating potential contributory factors implicated in medical errors, current government recommendations (e.g. DoH, 2001; IoM, 2004) advise that health care organisations should not rely solely on retrospective analysis. The Institute of Medicine in particular suggest 'reactive investigations must be supplemented by proactive error reduction activities'. Indeed, Reason suggests latent holes in each defensive layer are in 'continual flux' (Reason, 1998). Therefore, while reactive investigation can provide useful information on how specific errors have occurred it cannot mitigate against *all* possible future unsafe acts. Proactively identifying latent failures which consistently predict multiple unsafe acts or at least promote local triggering conditions would be a useful addition to an organisational error management toolkit.

2.5. Summary

The need for safety in high-risk industries is well established. In installations such as nuclear power, chemical engineering or aviation, what could be deemed a relatively small unsafe act could potentially cost thousands of lives and have dire financial consequences (e.g. Chernobyl reactor disaster, 1987; Three Mile Island nuclear plant disaster, 1979; Clapham junction railway accident,

1988; see Reason, 1997, pp. 86-90 for detailed case histories). By comparison, the level of error occurring on a daily basis within health care far exceeds that observed in other high-risk industries. Leape, Bates and Cullen (1995) estimate that patients in intensive care units could each experience up to 2 errors every day that they are hospitalized and 20 per cent of these errors could be of a potentially serious nature or even result in fatality. In terms of medication errors, Leape suggests that these may occur in as few as 1 in 1,000 to 1 in 10,000 patient admissions in the United States, but adds that even a 0.01 per cent failure rate would not be tolerated in other high-risk industries such as aviation or engineering (Leape *et al.,* 1995). Research suggests that human error is both inevitable and resistant to change and acknowledges that as a result individualistic error management strategies have been and will continue to be largely ineffective (Temple, 2004). Reason maintains that this 'person-approach' is still the dominant tradition in health care and suggests that 'continued adherence to this approach is likely to thwart the development of safer healthcare institutions' (Reason, 2000, pp. 769). He suggests that by considering only factors at the level of the individual as likely causes of error, opportunities for organisational learning are being lost and subsequent remediation attempts to reduce errors will continue to be unsuccessful.

While there has been some application of Reason's organisational accident framework in several high-risk industries to retrospectively identify latent failures, there has been no published evidence to support their association with future errors. In order to test the relationship between latent failures and active errors and therefore validate Reason's framework, it is vital that latent failures can be measured as proactively predicting future errors. There is to date no empirical evidence that any identified latent failure *predicts* errors or any other outcome measure in health care or any other high-risk industry. Chapter 3 will present health care research which has employed latent systems interventions to reduce medication errors. It is anticipated that this review will go some way towards testing the applicability of Reason's framework in health care research.

2.6. Research plan

This thesis will present findings from a systematic review of interventions which have been implemented within secondary care to target systems (e.g. organisational and workplace) variables to reduce medication errors. It is anticipated that this review will go some way towards informing subsequent exploratory studies to investigate the latent causes of medication errors. Using Reason's organisational accident model as a framework the research aims to explore the latent failures which exist in secondary care from a 'bottom-up' perspective in a similar way to the Ladbroke grove analysis (Lawton & Ward, 2004). Using similar methodologies as those employed in the Tripod research project, interviews with both senior management at the 'blunt-end' of the NHS organisation and also 'sharp-end' health care professionals will be conducted alongside incident report reviews and clinical observations. The research will then investigate whether tangible safety indicators can be generated for each identified latent failure. Since research has emphasised the lack of empirical evidence showing latent failures *cause* subsequent errors (due to low error rates in industry and no suitable proactive measure in health care), the final aim of this research will be to measure the predictive validity of latent failure safety indicators.

CHAPTER 3
SYSTEMS INTERVENTIONS TO REDUCE MEDICATION ERRORS: A SYSTEMATIC REVIEW

3.1. Introduction

Research suggests that medication errors represent at least 20 per cent of all adverse events involving hospitalised patients, many of which are often considered preventable (Leape *et al,* 1991; Bates *et al.,* 1995; O'Hara & Carson, 1995). It has also been suggested that these figures may underestimate the scale of the problem, relying predominantly upon formal incident reporting systems which tend to underestimate the actual number of errors occurring (Wakefield *et al.*, 1999; Brown *et al.,* 2008). In terms of how these statistics relate to patient admissions, Leape reports that medication errors may only occur in as few as 1 in 1,000 to 1 in 10,000 admissions in the United States, but adds that even a 0.01 per cent failure rate would not be tolerated in other safety-critical industries such as aviation or engineering (Leape *et al.,* 1995).

Whilst organisational and other aspects of the working environment are frequently cited latent variables associated with human error in many other high-risk industries (see chapter 2), this chapter explores evidence for the role of latent organisational variables in the occurrence of medication errors in secondary health care. This chapter presents a systematic review of the literature which has implemented systems interventions in secondary care to reduce medication errors. It is proposed that if systems variables *cause* medication errors then interventions which target these variables should in turn reduce error occurrence. The terms 'system', 'organisational' and 'workplace' factors are used interchangeably throughout this chapter. The term *systems variable* is intended to encapsulate *organisational* variables which operate at the 'blunt-end' of the organisational system (e.g. policy development), and *workplace* variables which operate at the 'sharp-end' (e.g. teamwork) (Reason, 1997).

3.2. Literature search and data extraction

Hoff *et al.* (2004) conducted a similar review to investigate the linkages between organisational factors, medical errors and patient safety. Their search strategy generated a list of 16 organisational terms highlighted within the United States health care quality improvement literature (e.g. communication, training, feedback, education, culture etc.). The reviewers then combined each organisational term with the term 'patient safety' and then with the term 'medical error' (32 electronic searches in total). This extracted a total of 2445 articles which, when subject to several selection filters was reduced to 42 empirical articles. Although a rigorous search strategy, one potential disadvantage of this method is that it assumes a *top-down* approach to searching. Generally, employing this method will only extract those articles in which the authors knew they had studied a known organisational factor and reflected this in their terminology. Hoff *et al* acknowledge that their method of searching for articles was developed to reduce the chance that searching would be biased by the reviewers own opinion on what organisational factors are important in medical error. They further acknowledge that wide variability in the use of organisational terminology in the medical literature may have restricted the number of articles extracted. It is possible therefore that Hoff's list of 16 organisational factors used to search for articles generated by the quality improvement literature remains biased by the terminology used to search for articles. This is in part supported by the fact that the current review found several articles which had implemented organisational or systems interventions which were not extracted by Hoff *et al.*

While the current review shared similar broad objectives to the review conducted by Hoff *et al.*, it also differed in three ways. Firstly, the search strategy (bottom-up vs. top-down; as described above) was broadened to search a wider range of databases and used a more exploratory method of searching for organisational or systems terminology. Secondly, Hoff *et al.* included studies which had utilized quality improvement measures as outcome variables (such as litigation costs, job

dissatisfaction, self-reported quality of care), which were not included in the current review. Since the main aim of this research was to explore the role of organisational and workplace factors in medication errors and not quality of care, only those studies which had employed suitable error measures were included. Finally, this review will focus upon errors committed during the administration of medication while Hoff's review explored the causes of medical errors in general (e.g. diagnostic, prescribing, administration etc.).

3.2.1. Search strategy

3.2.1.1. Aims

The aim of this search strategy was to identify studies published from 1990 onwards (when patient safety research started to gain momentum; Vincent, 2006) which had employed *any* organisational and/or workplace intervention to reduce medication administration errors (MAEs). A broad approach to searching the literature was employed using systematic review principles recommended by the Centre for Reviews and Dissemination[12] (CRD, 2001) to extract any studies which were potentially relevant. These guidelines recommend that reviews are undertaken and reported in an unbiased way with specific and explicitly stated search strategies and selection (inclusion and exclusion) criteria (CRD, 2001, pp. 8). Although there is a considerable literature base which has *theorised* organisational and workplace causes of MAE, this review focussed upon studies which had designed and implemented organisational or workplace interventions to reduce medication errors. These studies were considered to be those which, through intervention had empirically tested the role of these variables in error causation and so would be of the most relevance to this research.

---

[12] The Centre for Reviews and Dissemination (CRD) is part of the National Institute for Health Research (NIHR) and is a department of the University of York. CRD, established in 1994, is one of the largest groups in the world engaged exclusively in evidence synthesis in the health field, undertaking systematic reviews to evaluate the research evidence on health and public health questions of national and international importance. CRD produces guidelines for researcher conducting systematic reviews in health care research (http://www.york.ac.uk/inst/crd/index.htm).

3.2.1.2. Review question

CRD guidelines recommend developing an explicit review question before starting the review process. The review question for this review is as follows;

Which organisational and workplace factors have been targeted for intervention to reduce MAEs in secondary care hospitals since 1990?

3.2.1.3. Method

Multiple methods were employed to achieve a 'comprehensive' and 'unbiased' search of the literature (CRD, 2001, Stage II, Phase 3, pp. 4).

i. **Electronic searches -** MEDLINE (1950-2008); British Nursing Index (1985-2008); Cumulative Index to Nursing & Allied Health Literature (1982 -2008); EMBASE (1980 - 2008); Health Management Information Consortium (1983 – 2008) and PsychINFO (1985 – 2008). All databases recommended by the CRD, were accessed via the search engine Ovid.

ii. **Individual journal search -** The journals Quality and Safety in Health Care, Health Services Research, Journal of the American Medical Association and the British Medical Journal were hand searched between 1998 and 2008 for any articles not retrieved by electronic searches. Due to time and journal issue availability constraints, only the last 10 years were searched by hand.

iii. **Review articles** - Review articles and government reports were extracted both by electronic and hand searches which highlighted important intervention studies. These review articles were subsequently searched for any studies which might be relevant to the current review. The following reviews were particularly useful; Hoff *et al.,* (2004); Ioannidis and Lau (2001);

'*Building a safer NHS for patients*' (DoH, 2001); '*To err is human: building a safer health system*' (IoM, 1999).

iv. **References** - All relevant articles retrieved and subject to appropriate selection criteria were hand searched for any additional citations to other articles which previous search techniques had not already retrieved.

v. **Unpublished studies** - The CRD systematic review guidelines recommend the inclusion of unpublished studies but suggest this search step should be omitted if the length of time allocated for the review does not permit this.

3.2.1.4. Search terms

CRD guidelines recommend that search terms should be sensitive enough to identify all available relevant articles yet specific enough to exclude irrelevant articles which can hinder the search process. This process should be iterative and guidelines suggest several attempts should be made to develop an adequate strategy. Using a structured approach, the review topic was broken down into four 'facets', which included study population (area of care), interventions (systems variable targeted), outcomes (type of error) and study designs. Search terms representing each facet were then generated between the principal researcher, academic supervisors and other collaborating partners[13].

In order to access electronic databases, the following search terms were used for each of the four facets;

1. **Type of error** – these terms were considered to incorporate multiple variations of medication error. It was anticipated that whilst this would also access studies which had not necessarily

[13] As part of this research the principal investigator was invited to join the hospital's patient safety steering group which provided valuable input throughout the duration of the project.

measured MAEs (e.g. many had measured transcribing errors), trial attempts at running the search using terms *only* related to MAEs was too stringent and yielded few relevant studies.

Medication, medication error, drug error, drug administration error, drug event, adverse drug event, ADE, near miss: 8 exploded terms (to include any related terms) combined using OR; yielded 8,044 results

2. **Area of care** – these items were considered to reflect the area of care most relevant for this research. Several search attempts revealed these terms were specific enough to exclude any studies conducted outside of secondary care (i.e. primary care, ambulatory care).

Secondary care, secondary, tertiary, tertiary care, hospital, in-patient, medicine, ward, hospitalised: 9 exploded terms combined using OR; yielded 5,138 results

3. **Systems variable targeted** – although Hoff's review employed a top-down strategy to search 16 key organisational variables, a key objective of this review was to explore the literature and identify studies which may not have necessarily used well-known terminology. One way of achieving this was to impose less stringent criteria on this aspect of the search. CRD recommendations suggest researchers should find the middle-ground between developing broad search terms which will successfully retrieve a comprehensive selection of relevant articles (and ultimate several irrelevant ones), and developing overly-specific terms which may bias the retrieval (and retrieve too few relevant articles). Preliminary searches revealed that as predicted there was a general inconsistency in the terminology used to describe organisational and workplace variables (see section 3.4 for discussion of findings). In order to retrieve articles which had possibly measured the same variable as other articles but used different terminology, a broad approach to search terms was taken. Government documents

(DoH, 2000, 2001; IoM, 1999) were reviewed for appropriate search terms which would encapsulate the *nature* of systems variables (i.e. organisational, workplace, systems etc.) if not the variable itself.

Organisation, organisational, system, systems, systemic, environment, environmental, workplace, working condition, local conditions, ergonomic, ergonomics, ergonomic factors, human factor, human factors; 15 exploded terms combined using OR; yielded 80,822 results

4. **Study design** – terms commonly used to describe various empirical research designs were developed.

Intervene, intervention, randomised controlled trial, RCT, controlled trial, trial, case study, time series, cohort, control group, non-randomised, before/after study: 12 exploded terms combined using OR; yielded 3,930,246 results

All four search results were combined using AND; yielding 1223 results. CRD guidelines recommend applying a number of filters to ensure all abstracts to be reviewed in the first instance are relevant. Filter 1 removed all duplicate articles (since multiple databases had been searched), leaving 912 results. Filters 2 and 3 removed any articles not published in the English language and those published before 1990 (since several of the databases did not allow for this specification beforehand). A total of 714 articles remained.

3.2.2. Inclusion/Exclusion criteria

As recommended by CRD guidelines, article titles and abstracts were screened in the first instance for relevance, followed by a full review of the article if the following criterion was met;

i. **Intervention** – As discussed previously, only those studies which had proposed potential systems (organisational and/or workplace) causes of medication errors *and* implemented some form of systems intervention to reduce these errors were included. There was no inclusion restrictions on the type of intervention carried out. Studies which provided insufficient explicit detail to determine the intervention employed to reduce errors with enough detail to be *replicable* were excluded.

ii. **Field of health care** - Studies conducted within any in-patient secondary or tertiary care setting were included, provided that the intervention or systems problem described was not so specific that it would not be generalisable to other fields of health care. For each article extracted, a judgement was made by the reviewer on whether the systems variable identified and intervention employed was generalisable beyond the clinical specialty in which it was conducted. Since this research focussed upon errors made in secondary care, it was decided that studies conducted in residential care facilities, primary care or any other outpatient or community setting would be less useful for the development of the safety tool and so these studies were excluded. For example, Eccles *et al.* (2002) measured the effects of implementing computerized evidence-based guidelines on the prevalence of medication errors in the management of asthma and angina in primary care. Although employing a systems intervention, this study was excluded since it was not conducted in secondary care and referred to processes specific to primary care which were not considered transferable to secondary care settings. There was no selection criteria imposed on the types of health care staff members participating in the studies.

iii. **Outcome measure** – it was originally hoped that only studies which had employed a measure of MAEs would be included since this was the main focus of this research. However, during preliminary screening of extracted articles it became apparent that very few studies had measured errors at the administration process of medication delivery. The decision was therefore made to expand the search and include studies which had measured errors at *any* of

the stages of medication delivery including prescribing and transcription. Although the focus of this research was on errors occurring in the administration of medication, it was acknowledged that this process relies upon the combined success of preceding stages. It was anticipated that studies which had explored other aspects of this process would shed light on downstream causes of error at the administration stage (e.g. prescribing practices).

Most of these outcome measures were in the form of self-reported errors (either via survey or those formally reported in hospital reporting systems) since there are few other error measures available to health care researchers. Other studies employed the use of medical chart reviews or observational methods. All methods were considered eligible outcome measures. Several studies reported interventions to reduce the incidence of Adverse Drug Events (ADEs) but had not specified whether these 'events' were the outcome of error or were non-preventable unexpected events (e.g. giving penicillin to a patient without knowing they were allergic). Only studies which explicitly attributed ADEs to some type of error in prescribing, dispensing or administration of the drug were the focus of this review.

Studies which had used *only* quality improvement measures as their outcome variable (such as those included in Hoff's review) were excluded since the focus of this research was on the improvement of patient safety, in the form of medication error reduction rather than quality of care. For example, some studies extracted had investigated reductions in litigation expenses, length and costs of patient stay, patient or health professional ratings of quality of care or nurse stress and burnout levels.

Studies which provided insufficient detail of pre- and/or post-intervention comparisons on outcome measures were also excluded. For example, some studies simply stated that interventions had been 'successful' in reducing errors (Leape, 2000). Only studies which

stated explicitly the number (or percentage) of medication errors made before and after intervention were included in the full review.

3.3. Results and analyses

After initial filtering processes, 788 articles were subject to the selection criterion after which a total of 769 were excluded due to a lack of empirical reporting of improvements post-intervention, insufficient detail regarding the type of intervention, the methodology and /or intervention success and poor or non-existent definitions of the types of medication errors (MEs) measured. The final review consisted of 19 articles which had identified potential systems causes of medication errors and employed systems interventions to reduce them. All articles were analysed in terms of several salient categories such as the areas of hospital care they were conducted in, their methodological design, the systems variables targeted, any theoretical framework or approach used for intervention design and finally the types of outcome variables (types of medication error) measured. Although each of these general areas of research interest will be discussed briefly under their corresponding subheadings, this review will focus predominantly on the systems interventions employed in order to assist in the development of the error management tool. Each article was assigned an identification number from 1-19. These numbers are used in tables and discussion to improve synthesis and discussion of articles. Summary tables of articles reviewed (including identification numbers and references), comparison of key findings, and specific intervention details of each study are available in appendices I, II and III respectively.

3.3.1. Area of hospital care

Studies were conducted in a broad range of clinical areas. Table 3.3.1 presents a summary of these areas. More than a third of the studies reviewed were conducted in paediatrics or neonatal care settings. This is unsurprising since there is an abundance of evidence highlighting the significantly

higher risks of MEs occurring within this patient population compared to adult patients (Wong *et al.,* 2004; Gorman *et al.,* 2003; Marina *et al.,* 2000) and research has estimated the likelihood of these errors resulting in patient injury to be three times higher in paediatric patients (Fortescue *et al.,* 2003). Davey *et al.* (**11**) note that these increased risks are largely due to the need for more individualized calculations of drug doses based on patient weight and report previous findings that estimate there are potentially 50,000 paediatric dosing errors per year in England (Wong *et al.,* 2004; Costello *et al.,* 2007).

Table 3.3.1: Clinical specialties of reviewed studies

Area of medical care	Articles *	Percentage of total
Paediatrics / Neonatal	**3, 11, 13, 15, 16, 17, 18**	37%
General Medicine	**1, 2, 7, 10**	21%
Critical / Intensive Care	**1, 2, 5, 19**	21%
Unspecified Secondary Care	**8, 12, 14**	16%
Coronary care	**5, 19**	11%
Surgery	**7, 9**	11%
Emergency Care	**6**	5%
Anaesthesia	**4**	5%

*** Studies 1, 2, 5, 7 and 19 were conducted in multiple areas of hospital care**

One particularly important aim of this review was to explore areas of care which had implemented systems interventions to reduce medication errors, whose findings were generalisable to other clinical specialities. Only studies which reported findings not limited to a highly specialised field of patient care were included in the review. For example, several studies were identified which reported the effects of interventions on errors of one particular chemotherapy drug (Schiff, Aggarwal, Kumar & McNutt, 2000) whose intervention related to a highly specialised process of cancer therapy and were subsequently excluded. By comparison, although the risks of dosage

errors in pediatric settings are elevated, general dosage errors *per se* are not limited to that field of health care. Similarly, a study by Landrigan *et al.* (**5**) reported that many serious MEs were being made by medical interns (e.g. trainee doctors) on coronary care and intensive care wards, whose shift patterns meant they could potentially be expected to work up to 70-80 hours per week. The effects of sleep deprivation as a result of long working hours and / or shift patterns is not specific to coronary care and systems interventions to reduce errors are unlikely to be beneficial only in this area of care. For this reason such studies were retained.

3.3.2. Methodological design

Table 3.3.2 below details the methodological designs employed in the 19 studies reviewed.

Table 3.3.2: Methodological design of reviewed studies

Methodological design	Articles	Percentage of total
Before/After quasi-experimental study (no control group)	**1, 2, 4, 5, 6, 9, 10, 11, 12, 14, 15, 16, 17, 18**	73%
Randomised Controlled Trial	**7, 19**	11%
Non-Randomised Controlled Trial	**8, 13**	11%
Cohort Study	**3**	5%

Almost three quarters of the studies extracted employed before/after study designs *without* control group comparisons. These studies involved comparing the number of MEs occurring pre-intervention with those occurring post-intervention. Interestingly, there was a large degree of variation in the length of post intervention follow-up, ranging from only one week (**11**) to four year follow-ups (**2**). For example, Bates *et al.* (**2**) examined the impact of a computerized physician order entry (CPOE) system upon MEs over a period of four years. This study recorded a baseline

measurement of the number of reported MEs made pre-intervention in 1992, followed by three follow-up measurements of errors in 1993, 1995 and 1997. This level of follow-up is rare within heath care research due to time, financial resources and recruitment constraints common to the field. In this instance, it allowed the researchers to make additional changes to the computerized system to reduce errors further such as the addition of allergy warnings and drug-drug interaction warnings (referred to as decision support). By comparison, Davey *et al.* (**11**) measured the impact of a prescribing tutorial and introduction of decision support for junior doctors on the number of prescribing errors executed only one week after each change had been implemented. Although they report that the number of prescribing errors significantly decreased after the interventions, one might argue that a follow-up occurring after only one week is not long enough to detect the durability of that intervention. Without control group comparisons it is possible that significant effects may have occurred as a result of the *Hawthorne effect* (Landsberger, 1958). This effect has been described as a short-lived reactivity to quasi-experimental conditions which results in improved performance which does not occur as a result of the condition being tested but simply because participants are being observed and as such are being paid more attention. Studies have shown participants will respond positively to any novel change in work environment (Mayo, 1949; Gillespie, 1991). Parsons (1974) furthers this argument to suggest that if the nature of the experiment involves participant feedback (as it did in the case of Davey *et al.*) this effect will be more profound. Without extended follow-up periods, it is difficult to determine whether significant reductions in MEs observed after intervention have occurred as a result of the intervention itself.

One way of controlling for this confounding variable is to use an appropriate control group. Interestingly, only two of the studies employed randomised controlled designs (RCTs). These studies (**7 & 19**) involved comparisons of MEs between wards implementing systems interventions and control wards that had not. These interventions are described below under *human resources* and *decision support* subheadings respectively. Neither study found a significant effect of their

respective interventions on the number of MEs (**7 & 19**). ADEs, potential ADEs or 'serious' medication errors (**19**).

The fact that only two studies had attempted RCTs is unsurprising for several reasons. Firstly, they are notoriously costly and often require long-term funding for external positions and/or the employment of the most experienced members of staff to carry out the research (Pringle & Churchill, 1995). Bottomley (1997) suggests that appropriate recruitment and randomisation of participants to groups requires significant effort which means these studies are largely conducted only within 'centres of excellence'. Even when well designed, RCTs suffer problems with participant recruitment, attrition and 'opposition to randomisation' which Fallowfield (1995) argues can all lead to results which are difficult to interpret at best. Bottomley proposes that while RCTs are undoubtedly rigorous, their implementation may be at the cost of 'methodological problems which may limit generaliseability of findings' (Bottomley, 1997, pp. 228) which could explain why so few attempts to implement RCTs was observed in this review.

One particularly noteworthy article implemented multiple systems interventions and has been categorised as a before/after study since it provided baseline and post-intervention ME data for eight hospitals. Silver and Antonow (**14**) reported the results of a multi-hospital systems intervention scheme involving large scale recruitment of thirteen hospitals across Utah (although data was only reported for eight of these). These hospitals adopted a variety of interventions to reduce MEs and whilst data are presented *before* and *after* interventions, it is an aggregate of all eight hospitals. One criticism of the paper is that because not all hospitals employed the same type or number of interventions, the authors do not report which interventions yielded the most significant benefits in terms of error reduction, a finding which would be useful for those organisations hoping to utilize *some* of the changes in their own institutions. However, the absence of such methodological detail does not represent a failing because the objective of the article is to

present a quality improvement initiative to reduce MEs and is reported as such. Unfortunately, in terms of organisational learning, such a study may be dismissed by organisations seeking more rigorous methods of 'gold standard' research on which to base their own interventions.

3.3.3. Theoretical framework

As discussed previously, there has been a large body of published research on the potential systems causes of high-risk industrial error (see chapter 2); a large proportion of which has focussed on Reason's organisational accident framework (Reason, 1990, 1997) or other human factors approaches to understanding error causation (Perrow, 1984; Rasmussen, 1983). Extracted articles were reviewed for any evidence that identification of organisational/systems causes of medication error or the implementation of systems interventions had been driven by known theories of human error.

Almost all articles reviewed (18 out of 19) provided no indication that their designed intervention(s) were based around any theoretical framework or human error principles. Only one article (**14**) intimated that their interventions had involved redesign of their systems based on the Institute of Healthcare Improvement's Breakthrough Series model (Langley, Nolan & Nolan, 1992) which had evolved from human factors principles of task and process design identified by human factors engineers in industrial research[14]. However, the authors of this paper do not elaborate exactly *how* their interventions had been developed with this theory in mind; only that staff were 'encouraged to select strategies on the basis of the literature on human factors in non-medical settings' (pp. 333).

In a similar article by Leape *et al.* (2000), which was extracted but not included in the review because it did not meet inclusion criteria, 739 individual systems changes were presented such as

[14] In 1996, 38 US hospitals and other health care organisations formed a collaborative under the leadership of the Institute for Healthcare Improvement to reduce medication errors. Teams from multiple hospitals were recruited to develop 'best practices' to improve the quality of health care.

improvements in error reporting, increased standardization of procedures, checklist development and implementation of protocols, based on the IHI's Breakthrough Series model, conducted in forty hospitals over a fifteen month period. They report an overall 'successful reduction' of medication-related errors of 20 per cent but like Silver and Antonow (**14**), do not describe in any detail which particular interventions were most beneficial in reducing which types of errors in which areas of patient care. Without this transferable information, the data simply presents a showcase of how successful one organisation has been and offers little opportunity for patient safety learning. The IHI model of improvement itself however, does not offer guidance on the types of organisational or systems interventions which are most likely to be effective under certain circumstances or for particular environments. For example, adopting one of the model's 'best practices' seen to effectively reduce prescribing errors in paediatric care may be less effective in emergency or elderly care settings. The only way in which staff in a specific environment would know whether one particular type of intervention is warranted or will be effective in their field would be through reporting of *empirical* evidence or information on interventions which are *generally* effective.

3.3.4. Systems variables addressed

Extracted articles were reviewed to determine from the interventions implemented, which systems failures had been addressed. Of the 19 articles reviewed, there was very little use of common systems terminology in descriptions of interventions. Those studies which had attempted to classify their type of intervention showed little consistency in the terminology used. Additionally, the terminology used in each article did not always provide the most accurate reflection of the variable the authors had actually addressed. Therefore, the variables presented in Table 3.3.4 are a culmination of terms described within the reviewed articles (if presented), General Failure Types (GFTs, Groeneweg, 1992; see also chapter 2 of this thesis, section 2.3.1.1) in cases where these terms were a better reflection of the variable than the term used within the article, and the reviewers

own judgement. An independent reviewer was then given a list of systems terminology which included any terms used in the review articles, 11 GFTs (with definitions) and others which the primary reviewer had generated in the course of the review. The second reviewer was given each article's methods section and asked to categorise the systems variable(s) measured using one of the terms provided in the list. Agreement was reached on the variable being measured for all but 2 of the 19 studies. After discussion, agreement was reached on all 19 studies.

Each variable will be discussed individually under corresponding subheadings followed by a more detailed discussion of salient points and data synthesis in section 3.4. Specific details regarding each intervention are listed in Appendix III.

Table 3.3.4: Systems interventions employed by reviewed studies

System variable		Articles	Percentage of total*
Standard operating procedures (G)		**1, 2, 6, 9, 10, 12, 13, 14, 15, 16, 17, 18**	63%
Decision support		**1, 2, 8, 11, 16, 17, 18, 19**	42%
Education / training (G)		**4, 8, 11, 14**	21%
Design (G)	Equipment	**4, 14, 16**	16%
	Workplace	**3, 14**	11%
Human resources		**7**	5%
Local working conditions		**5**	5%
Team Communication (G)		**1**	5%

*** Some studies identified more than one variable**
(G) Systems variables which are specified as a General Failure Type (Groeneweg, 1992) and proposed to be the most likely error antecedents

3.3.4.1. Standard Operating Procedures

By far the most frequently implemented systems intervention involved making changes to standard operating procedures (SOPs). All of these articles (n=12) involved changes made to prescribing practices. This is not surprising since there is considerable evidence that approximately 50 per cent

of MEs occur at the stage of drug ordering (Bates *et al.,* 1995; Kaushal *et al.,* 2001). Notably, all SOP studies except three involved the implementation of Computerised Physician Order Entry (CPOE) systems (**1, 2, 6, 9, 10, 13, 15, 17 & 18**). CPOE is an electronic method of prescribing medication and aims to improve poor prescribing practices related to inappropriate drug chart design, illegible handwriting, delays in the medication delivery process (from ordering to administration) and difficulty keeping track of paper drug charts among others. Interestingly, few of the articles reviewed described any of the above benefits of CPOE as a rationale for its implementation or identified that these aspects of prescribing practice were a particular problem in their hospital. In general, these articles referred to government directives to integrate this technology into secondary care (DoH, 2000; IoM, 1999). It should be emphasised that standard CPOE systems can simply involve a change in prescribing practices from paper-based to electronic ordering of medication. More sophisticated systems can also 'add-on' decision support options including drug libraries and error-prompts. Since not all studies review had installed such advanced systems, studies which had measured the impact of decision support add-ons to standard CPOE systems are discussed separately under the decision support subheading.

All nine SOP studies which measured the effects of CPOE implementation on some form of ME used various definitions of this type of error and employed a variety of methods to measure them making it difficult to draw overall conclusions as to whether this type of intervention was successful. For example, Spencer *et al.,* (**10**) measured the effects of CPOE implementation on self-reported ordering/dispensing and administration errors. Franklin *et al.,* (**9**) however, measured the effects of CPOE on prescribing and administration errors as observed by independent pharmacists. It is unlikely that these two methods of measuring errors are comparable, largely due to reporting bias and other factors which may influence whether individuals are willing to report their own behaviour (see chapter 6, section 6.2).

Putting this variability in defining and measuring MEs to one side, CPOE successfully and significantly (statistically) reduced *some forms* of medication error in seven of the nine studies which had implemented the system. The remaining two studies reported that they had observed decreases in medication errors post-intervention (dosage errors of Gentamicin and general prescribing errors respectively) but did not report whether these reductions were statistically significant (**6 & 17**).

The three studies which had not implemented CPOE systems but have been classified here as employing changes to SOPs, employed the use of a standardised protocol of prescribing insulin (**12**), standardised medication ordering procedures for high-risk medications (**14**) and standardising drug concentrations (**16**). Two of these studies reported statistically significant reductions in prescribing errors (**12**), ten-fold dosage errors and medication-infusion errors (**16**). Although Silver and Antonow (**14**) report a statistically significant overall decrease in medication errors after standardising ordering procedures, the article combines data for multiple hospitals employing a variety of interventions so individual intervention success cannot be ascertained.

3.3.4.2. Decision Support

Eight studies measured the effects of providing decision support on MEs. Four of these studies involved implementation of CPOE systems with additional decision support options (**1, 2, 17 & 18**). CPOE systems are usually commercially manufactured electronic prescribing aids tailored according to the needs of individual health care organisation and most importantly designed according to budget. Of the nine studies which implemented CPOE systems, five did not incorporate additional information which could be defined as decision support or did not specify whether they had or not (**6, 9, 10, 13, & 15**). For example, the electronic prescribing system implemented by Bizovi *et al.* (**6**) comprised a 'low-spec' computer system of medication selection

allowing prescribers to choose medications, dosage, quantity and frequency from a drug 'pick-list'. This system alleviated some but not all problems associated with prescribing practice. Similarly, Franklin *et al.* (**9**) measured the effects of a computerised prescribing, dispensing and administration system which employed the use of hand-held prescribing computers, synchronised with other computers across the hospital. Although this computer system comprised pull down lists of available drugs, no actual decision support was incorporated into the system.

By comparison, all four of the studies identified here as decision support interventions which had employed CPOE had significantly more sophisticated versions of the system. For example, Potts *et al.* (**18**) describe their CPOE system as including several decision support functions such as drug allergy checking, contraindication and incorrect dosage warnings. Similarly, Bates *et al.* (**1**) describe their CPOE system as incorporating 'consequent order prompts' (e.g. orders that should follow from other orders according to the patient's previous clinical history) allergy checking and electronic presentation of patient test results which might affect the medication selection. In this way, doctors are presented with more sophisticated information than standard CPOE. It is acknowledged that although some CPOE 'add-ons' described involved providing physicians with decision support (e.g. drug libraries, dosage options etc.) other aspects of these decision support options might also have been categorised as the GFT *defences* (e.g. drug allergy warnings; see chapter 2, section 2.3.1.1). Since these elements provide the user with *support* in making the correct *decision* it was decided that decision support was a more apt classification of these elements.

It is difficult to determine whether the addition of decision support to CPOE systems is more, less or equally effective in reducing medication errors overall due to the considerable variability in outcome measures reported. Only one study specifically compared the cumulative effects of adding decision support to basic CPOE systems (**2**). In a before/after study, Bates *et al.* measured the effects of implementing a basic CPOE system on three hospital wards in intensive care and general

medicine, followed by two attempts to gradually increase decision support aid over the following three years. Their results revealed that the introduction of a basic CPOE system significantly reduced *some* types of medication errors (e.g. non-missed dose errors, serious medication errors, non-intercepted potential ADEs) by 64 per cent overall compared to baseline measures. The addition of decision support functions (e.g. an allergy warning system in phase 2 and forced ordering of potassium chloride in phase 3) significantly decreased more types of error (e.g. all of the above plus preventable ADEs and intercepted potential ADEs) by a further 64 per cent compared to standard CPOE.

Two of the remaining four decision support studies measured the effects of implementing 'smart' infusion pumps on medication infusion errors (**16 & 19**). Smart pumps address the increased risk of error posed by individualized medication infusion concentrations by calculating standardized rates of fluid administration and alerting users to incorrect dosage or flow rates, with maximum dosage ranges to eliminate ten-fold dosage errors. Larsen *et al.* (**16**) investigated the effects of combining smart pumps with the provision of standardised drug concentrations (based on 32 common intravenous medications calculated by patients weight) and 'human engineered' medication labels to facilitate the correct transfer of information to the pump. Although a 73 per cent decrease in medication infusion errors was observed, the independent effects of each intervention were not measured so it is difficult to determine which particular effort was the most effective in reducing errors. It is possible that the combination of standardised drug concentrations and medication labels designed according to human factor principles was both cheaper and a more effective strategy of error reduction than the implementation of smart pump technology. This idea is strengthened in a similar study by Rothschild *et al.* (**19**). This study investigated the effects of using smart pump decision support technology in cardiac surgery. Results revealed no significant impact on serious medication errors or Adverse Drug Events (ADEs) after implementation of *only* smart pump technology. Authors attribute this non-significant result in part to the observation that although the

technology offered *some* support in decision making, it was still possible to over-rule the guidance given and over-rides were frequently observed.

These findings do however raise the question of whether simpler, less costly methods of intervention based on human engineering principles would be both easier and more likely to be implemented and ultimately more effective in error prevention long-term. With this in mind, the final two decision support studies (**8 & 11**) measure the effects of much simpler, less technology based interventions on prescribing errors. Shaw *et al.* (**8**) tested the effects of 'academic detailing' on prescribing errors in drugs of addiction (DoA). Part of their intervention comprised one-to-one interviews with junior doctors to discuss and troubleshoot problems with prescribing DoA in terms of the legal requirements according to state law. This element of their intervention will be discussed under the subheading *education and training*. The other component of their intervention involved provision of a bookmark to all junior doctors with a ten-point summary of legal prescribing requirements for DoA on one side and a sample prescription on the other. Although they described their intervention as 'educational', they did not acknowledge that they had also tested the added effects of decision support in the form of a bookmark prompt. Findings revealed that there was a significant decrease in the rate of prescribing errors (from 41 to 24 per cent of the total written over a 4-week period) compared to a control group. However, authors report only the overall effectiveness of their educational *and* decision support interventions but do not report the respective effects of the additional decision support aid. This information would have undoubtedly been useful since decision support prompts are cost effective and require fewer man hours to produce compared with education interventions which require significantly more effort. It is possible that the prompt alone was a sufficient memory aid to reduce prescribing errors in this study.

Comparatively, Davey *et al.* (**11**) measured the respective effects of education and decision support for junior doctors on paediatric prescribing errors. Their intervention involved an interactive

pharmacist-led tutorial for junior doctors on appropriate prescribing practices which included legibility of handwriting, placement of decimal places and zeros, units of administration and dosage calculations (not currently taught on UK undergraduate medical training). This tutorial was followed by the formulation of standardised bedside guidelines for calculating 22 common drugs utilised on paediatric wards according to patient age and body weight. Results revealed a significant reduction in prescribing errors after the educational tutorial but no further reduction in errors was seen from introduction of the decision support tool. These findings would appear to suggest that while educational interventions alone can be successful in reducing MEs, the addition of decision support aids offers no further benefit to error reduction. However, post-intervention follow up data for both studies was just one week (**11**) and two months (**8**) so it is possible that clinicians had little need to rely upon the decision support aid since their knowledge had only recently been refreshed. In order to determine whether such simple and cost effective decision support aids could be *independently* useful in reducing error likelihood (or that the findings were not due to the Hawthorne effect), a longer follow-up would be required or a separate test of the effects of *only* decision support.

While Shaw *et al.* did not measure the cumulative effects of adding decision support to formal education and Davey *et al.* did, neither study referred to this aspect of their intervention as decision support, which is crucial to patient safety learning. If an organisation has identified a problem with prescribing practices and is looking for an inexpensive method of reducing errors as an alternative to the very expensive implementation of CPOE, it would be useful if this could be readily identified from the patient safety literature. By making little effort to label interventions appropriately, opportunities for learning will be lost.

3.3.4.3. Education and training

Four of the nineteen articles reviewed investigated the effects of education/training on medication errors. Two of these studies measured the *direct* effects of specific educational interventions. As already discussed previously, Shaw *et al.* (**8**) and Davey *et al.* (**11**) describe interventions which aimed to measure the impact of educational programs in the form of one-to-one troubleshooting interviews and prescribing practice tutorials respectively on prescribing errors. Both studies report statistically significant reductions in prescribing errors made by junior doctors. However, both studies combined their primary educational intervention with additional decision support aids. By comparison, Fasting and Gisvold (**4**) measured the *indirect* effects of combining 'medication safety education' as a secondary intervention with an intervention primarily focussed on the design of equipment. In this study, only the *combined* effects of the two interventions were reported. Authors provide only a cursory mention of their educational intervention, reporting that 'educational department meetings' were held to discuss medication errors and the mechanisms behind them. No further details were provided. Similarly, Silver and Antonow (**14**) report an educational intervention carried out by five of their recruited eight hospitals which had implemented 'comprehensive educational programs' aimed at identifying drug knowledge deficiencies in nursing staff to increase their 'awareness for error potential'. As with Fasting and Gisvold, Silver and Antonow provide no detail on *exactly* what form this comprehensive educational intervention took or how it was developed or implemented.

Providing this level of detail is especially important in educational interventions which can be both costly and time-consuming (Berkson, 1993) since research has shown there to be a vast range of educational approaches available to teaching skills and knowledge to health care professionals. For example, small group problem based learning (PBL) strategies in particular have been shown to be significantly more successful than more traditional educational methods in health care settings (Walton & Matthews, 1989; Vernon & Blake, 1993). However, Savin-Baden (2000) found there to

be five different types of PBL strategies alone currently being employed in UK medical institutions and Newman (2003) points out that established guidelines on the implementation of PBL strategies (Barrows, 2000) are not widely used. It is vital therefore that research which has involved educational interventions report *explicitly* the types of methods employed for the benefit of replication and as a guide to 'what works'.

3.3.4.4. Design

Four of the nineteen studies reviewed measured the effects of design interventions on medication errors. These studies have been subcategorised according to the nature of their design changes. Two articles measure the impact of changes to the design of the workplace (**3 & 14**). Walsh-Sukys *et al.* (**3**) measured the effects of ergonomically reducing sound and light levels on medication errors in a neonatal intensive care unit. They hypothesise that excessive light and sound levels may contribute to increased levels of stress in both patients and nurses, which may increased the risk of making errors. However, confusingly they also argue that reducing light and sound levels may also *cause* errors. Whilst a well controlled study, the authors do not report the success of the intervention in terms of the reduction of errors on the experimental ward; only the difference in error rate between the experimental and control wards which was non-significant. In a comparatively larger study involving seven of their eight recruited hospitals (although the exact number of participating wards or individuals is not reported), Silver and Antonow (**14**) describe two distinct strategies to reduce medication administration errors by making changes to workplace design. One intervention involved 'restriction of physical access to potentially lethal drugs' and the other involved separating look-a-like drugs in clinical storage cupboards. As has been mentioned previously, Silver and Antonow do not report data for each intervention, only the combined success of all interventions across all eight recruited hospitals.

Three studies describe ME reduction strategies based on the design of equipment (**4, 14 & 16**). Fasting and Gisvold (**4**) measured the impact of colour-coding syringe labels in an attempt to reduce medication administration errors in anaesthesia. Their findings showed no significant decrease in errors after the intervention, though this intervention was combined with an educational program which may have had an impact on the number of errors reported (see section 3.4. for more detailed discussion). In a similar study, Larsen *et al.* (**16**) measured the impact of designing intravenous fluid labels based on 'human engineering principles' to facilitate the transfer of correct information into IV pumps. This intervention was a tertiary strategy to reduce medication infusion errors and was combined with the primary and secondary decision support strategies of smart pump technology implementation and provision of standardised drug concentrations. Although an overall significant reduction in medication infusion errors was observed, as discussed previously the respective effect of this design change was not measured. The final study to test the impact of equipment design changes on MEs was reported by Silver and Antonow (**14**). In their collaborative report, the authors report an intervention to improve the 'design of medication administration records' to improve clarity for nursing staff. However, the authors do not report *exactly* how this was performed or whether this particular design strategy was independently successful in reducing errors.

3.3.4.5. Human Resources

Only one of the nineteen studies measured the impact of employing changes in human resources on MEs. Greengold *et al.* (**7**) measured the effects of changing the job role of nursing staff to yield reductions in MEs. This study hypothesised that MEs would decrease by having dedicated 'medication nurses' who could focus entirely on administering drugs to overcome problems associated with frequent prescription changes and an abundance of new drug products. This strategy did not significantly reduce MEs overall. However, this result may have been a knock-on

effect of combining a secondary educational intervention with the primary changes in human resources. This argument will be discussed in more detail in section 3.4.

3.3.4.6. Local working conditions

Only one study had attempted to measure the effects of what could be classified as a *local working conditions* intervention. The Harvard Work Hours, Health and Safety Group (**5**) measured the effects of reducing the number of hours worked consecutively in one shift and per week by doctors (interns). The intervention compared the number of 'serious' MEs (among others) during a traditional intern work schedule of 77-81 hours per week with the possibility of 34 continuous working hours with an intervention schedule comprising a maximum of 63 hours per week and no more than 16 hours of continuous work per shift. Results revealed that interns made significantly fewer serious MEs during the intervention period.

3.3.4.7. Team Communication

Again, there was only one study which claimed to have measured the effects of a team intervention on subsequent MEs. Bates *et al.* (**1**) investigated the combined effects of a primary CPOE intervention with a secondary team intervention developed to improve team communication. The authors provide sparse detail on this secondary aspect of their intervention (one small paragraph) and the information they do provide appears to describe multiple decision support additions such as standardised dilutions charts, a drip-rate calculation program and standardised labelling of IV bags, tubes and pumps. The only aspect of their intervention which seems to directly address the way in which a multi-professional team might function or communicate is in the form of a nurse to pharmacy communication log book to improve inter-disciplinary communication. Nevertheless, the authors did attempt to measure the independent effects of their 'team intervention'. The authors

report no incremental effect of the team intervention over the benefit of CPOE implementation compared to wards who had received CPOE alone.

3.3.5. Outcome measures employed

Each of the articles was reviewed in terms of their employed outcome measures. Table 3.3.5 summarises these findings which are discussed in more detail below in terms of the provision of error definitions, the utilisation of known error classifications and the methodologies used to measure these outcome variables.

Table 3.3.5: Outcome measures employed by reviewed articles

Error type	Articles	Percentage of total*
Medication Errors - All	**2, 3, 10, 13, 14, 15, 19**	37%
Prescribing Errors	**6, 8, 9, 11, 12, 17, 18**	37%
Medication Administration Errors	**4, 7, 9, 16**	21%
Serious Medication Errors	**1, 5**	11%
Adverse Drug Events	**2, 19**	11%
Potential Adverse Drug Events	**2, 19**	11%

*** Some studies employed more than one dependent variable**

3.3.5.1. Definitions and classifications

Studies were reviewed in terms of their definition of medication error and their use of known classification systems. Sixteen of the nineteen studies reviewed reported measuring only one type of medication error, while the remaining three studies reported measuring two (**9**) and three (**2 & 19**) types of medication error. Eight studies reported additional outcome measures including procedural and diagnostic errors (**5**), physician confidence in prescription writing (**8**), time taken to

conduct the drug administration round (**9**) and pharmacy turn-around times (**17**). Of the seven studies purporting to measure 'all' medication errors, five used the NCC MERP[15] definition (or variations of) which denotes; '...any error in prescribing, dispensing, administering and monitoring medications regardless of outcome...' (**2, 10, 13, 15 & 19**). The remaining two studies did not provide a definition of *all* medication errors (**3 & 14**).

Seven studies measured errors associated with prescribing medication. Two of these articles presented no definition of prescribing errors (**8 & 12**). Another study did provide a definition of prescription error; although the authors measured only dosage errors of one high-risk drug (Gentamicin) given on admission to a Neonatal Intensive Care Unit. This error type was so specific it required little defining (**17**). An additional two studies stated they had used established definitions of prescribing error and prescribing classification systems (Dean, Barber & Schacter, 2000; AAP, 1998), though they did not provide these definitions within the article (**9 & 11**). Finally, the remaining two studies measuring prescribing errors presented detailed definitions of prescribing errors and the ways in which these errors had been taxonomised (**6 & 18**). For example, Potts *et al.* (**18**) give a clear definition of the types of prescribing errors they measured and classified them as either potential adverse drug events (ADEs), medication prescribing errors (MPEs) or rule violations (RVs). The authors provide clear definitions and examples for each category such as providing incorrect or inappropriate information on ordering medication, failing to account for patient-specific information such as allergies, problems with interpretation of handwritten orders and non-compliance with standard hospital policies. Similarly, Bizovi *et al.* (**6**) present a detailed description of their outcome measure which comprised medication orders which required 'pharmacist clarification'. Authors also provide examples of several clarifications such as missing or incorrect information, legibility problems, incorrect dose or drug selection.

---

[15] The National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) is an independent body comprised of 23 national organizations spearheaded by the United States Pharmacopoeia.

Four studies measured the effects of intervention on medication administration errors (MAEs). One of these studies provided no definition of what constituted an MAE (**16**), while another provided a vague reference to 'intra-operative problems' in the form of MAEs but little else (**4**). Greengold *et al.* (**7**) describe their outcome measure as discrepancies between physician's medication orders with independent observations of drug administration rounds which included documenting 'variations from safe medication practices'. The authors do note that observers were given the NCC MERP definition of medication administration error (see footnote 15) but this definition was not provided in the article. Finally, Franklin *et al.* (**9**) describe medication administration errors as; '...any dose of medication which deviated from the patient's current medication orders... ' However, no further elaboration was provided and the authors do not cite the use of any system of error classification.

Two of the nineteen studies described their outcome measures as 'serious' MEs (**1 & 5**). In their study measuring the effects of reducing work hours on MEs, Landrigan *et al.* (**5**) present cursory definitions of ten outcome measures; describing serious MEs as those which '...related to the ordering or administration of pharmaceutical agents, blood products, or intravenous fluid..', but do elaborate on the word 'serious' to differentiate between these and 'non-serious' errors. The authors do not provide a definition of error at all and present no examples of serious errors in lieu of a definition. Similarly, Bates *et al.* (**1**) provide some details on their definition of serious MEs, making the distinction between those which are preventable and those which are not, however, as with Landrigan *et al.* (**5**), they make no reference to the way in which an error is classified as 'serious'.

The final two studies measured the impact of their interventions on Adverse Drug Events (ADEs) or Potential Adverse Drug Events (PADEs) (**2 & 19**). As their secondary outcome variable, Bates *et al.* (**2**) measured the effects of CPOE on ADEs and PADEs. Authors provide detailed, though convoluted definitions of both measures. Additionally, authors present a highly complicated system

of classifying these errors in terms of their preventability, potential for harm and whether they were intercepted before reaching the patient which makes it difficult to ascertain exactly what has been measured. By comparison, Rothschild *et al.* (**19**) present a detailed yet clear description of both ADEs and PADEs, elaborating on when an error becomes an ADE. Neither study alludes to any known system of error taxonomy.

3.3.5.2. Methodologies

As the secondary aim of this research was to record medication error prevalence (in order to improve the predictive validity of the error management tool), articles were reviewed for their methods of error data collection. Table 3.3.5.2 summarises these findings.

3.3.5.2: Methods of data collection for reviewed studies

Method of data collection	Articles	Percentage of total*
Drug chart/patient record review	**1, 2, 5, 7, 11, 12, 17, 18, 19**	47%
Self report (survey)	**1, 4, 5, 14, 19**	26%
Spontaneous reporting system	**3, 10, 13, 16, 19**	26%
Independent observation	**5, 7, 9**	16%
Pharmacist report	**2, 6, 9**	16%
Not stated	**8, 15**	11%

*** Some studies used more than one method**

It was the general aim of this particular phase of the review to determine which methodology provided the most accurate measure of ME prevalence. Cullen *et al.*, (1995) suggest that spontaneous formal incident reporting (non-anonymous) ordinarily captures less than five per cent of the errors which actually occur. Research also suggests that significantly more medication errors

are identified when an anonymous system of error reporting is implemented (Harris *et al.,* 2007). Similarly, Barker *et al.,* (1984) estimate that when observational techniques of identifying errors are employed the prevalence of errors could be as high as one error per hospitalised patient per day; significantly greater than those reported 'officially' through more formal channels. In a recent comparison of methods across thirty-six hospitals, Flynn *et al.* (2002) found the number of errors identified using independent observation was significantly higher (and identified the highest number of events compared to all methods of data collection) than traditional incident reporting systems by a factor of 457:1!

The original aim for this part of the review was to combine these findings in a meaningful way to compare the prevalence of MEs detected by each method and the validity and reliability of different methods; the method detecting the highest number of errors presumably reflecting the most thorough methodology and therefore providing the best reflection of *true* error rate. Unfortunately, due to the wide variability in the types of medication errors measured by each study it was not possible to synthesise these studies in this way. For example, eleven studies employed only one method of data collection. However, four studies used two methods (**1, 2, 7 & 9**), two studies combined three methods (**5 & 19**) and two studies did not specify their methods (**8 & 15**). By combining data in this way, studies have not measured respective error rates detected by each method in the same way as Flynn *et al.* (2002) (see above).

Of particular interest to this thesis was the method of self-reporting errors via survey methods. It was determined early in the current research that since the author was not a trained health professional, chart review or clinical observations would not be appropriate methods of measuring error frequency. Therefore methods which did not require clinical expertise (or recruitment of such a professional) would be the most likely method employed in this research. Of the studies reviewed which had used survey methods to identify medication error prevalence, two measured 'serious'

MEs (**1 & 5**), one measured only MAEs (**4**) and two studies measured 'all' medication errors but provided conflicting definitions (**14 & 19**). It was subsequently not possible to compare the number of errors each respective method had identified because of the huge variability in these studies.

3.4. Discussion

This discussion will focus on several key issues which have arisen from the review followed by overall conclusions.

Lack of theory, intervention choice and use of organisational terminology

Recognising the value of human error theory and developing systems changes which reflect these theories (learning from other high-risk industries) are recommendations made by UK and US governments. A strong theoretical foundation is also fundamental in the design of complex safety improvement interventions (Eccles *et al.,* 2003). However, all but one of the nineteen studies reviewed gave no indication that Reason's Swiss cheese model, organisational accident framework (Reason, 1997) or any other 'systems' theory (Perrow, 1984; Rasmussen, 1983) had been utilised as a theoretical basis for their intervention. In the only study that did suggest Reasons model had been employed as a means of developing their intervention, the way in which the theory was used or to what level was not described. This could be in part attributed to the finding that although many health care organisations are aware of the model, considerable overlap in interpretation of Reason's theory has resulted in a lack of confidence in appropriately applying it (Perneger, 2005). The only attempts to employ Reasons' organisational accident framework was observed earlier in the review process in studies which had used the framework to propose systems causes of errors during *retrospective* incident analysis (Dean *et al.,* 2002; Vincent *et al.,* 1998; Meurier, 2000). However,

since these studies had not employed interventions to test their hypotheses, they did not meet inclusion criteria.

As well as being predominantly atheoretical, very few of the articles reviewed here referred to key terms such as systems, organisational, human factors, teams, decision support or standard operating procedures which are prevalent in other industrial error research (Reason, 1997). None of the studies acknowledged that their interventions had involved *organisational* or *systems* changes, despite having little to do with changes at an individual level (e.g. cognitive strategies, work shadowing or vigilance training). Furthermore, most of the reviewed articles provided very little empirical evidence to support their chosen interventions. Rationales were largely driven by a limited range of government objectives such as the introduction or modification of high-spec technology such as computerised physician ordering systems (DoH, 2000; IoM, 1999). These studies provided little or no explanation as to why employing these particular interventions was necessary or might be successful in reducing errors in their particular organisation.

Employing interventions to reduce errors in clinical settings is undoubtedly costly and time-consuming. However, if attempts are not first made to measure the systems variables which actually *require* intervention, changes may be inappropriate or unnecessary (Wensing & Grol, 1994). Additionally, if a variable targeted for intervention is problematic within a particular environment, the intervention may even lead to dynamic changes to other systems variables. Almost half of the studies reviewed here reported that at least one of their interventions was not effective in reducing medication errors. In terms of planning appropriate interventions, a theory driven tool which can proactively identify multiple systems weaknesses at a local level would unquestionably be useful for health care organisations.

Single vs. multiple interventions

More than half of the studies reviewed examined the *combined* effects of two or more interventions[16]. For example, Shaw *et al.* (**8**) and Larsen *et al.*, (**16**) describe only the combined benefit of implementing multiple interventions. However, it is possible that individual components of their interventions alone could have been responsible for the successful reductions in medication errors observed. These studies by inference suggest that only when *all* their interventions are combined will successful results be observed. It is possible that the respective contributions of each component of multiple interventions could be incremental or synergistic. For example, when Davey *et al.* (**11**) measured the *respective* effects of a two-stage intervention, results revealed that while the initial educational component was successful in reducing prescribing errors, the subsequent decision support phase generated no further benefit in error reduction. Similarly, Bates *et al.* (**1**) measured the respective effects of implementing a basic CPOE system, followed by two further interventions to gradually increase decision support. While the implementation of basic CPOE (phase 1) significantly reduced MEs compared to baseline measures, the occurrence of MEs increased significantly during their phase 2 intervention before finally decreasing once more after phase 3.

It should be taken into consideration that there are significant constraints in this particular field of applied research such as access to participants, ethical constraints and the availability of financial resources. As a result, by assessing and reporting only the combined results of multiple interventions, smaller organisations constrained by significantly lower financial budgets may struggle to implement *all* changes implemented in larger institutions. Fundamentals of research design dictate that research should be well described and reproducible. Health care organisations which implement interventions to change clinical practice are responsible for reporting explicit

---

[16] It should be noted that not all of these studies acknowledged they had measured multiple interventions. Descriptions of interventions were disentangled where this author considered more than one intervention had been employed.

replicable methodologies for interventions. Where multiple interventions have been employed, individual components of those interventions should also be described, including the respective effects of each intervention.

Almost half of the studies reviewed investigated the effects of only one intervention on the reduction of MEs. This approach to error management may be viewed as over-simplistic, serving only to slow research progress. However, there is considerable industrial evidence to suggest errors are likely to occur as a result of a complex interaction between multiple workplace and organisational variables (Reason, 1990; Hudson *et al.,* 1994; Wieman & Wieman, 2004; Vincent, 2006). It would be helpful therefore for health care researchers to consider this complex relationship when designing interventions. Without first understanding the multifaceted interplay between organisational variables, many combined forms of intervention may not reach their optimal potential for reducing errors nor will it be known which aspects are successful or unsuccessful in reducing errors.

CPOE: Panacea or quick fix?

Kremsdorf (2005) proposes that in the US in particular, IoM recommendations to utilise technological systems in health care to improve medication safety have led to an 'industry rush' to implement CPOE systems. Implementation of these systems is internationally regarded as the technical 'panacea' to medication ordering errors. Almost half of the studies reviewed measured the effects of implementing CPOE systems on medication error prevalence. Notably, of the nine studies which implemented CPOE systems, four reported that their systems included decision support components such as drug libraries or drug-allergy warnings. Three of these studies report significantly larger decreases in medication error prevalence post-intervention than studies which did not integrate additional decision support features into their CPOE systems (see appendix III for

comparison on intervention success). These findings suggest that basic CPOE systems are less effective without the addition of decision support features. However, this may be an over-generalisation of the effectiveness of additional, costly upgrades to standard CPOE systems. As previously discussed, Bates *et al.* (**2**) showed that while some decision support options significantly reduced MEs compared to baseline measures and standard CPOE implementation alone, other decision support options actually increased the rate of errors. In support of this, Koppel *et al.*, (2005) observed that implementation of some CPOE systems actually facilitated ME prevalence due to poor user-interface design.

These findings raise several important issues. Firstly, basic CPOE systems and decision support features of CPOE do not target the same aspects of human behaviour within an organisation. For example, CPOE alleviates problems associated with handwriting prescriptions such as specification errors (trailing and leading zeros), transcription errors and problems with the design and tracking of paper-based drug charts. Conversely, decision support features are more specifically designed to reduce medication errors by providing interactive advice and assistance during the decision making process. These features consist of constraining or forcing functions such as incorrect drug/dosage/route/frequency prompts and selection of drugs with contraindications or to which patients are allergic (Koppel, *et al.,* 2005). Without first identifying where the problem in any given organisation lies, more sophisticated versions of CPOE systems which incorporate decision support options may represent a needless and costly purchase. By measuring the *respective* effects of standard CPOE system implementation, followed by incremental introduction of decision support options Bates *et al.* were able to tailor their intervention to some extent to meet the specific needs of the clinical setting. Institutions which have implemented CPOE systems should make the distinction between the efficacy of standard CPOE systems and the addition of extra decision support features clearer and should attempt to measure their respective effects.

The total costs of implementing CPOE with additional decision support features at a 500-bed hospital have been estimated to be in the region of $8 million, with continuing maintenance costs of $1.35 million per year (Kuperman & Gibson, 2003). On the other hand, research has shown that implementation of CPOE can significantly reduce the costs of patient care associated with iatrogenic injury in the long term (Tierney, Miller, Overhage, & McDonald, 1993; Evans *et al.*, 1998). However, not all health care organisations will be in a financial situation to afford such an expensive 'up-front investment' (AHRQ, 2001). Although more recent data has not yet been published, a survey conducted in 2002 at the Health Information Management and Systems Society (HIMSS) Annual Conference in the US revealed that while 67 per cent of respondents reported that their organisations were *planning* to implement CPOE, only 21 per cent said that they were currently implementing it[17]. Survey results further indicated a direct relationship between organisation size and implementation of CPOE systems with individuals working at larger organisations (with budgets over $500 million) more likely to report that they were implementing CPOE software compared with individuals working for organisations with annual budgets of less than $100 million. Kremsdorf (2005) argues that implementing CPOE systems has prevented some health care organizations from funding other interventions that may yield more significant benefits and that this in turn has bypassed the core of the patient safety problem and has become a 'distraction from solving the real systemic issues'.

The absence of team interventions

Despite a proliferation of evidence in health care and other high-risk industries implicating the influences of team factors on human error, particularly those involving communication, only one study in this review claimed to implement a team intervention (**1**). Although authors described what

[17] The HIMSS 2002 'Hot Topics Survey' addressed key topics influencing the healthcare IT industry, including disaster preparedness, HIPAA, and patient safety conducted at the Annual HIMSS Conference and Exhibition in Atlanta, Georgia, January 2002.

they deemed to be a team intervention, they did not report that their current team functioning or communication was substandard or the ways in which the intervention was expected to improve these deficiencies. Furthermore, most of the components of this intervention could be more appropriately labelled as decision support improvements (e.g. increasing pharmacist`s availability on the ward, standardising medication dilution charts, computerised drip-rate calculations etc.) rather than improvements of team communication.

This finding raises two important questions. Firstly, why was there only one published team intervention to reduce MEs when team communication is commonly cited in the medical literature as an important precursor of error (IoM, 1999; Helmreich & Schaefer, 1994; Lingard *et al.,* 2005; Sutcliffe, Lewton & Rosenthal, 2004)? Secondly, given the considerable degree of evidence which has attempted to define and categorise team communication (Reason, 1990; Sasou & Reason, 1999; Lingard *et al.,* 2004), reported ways to measure team functioning (Sexton & Helmreich, 2004; Helmreich & Schaefer, 1994; Healey, Undre & Vincent, 2004) and proposed methods of team intervention (Risser *et al.,* 1999; Moray *et al.,* 2002; Lingard *et al.,* 2005; Awad *et al.,* 2005) why was this sole 'team' intervention not really a team intervention?

It could be argued that team interventions are simply too difficult to design and implement. In such an applied field of research, getting teams within and between disciplines together is notoriously difficult due to time and staffing constraints. However, as described above, there is an abundance of research which has developed tools to measure team factors. Alternatively, it might be the case that *any* intervention will have an indirect impact on the way a team functions and so specific team interventions may not be warranted. For example, increasing the availability of a pharmacist on the ward will not improve the knowledge or skills of any other member of the health professional team but will provided valuable decision support for the team as a whole. In this way, the burden of knowledge for the rest of the team will be attenuated which may in turn indirectly improve intra-

discipline relations. Such an intervention could improve team communication by providing a communication channel that did not previously exist. However, it does not necessarily follow that mere addition of another health professional will improve team functioning if that communication channel is ineffective. In fact psychological research has indicated that inappropriately increasing team size can lead to coordination problems due to loss of motivation due to 'dispersion of responsibility' and problems with performance (Latane, Williams & Harkins, 1979; Sheppard, 1993).

Research has suggested that team effectiveness can be affected by multiple inherent design factors such as team size (Guzzo & Dickson, 1996; Temple, 2004), group cohesion (Gully, Devine & Whitney, 1995), organisational context and a match between material resources and task demands (Hackman, 1983). Perhaps it is this plethora of evidence emphasising the highly complex nature of teams and communication strategies which has deterred researchers from implementing such interventions. Knowing there are so many factors which will determine whether a team intervention is warranted and will ultimately be effective could be a deterrent to health service researchers, particularly when the alternative could be as comparatively simple as installing a computer system.

Inconsistent error definitions, detection methods and taxonomies

Classen and Metzger (2003) argue that assessing the effects of any intervention relies heavily upon the use of well defined, standardised and reliable medication safety outcome measures. This review revealed considerable disparity in the measurement of medication error. Of the 19 studies reviewed, only 5 used a standard published definition of medication errors or variations of this definition (NCC MERP). Additionally, all studies measured different aspects of medication errors from prescribing to administration and all nature of medication error types from 'serious' errors and

Adverse Drug Events to dosage errors of Gentamicin. For this reason it is not possible to provide a cohesive synthesis of the material (also noted in the Hoff *et al.* review). In fact, the majority of the studies reviewed provided scant definitions of their outcome measures.

This finding is also consistent with a recent review conducted by Yu *et al.*, (2005) who attempted to identify terms and definitions used by organisations associated with medication safety. From 160 websites searched, only 33 had provided one or more definitions for medication safety terms. Using error scenarios to categorise definitions according to similarity, 25 different terms with 119 different definitions were found. There were eight different definitions of the term *adverse event*, nine versions of the term *error* and twelve definitions of the term *near-miss*. Yu *et al.* concludes that there is an 'imperative need for consensus' on the standardised terminology used to describe medication safety. This nomenclature will in turn enable more meaningful synthesis of medication incident data and assist in the development of improvement strategies. Classen and Metzger (2003) argue that it is this current lack of 'meaningful synthesis' which will continue to hamper strategies aimed at medication safety.

3.5. Conclusions

Apart from one study, there was a distinct lack of any theoretical basis for the implementation of medication error management interventions. Because of this there was very little consistency in the terminology used to describe these interventions. Furthermore, the majority of research reviewed here provided very little empirical evidence to support the need or use of their respective interventions. There is evidently a tendency within health care to perceive technological interventions as the panacea to medication safety and many studies used 'evidence-light' government recommendations to justify their implementation. This reported wave of technological interventions significantly outweighed the use of more psychological improvements involving

teamwork or communication for example. It is argued here that the over-use of technological intervention and under-use of more psychological strategies is due to two reasons. Firstly, Government directives have recommended the increased use of technology in health care to reduce prescribing errors. However, this is often an over-cited recommendation used in lieu of other empirical evidence to support its widespread implementation. Both US and UK government directives (IoM, 1999; DoH, 2000) recommended *multiple* strategies to mitigate medication errors, many of which related to organisational learning from other high-risk industries in terms of primarily identifying organisational weaknesses. This is a directive which is rarely undertaken. Secondly, there is to date no comprehensive or reliable tool available to measure organisation or systems problems in the workplace. As a result organisations are not aware of exactly which interventions are necessary in their institutions and so cannot appropriately design or target required improvements. Organisations have 'panic-purchased' technological systems in order to effect changes to medication safety practices that may or may not have been warranted in their institution. The dearth of psychological interventions such as those which target team functioning for example, may have led to the mistaken belief that such interventions are not beneficial in reducing human error prevalence. Comparatively, the plethora of technological interventions may have led to the over-inflated belief that such interventions are the 'magic-bullet' which will represent the most effectual improvements.

A more 'grass-roots' approach to the identification of organisational and workplace causes of medication error involving workers at the 'sharp-end' of medication practices is clearly necessary. In this way, efforts can be directed towards the development of a diagnostic measure which can proactively identify organisational weaknesses, and target improvement resources before errors can occur. Without a valid and reliable diagnostic tool, it is likely that medication error management strategies will continue to be inappropriately applied and will therefore not achieve their optimum potential.

CHAPTER 4
LATENT CAUSES OF MEDICATION ADMINISTRATION ERROR: AN EXPLORATORY STUDY

4.1. Introduction

The main aim of this research was to respond to a call from numerous research and government agencies (see chapter 1, section 1.4) for NHS research to learn from error management techniques and theory employed in industries which have attempted to proactively identify organisational and workplace failures before they result in error. This research is intended to complement existing retrospective incident analysis systems and approach error management from the opposing direction (see chapter 2, Figure 2.3). A systematic research review was conducted in chapter 3 to identify patient safety research which had employed organisational or workplace interventions to reduce medication errors. This review revealed a research base which was largely atheoretical and unsystematic. Furthermore, this evidence base provided generally inconsistent use of terminology to describe interventions, outcome measures and subsequent findings. It is hypothesised that a more 'grass-roots' approach involving health care professionals at the 'sharp-end' of clinical practice will establish a clearer understanding of these latent failures and the ways in which these problems manifest in a clinical setting.

Using Reason's *organisational accident model* (Reason, 1997) as a framework for understanding the interplay between organisation factors, workplace conditions and unsafe acts, this stage of the research employed several qualitative methods to explore the latent predictors of medication administration errors in secondary care. Research has suggested that qualitative methods are useful for capturing the factors behind human error and system failure, particularly in health care research (Runciman, 1993). Armitage (2004) argues that such methods are especially useful in the exploration of phenomena which is often 'complex, contextual, and of both a physical and psycho-

social nature' as is the case in drug administration errors. This qualitative study was conducted in three phases. Since the aim of this study was to explore the causes of errors committed by nurses during the administration of medication, the first study phase was to conduct observations of nursing activities, with particular emphasis on medication rounds (where nurses take a trolley stocked with medication around the ward to give patients their prescribed medication). Phase two involved reviewing formally reported drug errors to identify potential error causes and familiarise the author with the types of error which occur in the process of medication administration. The third phase of this study was to explore the perceptions of error causation of health care professionals at the 'sharp-end' of clinical practice and senior management at the 'blunt-end'[18] of one NHS Trust.

To meet the research needs of the collaborating hospital (Bradford Teaching Hospitals NHS Foundation Trust), the decision was made to focus on medication administration errors (made by nurses) as opposed to *all* medical errors which was the original intention. There is substantial international evidence in support of this decision indicating adverse events resulting from medication errors to be a leading cause of iatrogenic injury (Malpass, Helps & Runciman, 1999; Vincent, Neale & Woloshynowych, 2001; Leape, Bates & Cullen, 1995). Furthermore, a recent study of the medication delivery process from transcription to administration has revealed errors in the administration of medication to be the most frequent (Marshman *et al.*, 2006).

4.2. Method

Observations of nursing activities (study 1a), semi-structured interviews (study 1b) and a review of formally reported incidents (study 1c) were conducted in this study. The main aim of conducting multiple qualitative methods in this way was to provide a multi-dimensional perspective on the organisational and workplace causes of medication administration errors (MAEs). Wilson and

[18] See chapter 2 - Reason's organisational accident model for full explanation of these terms

Hutchinson (1991) argue that combining two or more qualitative methods provides important breadth and depth required in applied research and decreases potential biases or disadvantages of independent methods (see section 4.3.2 for discussion of such disadvantages).

It was agreed between the researcher and the collaborating institute that qualitative data collection would be conducted within general medicine. General medicine provides acute medical services for adults across a wide range of specialties. This clinical area represented a non-specialised clinical environment with a wider variety of patient complaints, medical staff and involved a broader range of clinical procedures than more specialised clinical areas (e.g. neonatal care, obstetrics etc.). It was anticipated that findings from research in this environment would be more generalisable to other clinical areas. Carrying out the exploratory stage of the research in a more specialised field of medicine may have lead to the development of an error management tool which was limited to that field. Ethics approval was obtained from NRES[19] (formerly COREC[20]) and research /governance approval was obtained from the Bradford Teaching Hospitals NHS Foundation Trust's R&D department before conducting any of this research.

Each of the three phases of study 1 will be detailed under respective subheadings **in the order they were conducted** and the combined results and discussion of all three stages will be discussed in section 4.4.

4.2.1. Study 1a: Ward observations

Aims

Research has highlighted that an ethnographic approach to the observation of clinical practice can be most effective (Taxis & Barber, 2000; Dixon-Woods, 2003; Armitage, 2004), particularly when

[19] National Research Ethics Service
[20] Central Office for Research Ethics Committees

the phenomena of interest is particularly sensitive as was the case in this study (Savage, 2000). Observations were conducted by shadowing nursing activity using an ethnographic approach for the following reasons;

- To acquaint the novice observer with the research environment and occurrences and patterns of nursing behaviour thus enabling the development of a more 'environmentally-aware' interview schedule.
- As a complementary methodology to assist in the interpretation of medical and nursing terminology used in subsequent interviews.
- As a potential source of data exploring the *workplace*[21] causes of error.
- To facilitate an open, honest and trusting relationship between the researcher and nursing personnel which would be vital in later interviews when staff would be asked to share potentially sensitive information.

Participants

Using purposive sampling, matrons responsible for three general medical wards in two different hospitals were recruited and agreed for their wards to be observed intermittently over a period of three months. Random sampling of wards was not possible since there were only three general medical wards in the collaborating hospitals and all three were recruited. It was agreed between senior nursing staff and the researcher that any nurse could potentially be observed performing any nursing activity provided that observation did not encroach upon patient rights to privacy. This decision was at the nurses' discretion. The nursing team were instructed by the observer beforehand that if they felt uncomfortable being observed they were free to question the purpose of the observation and could ask not to be observed if they felt uncomfortable. It was emphasised that the observer was not watching for mistakes, but the ways in which their local working conditions

[21] It was considered unlikely that *organisational* causes of error would be directly observed since these are more likely to operate at a much higher level of the health care system, not at the ward-level.

affected their ability to carry out certain tasks, particularly with respect to the administration of medication.

Design

Delamont (2002, pp. 130-131) recommends that during ethnographical observations 'an initial period of relatively unfocused watching is essential' which will enable the observer to 'tune in' to cultural behaviour. This is commonly referred to in qualitative research as the 'descriptive phase' (Werner & Schoepfle, 1987, pp. 262-264). Assuming the role of 'complete observer' (e.g. remaining unobtrusive in the background throughout the observation in an entirely non-participatory role; Creswell, 1994), all three general medical wards were 'informally' observed over a period of three hours each (nine hours in total) to meet this objective. Field notes were taken for the duration of these observations before more focussed observations were carried out.

Nine semi-structured observations were then conducted to explore several facets of nursing care and the clinical environment under the following categories. Werner and Schoepfle (1987) refer to this as the 'focussing phase'. These categories were derived from the first round of 'unfocussed' observations and also health care research which has found similar observational categories useful in obtaining a well-rounded perspective of nursing behaviour (Wolf *et al.,* 2006).

1. **Medication administration behaviour**. This included any nursing activity, disparate from other patient care activities, which involved the preparation or administration of medication. Any actual or potential drug errors observed during the shift were also recorded with permission from the nurse involved (although in reality this rarely happened during observations). This information was often in the form of volunteered information.
2. **Local working conditions**. Any immediately obvious aspects of the environment which had the potential to increase the occurrence of drug errors (e.g. noise levels, design of equipment or patient documentation, interruptions or distractions during complex tasks etc.)

3. **Team functioning**. Any particularly salient or noteworthy aspects of team behaviours which were judged particularly effective or defective. This included communication within nursing teams, with other health professionals and with patients.
4. **Administrative tasks.** Tasks which required nurses to complete paperwork.
5. **Comments**. Any unsolicited comments made by nursing staff to the observer regarding patient safety, their experiences of making errors, and their perceptions on error causation reduction and/or prevention. Although the researcher did not instigate these conversations, nurses often approached with information they thought was important to providing safe patient care. Permission was gained from staff before documenting any of these comments.

Each ward was observed once a month over a three month period (not including the initial 'unfocussed' observation), once on a morning shift (7.30am-1.30pm), once on an afternoon shift (1.30pm-7.30pm) and again for a period which straddled both morning and afternoon shifts (10.30am – 4.30pm) so that a comprehensive overview of that wards daily activities could be obtained. In this way it was possible to observe three of the four medication rounds conducted daily. Nursing staff felt that it was not appropriate to observe night shifts since there was very little ward activity, there were fewer staff to observe and patients should not be disturbed while sleeping. Each period of observation was approximately 6 hours, therefore each ward was observed for 18 hours in total (54 observation hours in total across all 3 wards). All wards observed were traditional Nightingale wards on which between 8 and 13 patient beds were positioned against both the left and right hand walls in a large rectangular open-plan room (as opposed to more modern wards which are divided into several smaller bays and private rooms). All observational data was recorded in the form of handwritten field notes.

Procedure

For initial unfocussed observations, the observer reassured nursing staff on arrival on each ward that the observations were simply for the benefit of the researcher to familiarise themselves with the

ward environment and nursing activities. During focussed observations, nurses were also informed of the categories of behaviour which were being observed. Participants were advised that observational field notes would also be taken. Particular emphasis was placed on assuring staff that the researcher was independent from the hospital, did not represent the hospital management and was not auditing or assessing their performance. This was felt necessary to ensure staff felt at ease and were more likely to behave in a way they normally would, rather than be on their 'best behaviour', a common problem in observational methodologies. It was also emphasised that all data obtained (either directly observed or disclosed by staff members) would be treated confidentially and any personal details which could make any individuals identifiable would not be recorded on any field notes which would not be made available beyond the observer. Nurses were advised of their right to refuse to be observed without giving reason for doing so. However no staff member on any of the three wards refused to be observed. It was considered extremely important for the observer to avoid becoming a potential *cause* of errors and so the observations were designed to be unobtrusive as possible. Observations were generally conducted (approximately 70 per cent of the time) from a central nursing station in the middle of each ward. Field notes and reflections of observations were transcribed after each observation and were entered into the qualitative analysis package NVivo for the purposes of organising and reorganising the data in subsequent analysis. These notes were analysed for potential workplace causes of drug errors using thematic content analysis.

4.2.2. Study 1b: Semi-structured interviews

Participants and recruitment strategy

To obtain a broad perspective of the latent causes of MAEs, interviews were conducted with both sharp-end nursing staff and senior managers at the blunt-end of the health care organisation. In doing this it was anticipated that factors across the whole organisation from management decisions

and organisational processes at the blunt-end to local working conditions and the nature and occurrence of unsafe acts at the sharp-end of the organisational would be explored (see chapter 2 for Reasons organisational accident model).

Twelve senior NHS managers from the collaborating NHS Trust were approached and invited to participate in interviews. Managers were sent invitation letters and participant information sheets describing the purpose of the research and interviews (appendices IV and V). Eight managers agreed to be interviewed. Participants were the Director of Clinical Governance, Head of Nursing Practice Development, Chief Nurse, Head of Risk, Clinical Director, Clinical Risk Manager and two Patient Services Managers. It was anticipated that senior NHS managers would be in a significantly better position than nursing staff to comment on latent causes of errors at the uppermost end of the organisational system due to the nature of their responsibilities.

The two matrons in charge of the three general medical wards previously recruited for observations were revisited and asked if their wards would like to participate in interviews. The matrons agreed that two of the wards could potentially be recruited but the third ward was due to close in the coming months and so was not approached. A recruitment strategy was agreed between the researcher and the matrons that only qualified nurses who could discuss the potential causes of drug administration errors would be invited to participate. Matrons preferred to nominate those nurses they felt were the most experienced, knowledgeable and articulate. Letters inviting nurses to participate in the research explaining the purpose of the study and the ways in which interviews would be conducted were sent to 25 nurses from the two recruited wards (appendix VI). Since it was agreed that interviews would be conducted during nurses' lunch breaks, lunch was provided as an incentive for interviewees. Furthermore, matrons, senior managers and the ethics committee agreed that £50 could also be offered as a prize draw incentive which is common practice to increase recruitment. Eleven nurses agreed to be interviewed and comprised one of the two

matrons, three Sisters, three senior staff nurses, three staff nurses and one student nurse who was due to qualify the same year and had been involved in the administration of medication on several different types of clinical specialty.

Design

Interview questions were designed to elicit participant's views on the causes of medication errors. This subject was deemed particularly sensitive, especially within the nursing group and so vignettes containing hypothetical error scenarios were designed to complement interview questions. Research has indicated that using a vignette technique enables participants to disclose views which they feel may be potentially threatening using reference to a 'non-personal story' (Rahman, 1996; Hughes, 1998; Schwappach & Koeck, 2004), particularly in nursing research (Gould, 1996; Gott *et al.*, 2004). In this way, participants were able to choose the degree to which they drew upon their own experience of error.

Error Scenarios

Liaising with the two recruited matrons, eight hypothetical examples of medication error were developed (appendix VII) from the medical literature (e.g. NPSA[22], AHRQ[23] and other patient safety websites), the matrons own extensive clinical experiences and the researchers experience of working in a hospital according to the following three parameters;

- Written in clear generic language which did not require significant clinical practice to understand. This was considered important since vignettes should be equally interpretable for both qualified nurses and also managers who may not have a clinical background.
- One paragraph in length which could be read in approximately five minutes. It was decided that vignettes should provide sufficient detail to describe one or more errors in the process of medication delivery.

---

[22] See foresight patient safety training - (http://www.npsa.nhs.uk/patientsafety/improvingpatientsafety/humanfactors/foresight/)
[23] Agency for Healthcare Research and Quality, USA

- No information presented which could potentially bias interviewees' responses, such as patient outcome[24] or stated potential latent variables. In this way, participants were free to generate their own hypotheses on causation rather than ones which were already stated or biased by the impact of the error on the patient.

The interview schedule

The interview schedule was semi-structured and based around Reason's organisational model of human error (see chapter 2). Questions aimed to guide participants through each stage of the model in a retrospective discussion of the vignette from the unsafe acts immediately preceding it to the local working conditions which may have triggered these and finally on to latent aspects of the organisation which may have brought about these conditions. Interestingly the interview schedule had to be amended due to the considerable difficulty both nurses and senior managers had interpreting organisational terminology such as system, human factors, latent, and local working conditions. Revised interview questions took a less structured approach and invited participants to discuss causes of medication errors from individual, environment and organisational perspectives (see appendix VIII for interview schedule). One benefit of this type of schedule was to enable participants to understand the hierarchical nature of error occurrence and to explore the ways in which downstream variables might impact on working practices.

Procedure

Management interviews were conducted during working hours in the manager's offices or suitable conference room in Bradford Royal Infirmary and lasted between 20 and 60 minutes depending on how much time the interviewee had to spare (and how much they wanted to say). These interviews were not incentivised since the ethics committee agreed that it was the responsibility of these

[24] Research has shown that the more severe the consequences of an error are, the more likely responsibility will be attributed to the individual(s) involved, regardless of the actual cause of the error. This is known as severity bias.

managers to be involved in this type of patient safety research. Interviews with recruited nurses were conducted during working hours, usually during the nurse's lunch break and lasted between 30 and 90 minutes depending on how much time they could spare from the ward. Interviews were conducted on the ward in a private room away from other staff members[25].

All interviews were recorded using a digital voice recorder and the researcher informed participants that notes would be taken during the course of the interview to aid understanding of the issues and ask relevant follow-up questions. As per BPS[26] ethical guidelines, all participants were assured that any information discussed would be anonymous and confidential and that only the interviewer would have access to transcriptions and audio recordings. Participants were advised of their right to refrain from answering any question, to refuse to be recorded, to withdraw from the interview altogether or withdraw their data from the study at any point without providing reason for doing so. The interviewer read participants the information sheet which they had previously received during the recruitment phase in case they had not read it to ensure they understood the purpose of the study. Participants were asked to sign a consent form before commencing interviews to ensure they understood their rights (appendix IX).

The first few minutes of the interview were spent introducing participants to the idea that errors are inevitable. This was considered especially important given that a common viewpoint in health care risk management concurs that errors should be avoided at all costs and individuals are often blamed for their mistakes (Reason, 2000; Spath, 2000; Vincent, 2006). In order to introduce the idea of error inevitability, participants were given a sheet of paper with a list of everyday errors which people execute whilst driving (Reason *et al.*, 1990; Parker *et al.*, 1995). Since driving is behaviour which most people can identify with it was thought that examples of these types of errors would

---

[25] Although it would have been preferential to conduct interviews away from the hospital setting in a neutral location to encourage a more open and frank discussion of such sensitive issues (Elwood & Martin, 2004), this was not possible since most nurses reported they would not participate if they had to travel elsewhere to be interviewed.

[26] British Psychological Society

encourage interviewees to feel more confident that error behaviour is common (appendix X). Furthermore, it was intended that discussing errors in the first part of the interview would act as an effective 'ice-breaker' before the topic moved on to more sensitive issues.

Once the ice-breaker reached a natural end, participants were advised that the interview and recording of the interview would commence. They were told that they would be given a hypothetical example of a medication error and that they should try to think about what might have been the main causes of this error. Vignettes were selected randomly by each participant by choosing a number from one to eight representing each vignette. Once one participant had selected a particular vignette it was removed from the next interviewee's selection until the 8th interviewee had only one vignette to choose from. This ensured that all vignettes were discussed twice (since there were 19 interviewees). In this way it was hoped that a broad discussion of variables would be generated rather than a discussion of the same vignette. Participants were given a few minutes alone to read the vignette (it can be difficult to take in written information when being watched) and were advised they could make notes if they wished, while the interviewer left the room. Participants were advised they could use examples from their own experience of error to illustrate their views of causation, but were advised they could simply use the scenario presented to them if they were not happy to do this. Once interviewees were ready, audio equipment was started and questions began.

A printed diagram of Reason's organisational model was given to participants who struggled to understand the concept of error causation as an aid to understanding the information the interviewer was trying to elicit (appendix XI). No latent causes of error which might bias responses were detailed within this diagram. If interviewees continued to find the interview questions difficult, it was decided that as a further aid to understanding the error causation trajectory, participants could be provided with examples from the Ladbroke grove rail crash incident analysis (see chapter 2, section 2.5.1.2.). These examples were intended to give interviewees an idea of how multiple

organisational factors can impact on local working conditions and ultimately unsafe acts. Examples were not considered to be those which would be evident in secondary care and so would not bias responding (e.g. bright sunlight obscuring danger signals) but illustrate what was meant by 'causal factors'.

Once interviews had reached a natural end or interviewees had to leave, participants were asked if they had any questions about the research and whether they felt the experience of discussing errors had been a positive one. All interviewees said they felt it had been a positive experience, particularly the nursing staff who said they were rarely given the opportunity to be so open.

4.2.3. Study 1c: Incident report review

Incident reporting in the NHS: A brief overview

The system of reporting incidents in the NHS (in 2005 when this study was carried out) relies upon NHS staff documenting events which are classified as 'significant' under a number of headings such as medication-related events, patient complaints, slips and falls, issues with medical records, delayed diagnoses and equipment problems. Incidents are not reported anonymously and staff members are required to provide their own details and those of the patient concerned with a brief description of the event they consider to be problematic. Completed incident forms are reviewed by immediate nursing managers and can either be dealt with at a ward level if the incident is relatively minor or can be investigated by the risk management department. The risk management team categorise incidents based on the severity, likelihood of recurrence and implications on organisational costs and patient safety outcomes of the incident based on a 'traffic light' system (from red-'severe/highly likely/high costs' to green-'minor/unlikely/low costs'; National Patient Safety Agency (NPSA), 2007) and collate and submit all reports to the NPSA.

Aims

To obtain a well rounded perspective of the types of drug errors occurring and their respective causes, a sample of formally reported medication incidents occurring in general medicine were reviewed. Since observations were and interviews would be conducted within general medicine, it was decided that incident data should also originate from the same source to ensure relevance and appropriateness of reports. Only errors which were associated with medication administration deemed to directly involve patient care (some drug errors did not directly involve patients) were selected for review. This review would serve two additional purposes;

- To corroborate latent factors generated during interviews
- To identify any factors *not* identified during interviews (see section 4.3.1 for descriptions of data triangulation for completeness and confirmation).

Data Source

The risk management department of the collaborating hospital agreed to allow the researcher access to a large dataset of medication incidents formally reported between 1999 and 2003. Consent was given to review a sample of these reports on the provision that any staff or patient details which would make identification possible were removed before review. It was initially decided that only a small sample of incident reports would be analysed (n=30) since this was not the main source but was intended as a complementary source of data with which to achieve a holistic perspective of medication errors. After preliminary review of these incident reports it became apparent that very little information was actually recorded (usually only two or three sentences); many reports containing extensive amounts of abbreviation and sparse detail regarding the incident itself or potential causation which made any in-depth qualitative analysis virtually impossible. Early analysis revealed it was only possible to hypothesize error causation from approximately 20 per cent of these reports. The decision was subsequently made to increase the number of incidents to be reviewed. A 50 per cent random sample of *all* medication incidents reported on general medical

wards between 1999 and 2003 in 2 hospitals was provided (n=155). From this sample, 18 reports (12 per cent) were excluded because they did not contain enough detail to be understandable or to be analysed and 8 (5 per cent) because they did not directly involve patient care. A further 22 (14 per cent) were excluded because they did not comprise an actual preventable error (e.g. patient had an adverse reaction to an administered drug which they had never reacted to before). The remaining 107 medication incidents were entered into the qualitative analysis package NVivo and their content analysed for potential organisational factors and workplace conditions of drug errors using thematic content analysis.

4.3. Data Analysis

Handwritten raw observational field notes, recorded interviews and incident reports were all transcribed in Microsoft Word and saved in the qualitative analysis package NVivo7. The method used to analyse data will be described in more detail in section 4.3.3 although in short an inductive approach was taken. Inductive analysis uses the data to generate ideas as opposed to deductive reasoning which uses data to 'confirm or negate pre-held ideas and hypotheses' (Thorne, 2000). Given the exploratory nature of the study, an inductive approach was considered the most appropriate method and is one that is often employed within nursing research (Appleton & King, 1997; Thorne, 2000; Upenieks, 2002). When selecting the best method of data analysis there were a number of considerations including the best possible use of the data to achieve a holistic perspective of error causation and also how the data sources should be combined in order to accomplish this.

4.3.1. Triangulation

One particular methodological consideration was how the three data sources should be combined to obtain an unbiased, meaningful and realistic perspective of medication error causation. One particular method of combining data sources in order to strengthen the design and reduce

subjectivity in interpretation of the data is known as triangulation (Denzin, 1970; Kimchi, Polivka & Stevenson, 1991). Although there are several types of triangulation, methodological triangulation involves the combination of more than one research method or data collection technique (e.g. structured instruments, observations and interviews etc; see Thurmond, 2001). In this way, different methods of data collection can tap different aspects of the phenomenon being studied to obtain a multi-dimensional viewpoint (Knafl & Breitmayer, 1989).

However, Shih (1998) suggests it is important to first identify the rationale behind triangulating data sources. He identifies two disparate purposes; confirmation (Knafl & Breitmayer, 1989; Webb *et al.*, 1981; Anderson, 1997) and completeness (Jick, 1979; Fielding & Fielding, 1986; Murphy, 1989). He argues that while triangulating data can be conducted solely for the purposes of validating findings as is often cited in qualitative literature, it should also be about obtaining the most holistic picture possible. Achieving this holistic viewpoint has been noted as a particular priority within nursing research which is notorious for various sources of bias (Jick, 1979; Fielding & Fielding, 1986; Murphy, 1989). Shih (1998) warns that when triangulating data sources for the purpose of completeness health care researchers should not anticipate multiple sources of data to confirm one another. He recommends that multiple data collection strategies should be selected and combined because of their 'inimitable contribution' towards addressing the research question. A brief synopsis of each data source collected is presented to determine whether data sources were likely to converge; thus dictating whether triangulation should focus on validating findings or using data sources to obtain a complete picture of medication error causation.

4.3.2. Methodological disparity

Although the combined findings will be described in more detail in section 4.4, the nature of both incident reports and observations will be described along with their respective similarities and disparities with interview data which represented the richest data source.

Incident reports

Medication related incident reports were by far the 'thinnest' qualitative data source collected. It was anticipated that these reports would represent the link between observations and interviews highlighting the unsafe acts and local working conditions preceding errors which were difficult to observe and those which nurses may have been particularly uncomfortable discussing during interviews. Unfortunately a number of factors meant that this source of data offered very little to the understanding of medication error causation. Firstly, the information provided on incident forms was minimal at best and often comprised only one or two sentences stating only the bare facts of the incident. For example;

> '...patient given stemetil on 5 occasions by 3 nurses. Brought to ward sister's attention by SN (Staff Nurse) who was "extremely upset on the telephone". Dr informed of incident and discontinued the drug... '

In this case and for more than 70 per cent of the other incident reports it was difficult to infer any potential cause of the error from the details provided. This is due in part to the wording of incident reports which specifically requires those completing the form to provide the facts without surmising causation. On the rare occasions that error causation was implied in reports this was often a blame reference to another department or health professional in order to assuage responsibility. For example (most relevant section is highlighted);

> '...enrolled nurse T, allowed a nurse who has not received a PIN number or been assessed as competent to give medications without supervision. Nurse who gave the medication has been told by myself and other members of staff that she must not do it alone. **Cause**: nurse had been told on previous occasions that she should not do medicines unsupervised, does not recognise her limitations in practice…'

This finding is particularly relevant and in terms of assuaging personal responsibility, was witnessed first hand. During two of the nine observation sessions, this researcher observed senior nursing staff 'coaching' more junior staff in their wording of incident reports. On one particular occasion a student nurse wrote three drafts of incident report which the Sister of the ward reviewed

and edited before submitting a 'polished' (and considerably shorter) version four. When asked why this was necessary both members of staff insisted 'you have to be careful what you put so you don't get sacked'.

Combined with the paucity of details provided on incident forms, the way in which errors were reported revealed significant biases which made any reasonable attempt at causation analysis near impossible. In the broader scheme of error management, such a finding suggests that opportunities for organisational learning are being lost. To some extent this was borne out in other incident reports which stated the cause and subsequent action taken after particular incidents;

> '...cefotaxime 2 grams was accidentally given to the patient, noticed 12 hours later. Doctor informed. Statement reads – "on night duty I had gone to assist SN (Staff Nurse) and given the task of administering IV antibiotics, I was given the correct details but gave the medication to the wrong patient. I make no excuse for the mistake made by me, despite the fact that we were busy".
> **Cause**: did not follow correct checking regime.
> **Action**: SN has been stopped from further IV procedures and an education package has been developed for her based on revisiting the medicines policy...'

Although in this incident the immediate cause of the error was explicitly stated by the nurse (that she was excessively busy), the action taken is not reflective of this, ignores the cognitive constraints cited by the nurse and recommends a course of action which is not consistent with the cause – education. A similarly narrow approach to error management is evidenced here;

> '...I have given the patient the wrong dose of the correct insulin. I (staff nurse) read a fellow patient's drug chart, and had given her 60 units by mistake. I informed senior nurse and house officer immediately. Relatives informed.
> **Cause**: staff nurse has had two separate insulin prescriptions for different patients and was administering the insulin one after the other – I reminded her that this is unsafe practice (matron), I said all drugs should be administered on an individual basis.
> **Action:** suggested she should work through one drug chart at a time, complete this and move to next, double check her IV drugs and insulins. Staff nurse T said she has already learned from the error...'

As with the previous incident report presented, this excerpt suggests that learning has taken place at an individual level but makes no reference to latent preconditions of this mistake and so will more than likely occur again due to inevitable cognitive constraints. Interviews yielded significantly more information on error causation which was simply not evidenced within incident reports.

Observational data

There are several important points of interest regarding the observational method of data collection. Firstly, the researcher was not clinically trained as is usually the case for this type of observation in health care settings. For this reason, it was not always easy to know what behaviour should be observed and when something of importance had occurred. For example, when initially observing nurses conducting their medication rounds, it was common for nursing staff to be interrupted by various other health professionals, patients and visitors. Although it seemed obvious to the observer that distraction during complex tasks would place considerable cognitive strain on an individual, it was not immediately apparent whether such practices were detrimental to medication administration errors since the researcher was not trained to detect errors. It was also noted that this particular method of collecting data was especially difficult because there were so many nurses and activities occurring simultaneously and for a detailed and reliable observation technique to be employed (for recurrent practices) a team of observers was probably necessary.

One particular criticism of this type of overt naturalistic observation approach and inextricably linked to the observers non-clinical background is the possibility that 'best' behaviour rather than natural behaviour was being observed. In the early stage of observations (the descriptive phase) several nurses asked the observer how she thought they were 'doing' and seemed concerned their behaviour was being audited; a regular practice within NHS hospitals. This was indicative that at least to some extent their behaviour was not as reflective of their everyday practice. It is possible that this behaviour was a display of the Hawthorne effect (Landsberger, 1958); participants

behaving positively because of the extra attention being received during observations. Sim and Wright (2000) suggest that this form of observer reactivity can be mitigated if observers maintain an unobtrusive position in the field of observation. Furthermore, research has shown that spending time in the research setting before formal observations will allow participants to become habituated to the presence of the observer and behaviour will become more natural over time (Bogdewic, 1992; Smith, 1996). It is hypothesised that since both of these recommendations were carried out, observed behaviour was an accurate reflection of usual behaviour.

Although observations proved useful in observing workplace problems and unsafe acts at the ward level, it became apparent that it was difficult to directly observe the manifestation of organisational failures. For example, nurses were observed administering medication without first checking patients I.D. wristbands. However, it was not observable why such a workplace condition should occur. It is possible that this was a cultural problem, i.e. this deviation from standard practice had become culturally accepted over time and nurses hardly noticed they were not conforming to standards which were not reinforced. Alternatively, it is possible that nurses were consciously aware of their violation, but felt justified in this behaviour due to workload or other related issues such as staffing levels. Since it was vital for the observer to remain impartial during observations to gain the nurses trust and belief that they were not being assessed, it was not always possible to follow up observations with questions about their reasons for behaving in any particular way. Nevertheless, this information was useful in subsequent interviews where follow-up questions were more appropriate.

Data obtained during observations proved invaluable for subsequent interviews both in terms of aiding the interviewers understanding of what was being discussed and also helping to generate well-informed questions. When presented with behavioural observational evidence during interviews, interviewees were often surprised they had not really noticed the behaviour in question

or that they had never really thought of it as problematic. For example, it was observed that doctors were starting ward rounds at 9am while nurses were in the middle of drug rounds. Doctors were then observed asking nurses to break off from the drug round to attend to their needs before returning to the task. It was obvious to the observer that such poor scheduling of two major ward activities would be likely to increase the likelihood of nurses making a subsequent error in medication administration; the impact of such poor timing of activities was rarely elicited without prompting by nurses during interviews. However, when presented with the observation of this scheduling clash during interviews, almost all nurses agreed it was a major problem but one which had been accepted practice for so many years they didn't immediately think of it as a patient safety concern. By using such observations to complement interviews nurses were able to suggest a multitude of reasons this problem might have occurred including communication problems, role insight and the perceived hierarchical relationship between nurses and medical staff.

Although observations were useful to a certain extent in supplementing interviews and alerting interviewees to potential workplace conditions which might promote error, the data obtained from incident reports was generally less comparable. The difference in quality and quantity of the information obtained from the three sources made it clear that it was not possible to use one source of data to validate the other. For this reason, it was decided that data would be combined for the purposes of obtaining a holistic perspective of medication errors rather than to validate each data source. The following thematic analysis technique was applied to each data source in turn to explore error causation.

4.3.3. Data analysis technique

Several qualitative analysis methods were considered for the interpretation of data collected in this study. It was a broad aim of this analysis to identify multiple commonly discussed preconditions of medication error from the data. Thematic content analysis (TCA) was deemed the most appropriate

method of achieving this. Research has described this form of inductive analysis as a method of interpreting qualitative data through the 'systematic classification process of coding and identifying themes and patterns' (Hsieh & Shannon, 2005; pp. 1278). Qualitative researchers have also recommended TCA for the analysis of studies with substantial pools of data, as was the case in this study (Patton, 1990; Weber, 1990; Berg, 2001). As study 2 would involve the development of safety indicators which represented each of the themes identified in this analysis, it was important that these themes were valid and reliable. TCA is a method which is largely thought to achieve this level of methodological rigour because it involves the 'application of set procedures for processing data using step-by-step rules' (Mayring, 2000, pp. 5).

In a recent and particularly comprehensible research paper, Braun and Clarke (2006) outline the 'rules' of applying TCA. The authors recommend TCA is conducted in five stages of analysis. Notably, examination of the qualitative methodology literature revealed that there were many articles which had specified step-by-step guidelines. However, there was a general agreement that the five stages proposed by Braun and Clarke would be sufficiently rigorous to achieve valid and reliable interpretation of the data (Tesch, 1990; Miles and Huberman, 1994; Schilling, 2006).

Phase 1 – familiarisation with the data

Braun and Clarke recommend it is vital for researchers to immerse themselves in the data before analysis takes place, which usually involves repeated reading in an *active* way, initially searching for meaning and patterns of key words. They recommend this is most effective when the sample size is moderate and any field notes or verbal data are transcribed into a written form by the researcher (as opposed to employing a transcriber). In this study there were a total of 281 pages of transcript which comprised 107 incident reports (29 pages), 9 observational field notes (23 pages) and 19 interviews (229 pages; 74 for senior managers and 155 for nurses). All data was transcribed verbatim by the researcher.

Phase 2 – Generating initial codes

Braun and Clarke recommend researchers next generate a list of initial ideas of interest across the entire data set and extract data excerpts which represent these ideas. These excerpts are then condensed into initial 'codes', defined as 'the most basic segment, or element, of the raw data or information that can be assessed in a meaningful way regarding the phenomenon' (Boyatzis, 1998, pp. 63). Table 4.3.3.a illustrates how this was done in this study. Braun and Clarke recommend that researchers should aim to code for as many different themes or patterns as possible, even if a code seems like a 'one-off' – it may be useful during later stages of analysis. In doing this, more than 150 different codes from all three data sources were generated in this study.

Table 4.3.3.a: Phase 2 generation of initial codes

Data source	Data excerpt	Code
Incident report	Patient prescribed insulin on insulin chart (attached), 81 units prescribed and nearly given – should have been 8.1 units.	1. Inappropriate design of drug charts 2. Legibility of handwriting
Observation	Staff Nurse trying to locate a certain piece of information in patient medical notes but loose sheets keep falling out. She places them back anywhere in the notes - not necessarily where they came from.	1. Design of medical notes 2. Procedure of filing
Interview	On a drugs round, if a patient is receiving 4 or 5 different drugs, all the information you have to process on that drug round are things like; diagnosis, the medications that you're actually giving them - is it appropriate for them, drug interactions, sometimes you'll go to the drugs trolley and there will be 2 items missing, so you're broken off, close the drugs trolley, go all the way back to the clinic, retrieve the items you need or you might not stock a drug so you will have to go and order it from pharmacy…so your level of concentration deteriorates just in that one correct drug administration.	1. Ability to concentrate on one task 2. Complexity of patient condition 3. Availability of stock items in drug trolley 4. Ward design

Phase 3 – Searching for themes

This phase begins when all data sources have been reviewed and initial codes extracted. In this way, the analysis considers how codes might be distinct or combined to form hierarchical themes. This phase also considers the different levels of themes. Since there were so many codes in this study it is beyond the scope of this chapter to fully present the way in which this was done. As an illustrative example, Table 4.3.3.b indicates how this was done for 2 themes;

Table 4.3.3.b: Transforming related codes into themes

Code	Sub-theme	Theme
1. Inappropriate design of drug charts 2. Legibility of handwriting 3. Excessive paperwork 4. Inter-departmental information sharing channels 5. Over-reliance on informal verbal information sharing 6. Problems with multi-cultural information sharing	**Communication**	
1. Information overload 2. Being interrupted mid-task by other health professionals 3. Task complexity	**Mental workload**	**Workload**
1. Ability and time to plan ahead 2. Unpredictable environment	**Workload planning**	

Phase 4 – Reviewing themes

Once 'candidate themes' have been developed, phase 4 of the analysis involves revisiting the data in two stages. Firstly, coded excerpts are reviewed to ensure that they represent the theme or sub-theme category they were initially assigned to. This allows the researcher to check whether the

proposed theme or sub-theme makes sense. It may be the case that sub-categories or codes which were initially assigned to a particular theme may not fit the overall theme very well or may even comprise a standalone theme. Once this level of analysis is complete, original data sources are reviewed to ascertain firstly whether themes 'work' in terms of the entire dataset and secondly to code any additional excerpts which may have been missed during phase 2 but which add strength to proposed themes. Table 4.3.3.c indicates how this was done in this study;

Table 4.3.3.c: Refining themes and sub-themes

<table>
<tr><th>Code</th><th>Sub-theme</th><th>Theme</th></tr>
<tr><td>1. Inappropriate design of drug charts
2. Legibility of handwriting
3. Excessive paperwork</td><td>Written Communication</td><td rowspan="2">Communication</td></tr>
<tr><td>1. Inter-departmental information sharing channels
2. Over-reliance on informal verbal information sharing
3. Problems with multi-cultural information sharing</td><td>Verbal Communication</td></tr>
<tr><td>1. Being interrupted mid-task by other health professionals
2. Being expected to perform tasks which are not within your role responsibility</td><td colspan="2">Inter-professional regard/Role insight</td></tr>
</table>

This form of refinement allows themes to be collapsed into overarching themes or segregated into smaller sub-themes. As Table 4.3.3.c illustrates, it became evident at this stage that the theme *communication* could be subdivided into two distinct forms; *written* and *verbal*. Similarly this phase of analysis enabled a clearer picture of the code 'being interrupted mid-task by other health professionals' which had previously been assigned to the theme *workload* since this code was often discussed in terms of its impact on planning workload. After careful review and in the context of the whole dataset, it was decided to move this code into the theme *inter-professional regard/role insight* since it became clear that this behaviour was more a complex relationship between health

professionals who were not making their role boundaries clear. After continuous review during this phase of analysis, this code was later subsumed under the overarching theme of *ward climate* and will be presented in more detail in the following results section.

Phase 5 – Defining and naming themes

Braun and Clarke recommend that once a 'thematic map' of the data has been developed themes and sub-themes should be refined further to ensure their titles encapsulate the data they represent. They recommend that it is important for researchers to ensure that if this task is particularly difficult it is likely that there are too many sub-themes under a given theme. This was evident in the development of an initial theme of *team functioning* which initially incorporated *communication*, *supervision and leadership* and *professional regard*. Braun and Clarke warn against attempting to make a theme 'do too much' and so this theme was divided into smaller themes. This was particularly important for the next phase of the study which aimed to generate safety indicators for each theme. It was hoped that a relatively equal number of indicators would be generated for each theme and so a disproportionately large number of indicators would have been needed to represent the theme *team functioning* which may have affected subsequent statistical analysis.

Weber (1990, pp. 10) suggests that 'the best content analytic studies utilise both qualitative and quantitative operations on text by including the calculation of frequencies and percentage frequencies of comments coded in each category'. Whilst thematic analysis was the predominant qualitative method used to analyse the data in this study, the endorsement frequencies (the number of times the item was classified in a data source) of each theme, sub-theme and code were also generated for each of the three data sources. In this way an additional comment could be made on the ability of each respective method of data collection to identify latent causes of error. The frequency of higher-order theme endorsements during interviews for managers and nurses were also

compared to note any differences between the latent failures attributed to medication errors between the two groups.

4.4. Results and discussion

Results will be presented and discussed here according to the type of analysis conducted. Findings from thematic analysis will be presented first followed by findings from content analyses.

4.4.1. Thematic analysis

After completing all stages of thematic analysis, 10 overarching 'higher-order' themes emerged. Several themes comprised 'secondary' and 'tertiary' sub-themes. In a typical report of thematic analysis, each theme and its related descendents would be addressed fully in terms of how it was manifest within the data, providing additional supporting evidence where appropriate. Unfortunately, since so many themes and sub-themes were generated from this analysis, describing each theme in full is beyond the scope (and word limit!) of this thesis. Only one theme – *ward climate* will be discussed in full with additional supporting evidence to give a general idea of how all themes were identified as latent preconditions of medication errors. This theme was chosen in particular since it was one of the most complex and frequently endorsed themes. A thematic map is also provided to illustrate the hierarchical nature of the various themes and sub-theme descendents (Figure 4.4.1). The remaining nine themes and their related sub-theme descendents will be summarised with brief definitions, a sample of the most reflective excerpts and appropriate citations where supporting evidence exists in Tables 4.4.1.b - j. The most reflective excerpts representing higher order- and secondary themes will be presented. For the purposes of continuity and because this was the most detailed source of qualitative data, only excerpts from interviews will be presented.

Theme 1: Ward climate

Definition: The overall atmosphere of a hospital ward determined by predominantly unspoken multi-disciplinary shared assumptions, rules and norms of 'the way it is', which have evolved over time and forced individuals and teams to adapt to this environment.

Table 4.4.1.a: Interview excerpts reflective of the theme *ward climate*

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Ward climate	**General multi-disciplinary ward beliefs**	**Over-dependence on senior nurses**	…people come to you [as a senior nurse] all the time for every little thing, to the point where it gets pathetic, they come to you for the most ridiculously easy things that anyone would know…they come to you for everything…anybody who's seen in a sister's uniform is automatically a target for everybody, for doctors, relatives, visitors, everybody wants to speak to the sister.
		Commitment to caring for patients	…[providing safe care] is partly down to the staff themselves, whether they feel comfortable in that environment or not, whether they just come to work to come to work and go away again or whether they want to do the best that they can.
		Ward ethos	…you can have the climate on the ward isolated but you can have a climate that's a bit wider than that, say a group of wards on a unit where the prevailing culture is on speed rather than on safety and individualised care.
	Nursing attitude towards challenging others		...certainly some of the more junior staff would think "well the doctor knows best" and so their instructions must be right. Its tradition isn't it. People don't like militant nurses.

Table 4.4.1.a: continued

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Ward climate	**Nursing insecurities**	**Explicit proficiency**	…newly qualified staff think everything has to be done and dusted by the time the sister comes on so they can hand over a straight shift, everything done and organised. They think if they haven't done that, then it's a reflection on how they cannot manage their time. This often means people are rushing to get these jobs done.
		Implicit proficiency	…anybody in any working environment is trying to prove themselves…an inexperienced nurse would be trying to prove to other people they are capable and so they may take risks because of that. I'm sure there are senior nurses who are maybe doing the drug round who won't ask [for help] because they think "well, they'll think I'm stupid, I should know that".
	Nursing attitudes towards reporting mistakes		…I can say quite comfortably and I'm sure it's happened that loads of people have made drug errors and haven't reported them because they're just frightened of the consequences… I will be extremely cagey about how and if I reported another drugs error, not just of mine but of anybody else's. My reaction now is say nothing, it didn't happen.
	Professional regard	**Agenda conflict (incl. interruptions and role insight)**	…people come and interrupt you all the time when you're doing the drugs round, it's just dangerous, really dangerous. It's one of the main [causes of error], being interrupted. People are just working to very separate agendas.
		Medic-nurse relationship	…you might come on to shift on a weekend and find out which doctors are on and you might think "oh, great, he's alright, I like him" or it might be somebody you've no confidence in…when I'm interrupted it's usually by rude and ignorant medical staff.

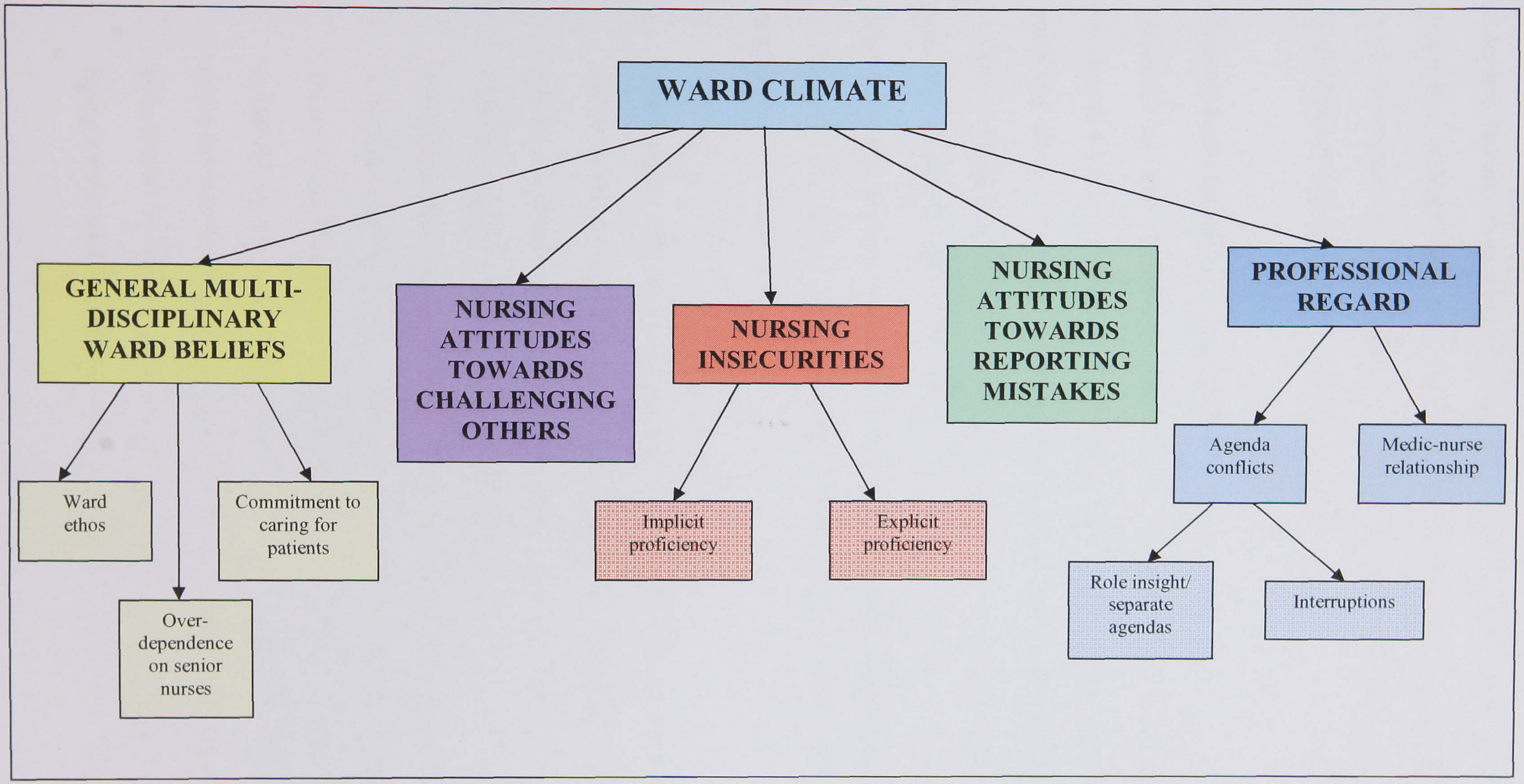


Figure 4.4.1: Thematic map of the theme ward climate with secondary and tertiary theme descendents

It became apparent from thematic analysis that the higher order theme of *ward climate* was a synthesis of multiple facets of the ward environment. Each facet will be addressed separately in terms of how each secondary theme was defined, any tertiary theme descendents and how these would manifest in the process of delivering patient care.

General multi-disciplinary ward beliefs

This particular secondary ward climate theme can be defined as an overall implicit 'feeling' about the general way patient care is delivered on a ward. Interviewees suggested that this type of embedded climate was 'driven' by particular individuals and represented unspoken values which should be adopted by all individuals working in that ward or unit. They added that these individuals were not necessarily supervisors (which is why these ideas were not subsumed within the *supervision and leadership* theme) but were often particularly assertive nurses to whom other nurses would look up to. The theme comprised three tertiary themes which will be discussed separately:

- **Ward ethos**. This was described as an overall ward atmosphere, driven by matrons and senior sisters who would be more concerned with either the speed or the safety of delivering patient care. Several interviewees suggested that nurse managers who had the most extensive experience would be more likely to encourage other nurses to focus on delivering safe care regardless of the time taken to do so. Comparatively, nurse managers with less experience would focus more upon task completion and less on the methods selected to complete them. Several nurses alluded to the likelihood that focus on speed over safety would inevitably lead to 'cutting corners' and violating safe practices (e.g. 'speeding up' during the drug round). One particular interviewee hypothesised that this may have been because initial training and early clinical experiences of nurses who had been nursing for many years would have been during a time where the provision of health care was less driven by government targets and

there were fewer, less acutely ill patients. These nurses would therefore have been and are more likely to remain less target-focussed. Although there is very little written on this type of sub-climate within health care, this finding is supported to some extent by evidence in the manufacturing industry. Zohar (2000) argues that work groups can develop sub-climates which are distinct from the overall safety climate of the organisation and driven largely by supervisory commitment to safety and in particular their expectations of productivity over safety.

- **Commitment to caring for patients**. Notably, this theme was only cited during management interviews. Several managers claimed that nurses who perceive their role as 'just a job' are therefore not committed to the role of caring for patients and as such are less likely to adhere to safe practices. Evidence has suggested that nursing cannot be exclusively understood as the delivery of a number of expert cognitive and technical skills but should be considered as an integration of these concrete skills with an 'inner attitude of caring' (Morrison, 1991; Gastmans, 1999). Although Gastmans argues that being committed to care for patients enables nurses to reach the 'goal of nursing practice', he does not intimate whether this commitment facilitates safe practice as suggested by senior managers in this study. He adds that if nurses act without commitment to care, there is a risk that they will 'lose sight of the patient as an individual and become fixated on "the problem"' (pp. 217). Neither Gastmans nor managers interviewed in this study hypothesize whether commitment to care attitudes are learned or innate or suggest ways in which other factors such as excessive workload to staff ratios and organisational expectations and targets might attenuate the motivation to care. For example, Laschinger *et al.* (2000) found that nurses who considered themselves 'empowered' by leaders in the workplace (perceptions of power, access to information, support, resources and opportunities) reported significantly higher levels of trust in the organization and commitment to their role.

- **Over-dependence on senior nurses**. Although intrinsically linked to 'task delegation' in the higher order theme of *supervision*[27], this phenomena was both observed on the wards and also discussed in nurses' (but not managers) interviews. Nurses described an unspoken 'rule' of the ward that all queries should be routed through the most senior member of nurse on shift, regardless of the nature of the problem and the task the senior nurse was currently involved with. During observations junior nurses, patient's visitors, doctors and other health professionals were all witnessed approaching the most senior nurse on shift with a range of queries such as where the fax paper was kept and whether test results had been sent to the ward despite the fact this nurse was in the middle of conducting the medication round. Senior nurses claimed that they were 'over-used' because of a long-standing belief of everyone entering the ward that senior nurses would 'know the answer to everything'. They added that while the role of senior nurse involved aspects of ward coordination they were unable to fulfill this role effectively because of staff shortages which meant they were also allocated a patient load.

It is possible that this over-reliance on senior nurses was to some extent facilitated by senior nurses themselves. For example, junior nurses stated that senior nurses often appeared unwilling to allow them to take on responsibilities which they were capable of carrying out because they said it was quicker to do it themselves. This could be due in part to traditional elitist perceptions of senior nurses and the impact of poor task delegation. In a qualitative study of the perceptions of senior nurses in adult intensive care, Bowler and Mallik (1998) suggest that senior nurses adopt an 'elitist position' in relation to junior nurses who as a result may experience feelings of 'alienation, subordination and oppression'. These feelings ultimately result in junior nurses who have low self-esteem and who are disempowered to take on

[27] It is important to distinguish between this tertiary cultural theme and that of 'task delegation' which was defined more in terms of the concrete ability to delegate tasks mediated by factors such as training, workload and skill mix.

responsibility and so will ensure all decisions are pre-approved by senior nurses. Bowler and Mallik suggest that by relinquishing overall responsibility for *every* activity, senior nurses would have to accept a less elitist position in the nursing ranks which contradicts the hierarchical grading structures in nursing. Senior nurses interviewed in this study claimed the over-reliance on their skills and knowledge was not instigated by them and it was a hindrance to the efficient performance of their nursing duties (see also interruptions under 'professional regard' tertiary theme). Conversely, while junior nurses agreed that this over-dependence on senior nurses was a long-standing and unspoken rule of the ward, they maintained it was nevertheless a deliberate attempt by senior nurses to justify their higher position.

Nursing attitudes towards challenging others

This secondary theme was defined as the perceived ability or confidence of health care staff to challenge the decisions of colleagues they believe to be incorrect and the openness of those individuals to act upon this contradictory advice. During interviews, both managers and nurses suggested that nurses were less likely to challenge decisions made by doctors and proposed two main reasons for this. Firstly, nurses suggested there was a long-standing tradition of deference where 'the doctor knows best' within the field of nursing which can have a disempowering effect on nurses who feel it is not their place to question someone in a position of perceived power. Senior nurses suggested there was a perceived 'expertise gap' between doctors and nurses which ultimately affects the confidence of junior nurses *to challenge* doctors and affect the likelihood that doctors would be open *to be challenged* on their decisions by junior nurses. Several senior nurses claimed that if a junior and senior nurse approached a doctor with exactly the same problem and advice, he/she would be much more likely to act on the advice of the senior nurses due to the perceived smaller expertise gap. Such a proposition is consistent with evidence by Sasou and Reason (1999) who found 'excessive professional courtesy' (e.g. 'doctors know best'), and 'excessive authority gradient' (the real or perceived difference in power between two or more individuals) were

significant predictors of failure to highlight and correct mistakes. Similarly, Mearns, Flin and O'Connor (2001) postulate that where many sub-groups interact with one another, it may be unclear within that organisation exactly who has ultimate authority. This is particularly true of health care which has a multitude of different professions, each with their own line- middle- and senior-managers responsible for their decisions and actions. Although this was not frequently endorsed within nursing or management interviews, two senior nurses did suggests that although they felt confident enough to challenge a decision they thought was wrong, they were unsure whether they were 'allowed' to challenge doctors because they were not in charge of them. Helmreich and Merritt (1998) suggest that 'status differentials' between doctors and nurses make it virtually impossible for those considered culturally (by themselves or by others) as having a lower status of authority to challenge their real or perceived superiors when they make errors.

By comparison, several managers interviewed suggested that nurses do not challenge mistakes that doctors make since they believe it is the doctors own responsibility and that of their own management (e.g. registrars and consultants) to monitor and question their behaviour and decisions. In this way, managers suggested this was a deliberate distancing of responsibility (e.g. 'it's not my job as a nurse to check doctors are doing everything right') by nurses for ensuing errors. Managers argued this attitude to challenging potential mistakes would undoubtedly be affected by the relationship held between nurses and medical staff on any given ward (see also *professional regard* sub-theme).

Nurses also stated that over the last few years within their hospital (as with many other NHS Trusts) there had been a large influx of nurses recruited from the Philippines (Buchan, 2002). Alongside language barriers (subsumed under the higher order theme *team communication*) nurses hypothesised that these nurses had been trained in a hospital climate which does not encourage the challenge of doctors' decisions and where nursing staff are viewed as much lower down the chain

of expertise. Unfortunately it was not possible to interview any of these nurses since they were not nominated by the matrons as potential candidates. During observations, a clear cultural divide between English and Philippine nurses was witnessed and significantly less communication was evident between the two groups. Evidence suggests that nursing in Asian countries is 'high-context' which means individuals depend to a large degree upon nonverbal signals rather than the spoken word (Bola, Driggers, Dunlap & Ebersole, 2003). Research indicates that nurses trained in such environments can be perceived by western nurses as inattentive and subservient to medical staff (Yi & Jezewski, 2000).

Senior nurses and managers suggested that promoting an acceptable challenging climate on a ward relies in part on increased emphasis during initial nursing and medical undergraduate training. Managers further speculated that improving communication between the different disciplines would improve nurse-physician relationships overall. This improved relationship would both empower nurses with confidence to challenge decisions they thought were potential risks to the safety of patients and also attenuate any perceived threat to expertise felt by challenged doctors. Furthermore, senior nurses felt that challenging the decisions of others was a skill which could be learned like any other clinical skill. They believed it was their responsibility to 'lead by example', and demonstrate to junior members of staff the 'right' way to challenge colleagues and other health professionals.

Nursing insecurities

This theme is distinct from other secondary themes in the sense that it relates only to idiosyncrasies pertaining to nurses and as such is not driven or mediated by any other health professional. Nurses described feeling a constant pressure and awareness that they should prove their worth and was described in two main ways;

- **Explicit proficiency.** This was described by nurses as the perceived need to prove their level of concrete skill and knowledge. Nurses described feeling that they must prove daily to other nurses that they have a good level of skill and expertise. This was thought to occur in a cyclical way. Senior nurses suggested that they were unlikely to seek advice for something they didn't know from a nurse who was one or two grades below them for fear they would be judged as 'unworthy' of their superior position and salary. They reported that they would feel 'ashamed' asking for help from a more junior nurse, and that nurses were judging them on something they should already know according to their grade. Junior nurses claimed that they would be similarly unlikely to seek advice from senior nurses due to a desire for respect and a need to prove themselves. As a result of this expertise 'stand-off', both senior and junior nurses suggested they would sometimes rather take the risk of being wrong than ask for help.

- **Implicit proficiency.** This was described by nurses as the perceived need to prove their intrinsic ability to manage under pressure and that they were a 'good' nurse. Both senior and junior nurses suggested there was an unspoken expectation within nursing that each shift should start on a 'clean sheet' without outstanding jobs from the previous shift. Nurses suggested that because of this expectation they would be likely to cut-corners and speed up during tasks (including the medication round) in order to appear to the next shift that they had managed their time well. Nurses reported feeling defensive and experiencing high levels of anxiety when unpredicted events occurred on the ward which might mean they would be behind in their shift plan and would be unlikely to complete all tasks before the next shift began.

Although there is little published evidence on nursing insecurities which might explain risk taking behaviour, such a finding is consistent with research conducted in adolescent avoidance of help-seeking in the classroom. Ryan *et al.*, (2001) note that low-achieving

students or those who believe they are not particularly competent are more likely to think that other students will perceive their need for help to solve a problem as a sign of incompetence and therefore will be less likely to seek help (Butler and Neuman, 1995; Butler, 1998; Karabenick and Knapp, 1991). In previous research, Ryan and Pintrich (1997) found that in asking for help, these students would worry about being negatively judged as 'dumb' by teachers and other students. This could explain the finding in this study of the unwillingness of junior inexperienced nurses to seek help from more experienced nurses. Research has also indicated that the avoidance of help seeking may also be governed by achievement-goal orientations. Hicks (1997) proposes help avoidance may be largely driven by a status-goal orientation. Individuals with an achievement-goal orientation are overly concerned with maintaining a particular image or reputation of expertise and have a heightened awareness of the self in relation to others (Hicks, 1997). Ryan *et al.*,(1997) report that students who have a status-goal orientation are more likely to avoid help-seeking because they perceive it as a threat to their perception of self-worth. It is possible that senior nurses' reluctance to seek help and advice from others can be explained by a desire to protect their perceived social status.

Attitudes towards reporting mistakes

This sub-theme relates to the attitudes held by various health professionals on a ward or unit which govern the likelihood they will report errors. Surprisingly, managers did not cite reporting mistakes as an important precursor of future error. However all nurses interviewed suggested reporting climate was vital in understanding why the same errors repeatedly occur and targeting appropriate interventions to prevent them. Nurses identified eight main reasons reporting at a ward level could be affected;

1. **Definition of error.** Nurses indicated that they were unsure exactly what constituted a medication error which should be reported. It was observed in the review of incident reports in this study that although it was recommended 'significant incidents' should be reported, no definition of significant, incident or error was presented on this form. Forms provide a classification system for reporters to categorise the type of error or incident they are reporting. However, it became evident during the incident report review that no clear guidelines or definitions for these were offered to reporters. Furthermore, analysis revealed that this list of potential categories of incident comprised outcomes (e.g. patient slip/fall, cut with sharp object), causes (e.g. medical records problem, staffing shortages, communication failure) and incidents not related to errors (e.g. physical assault by patient, theft of property).

2. **Severity of error.** Nurses claimed that in general they knew what comprised a 'serious' error but they were unsure what constituted a near miss. Furthermore, nurses generally stated they would be unlikely to report a medication error if it had been averted before reaching the patient or administered to the patient without consequence. Nurses suggested that while they would not officially report such errors, they would disclose them to colleagues as an informal local warning. This would imply that the number of medication errors being reported is a significant underestimation of the actual number occurring.

3. **Reporting responsibility.** Nurses suggested they would be more likely to report medication errors which they or another nurse had been responsible for but would be unlikely to report an error made by a doctor or other health professional. This was in part due to a lack of time to complete forms (see design of incident forms) but also as a professional courtesy. Nurses suggested that reporting the actions of another health professional, particularly doctors would not benefit the ward climate, especially the relationship between the two professions.

However senior nurses stated they felt it was their duty to try to persuade doctors to complete their own forms but added that this rarely resulted in a form being completed.

4. **Design of incident forms.** Nurses suggested that the likelihood of reporting a medication error was adversely affected by the design of incident forms which often meant that completing forms could take in excess of 20 minutes depending on the level of detail provided. This was supported by the incident report review conducted in this study which observed that very little detail was provided in any of the reports. For this reason, nurses stated that they would make a quick mental calculation of the severity of the error (in terms of patient outcome) and the likelihood of it occurring again before considering whether to report it.

5. **Confidentiality of reporting.** Nurses suggested that they were reluctant to report medication errors or near misses which could have had severe or even fatal consequences because reporting was not anonymous. They added that they would be more likely to report any error executed by themselves or others if they did not have to submit their own details.

6. **Poor experience of reporting.** There was a general acknowledgement that although reporting errors was the 'right thing to do', the likelihood of reporting was significantly affected by either a punitive personal experience or a particularly negative experience of a colleague. As an illustration, one senior nurse described her experience of reporting an error the previous year. This nurse had administered a double dose of IV antibiotic to a patient mistakenly believing she had not yet administered it. After reporting the error this nurse was reprimanded, demoted and was banned from administering medications for six months after which time she was shadowed by another senior nurse until it was felt she had achieved a

suitable level of competency. She revealed she would never report another medication error regardless of who was involved or the outcome for the patient.

7. **Lack of feedback.** Nurses in general perceived the lack of feedback from previous error reporting to be a substantial barrier to reporting future errors. All nurses said they had completed one or more incident forms in the last 12 months and all nurses except one stated they had not received any feedback, either positive or negative regarding the incident. In general it was felt that this would be a deterrent to reporting all but very serious errors in the future.

8. **'Blame culture'.** Nurses reported that there was a palpable blame culture surrounding error reporting which acted as a barrier to reporting some types of error. They added that this was not necessarily driven by a poor personal experience of reporting but it was just something they 'felt' within the hospital. This awareness was directly observed during this study on more than one occasion whereby junior nurses were 'coached' by senior nurses on the 'correct' way to write incident forms (see section 4.3.2.). During observations, senior nurses suggested this coaching was necessary to protect nurses from culpability.

There is a considerable degree of evidence to support both the existence of a 'reporting culture' within health care organisations and the likelihood that this will indubitably affect organisational learning and the probability of future errors. In support of the reporting issues identified in this study, Jeffe *et al.* (2004) found that uncertainty surrounding less serious errors and the need to report them, fear of reprisals, lack of confidentiality, time taken to complete forms and absence of feedback were cultural barriers to reporting errors for both doctors and nurses. In a similar focus group study, Uribe *et al.* (2002) identified multiple barriers to error reporting including incident

form complexity, error severity, organisational blame culture, ambiguous definitions of error and lack of guidelines on who is responsible for reporting.

Professional regard

This secondary theme was distinguishable from the way different health care disciplines *explicitly communicate* with each other and concerns the way in which these professionals *implicitly relate* to one another on an inter-personal level and the subsequent impact of this implicit relationship on providing safe patient care. This sub-theme was especially complex and comprised two tertiary themes which overlapped considerably with other ward climate sub- themes;

- **Nurse-medic relationship**. Nurses referred to their relationship with medical staff as being a particularly important predictor of medication errors. Senior nurses suggested that junior nurses were not taken seriously by doctors who were more likely to act on the advice of a senior rather than junior nurse, even if that advice was the same. This relationship was described as having a knock-on effect on the confidence of junior nurses to challenge doctors' mistakes (*nursing attitudes towards challenging others*) which in turn lead to an over-reliance of senior nurses by junior nurses and the medical team (*general multi-disciplinary ward beliefs*). All nurses reported that there were some doctors they were happy to work with and others in whom they had no confidence or rapport. Several nurses said they would rather wait for another doctor to come on shift than bleep the doctor they did not have a good working relationship with if they needed medical support or advice regarding a patient. Several nurses suggested that a perceived steep hierarchy gradient of doctors meant that nurses were generally less likely to see consultants as 'approachable'. As a result nurses would be reluctant to seek their help and advice, even though the consultant would be the most

knowledgeable source. Several senior nurses referred to the risk of having their 'head snapped off' during ward rounds if they were to question a consultant.

The working relationship between nurses and doctors has been well documented. In some of the earliest literature, Stein (1967) describes this relationship as the 'doctor-nurse game' whereby nurses learn to show initiative and offer advice while appearing to 'defer passively to doctor's authority'. Stein argues that this game ensures that disagreements are avoided at all costs which although appears on the surface to be effective, is indicative of poor communication which may ultimately affect patient care. Revisiting his theory in 1990, Stein noted there had been a significant 'deterioration of public respect for doctors and recognition of their fallibility' and an increase in the number of female doctors and male nurses which had shifted the dynamic relationship between doctors and nurses (Stein, Watts & Howell, 1990). Stein *et al.* suggest these changes have resulted in an empowerment of nurses who feel less passive in their relationship with doctors and more actively involved in the provision of patient care and ultimately 'more valued in their own right'. These changes have therefore lead to a more overt awareness of the poor relationship between doctors and nurses. More recently Mackay (1993) suggests that the historical change in the doctor-nurse relationship over the last 20 years may have been affected by educational developments such as the expansion of university based nursing degrees and the introduction of diploma courses based in institutes of higher education. Mackay argues that these changes have increased knowledge and autonomy for nurses and may have altered their power relations with doctors.

In terms of how these changes might be expected to affect patient safety, Adams, Bond and Arber (1995) emphasised the impact of inter-disciplinary working relationships in a qualitative study to identify organisational features which determine effective nursing practice. In particular, Adams *et al.* refer to the importance of the effective collaboration

between doctors and nurses in terms of planning patient care activities, asking for advice or opinion and sharing of ideas. In a review of the literature surrounding the determinants of patient mortality, Tourangeau, Cranley and Jeffs (2006) cite two studies which found hospitals with the highest patient mortality rates had the worst nurse-physician relationships (Knaus *et al.,* 1986; Mitchell *et al.,* 1989).

- **Agenda conflicts.** This tertiary theme represents the inter-disciplinary disparity in planning essential patient care activities in order to achieve the same goal. This theme was discussed as a manifestation of two main problems; lack of role insight and interruptions and will be discussed separately.

1. **Role insight.** Nurses suggested that although the aim of nurses, doctors and other health professionals working in the hospital was essentially the same – to provide safe, timely and effective treatment for patients, the methods each discipline applied to achieving this goal was not considerate of other disciplines. For example, during observations and also discussed frequently in interviews, it was noted that doctors arrived on the ward to conduct ward rounds while drug rounds were still being carried out. In one particular observation, ten house officers, two registrars, two physiotherapists and a phlebotomist arrived on the ward at exactly the same time as the senior nurse conducted the drug round. All professions were observed on a number of occasions approaching the senior nurse in order to achieve their own objectives, significantly impairing her ability to concentrate or even complete the task. During interviews nurses said this was 'just the way it was' and previous attempts to coordinate the timing of these activities had failed.

Nurses suggested that other departments simply didn't understand the priorities of nurses and the nature of their role and often asked nurses to undertake tasks which were not explicitly their responsibility but ones which over time they had 'adopted' because no one else had. For example, during one observational period from the central nursing station it was noted that a ringing telephone was left unanswered by a group of seven doctors who were standing nearby, three of whom were having a non-work related conversation. The telephone rang continuously for several minutes before a nurse broke off from changing a patients bed sheets at the other end of the ward to answer the call, which ironically was for one of the doctors. It was apparent that linked with this 'role ignorance' was the absence of formal communication strategies which would promote the sharing of role responsibility information.

West (2000) argues that lack of role insight may in part be due to the fact that members of health care delivery teams are generally educated separately without reference to other disciplines. She adds that the role of nursing has become much more specialised over the last 20 years. While nurses are aware of the aspects of patient care which are *currently* their responsibility, it is probable that other disciplines are not aware of the features of nursing for which nurses are no longer responsible (e.g. ward cleaning, ancillary services, administrative duties etc.). Vaughan (1996) refers to this as 'structural secrecy' and suggests that compartmentalising knowledge in this way can lead to increased potential for errors when tasks or information falls between the 'gaps' in role responsibilities. West (2000) advises that to improve this cultural phenomenon, health care organisations need to implement or improve formal methods of inter-disciplinary communication. Interestingly, one particularly senior nurse said that she had never had a meeting with any other department regarding their general working

relationship, role boundaries or coordination of multi-disciplinary activities (which was not related to the care of a particular patient) in the 15 years she had been a nurse.

2. **Interruptions.** It should be noted that during inter-rater reliability tasks (presented in section 4.4.3) there was some disagreement between raters in the most appropriate categorisation of interruptions. It was postulated that interruptions may have represented poor team communication, written policies and procedures (i.e. the non-adherence of standard policy *not* to interrupt) or local working conditions (i.e. interruptions are a condition of the physical working environment). However, it was eventually agreed between raters that the predominant factor governing interruptions was an implicit unseen and unspoken ward phenomenon where there was no specific protocol stating nurses should not be interrupted during complex tasks and nurses did not specifically ask not to be interrupted during tasks. Moreover, this behaviour had become accepted practice over time and although nurses and senior managers who were interviewed acknowledged this behaviour was 'poor practice', they suggested it was the cultural 'norm'.

Both managers and nurses to a much greater extent referred to the prevalence of interruptions during complex nursing activities such as the preparation and administration of medication. During one particular observational session, a senior nurse was interrupted over the course of 60 minutes on 17 separate occasions by several health care professionals, patients and visitors while conducting the drug round. The nature of these interruptions ranged from patients' visitors asking how their relatives were faring to a ward clerk asking where the fax paper was stored. All nurses emphasised that interruptions were frequent and a constant source of annoyance and stress. In support of this finding, Wolf *et al.* (2006) found during clinical observations

that nurses were interrupted mid-task on average 3.4 times per hour which resulted in them changing focus from one patient to another an average of 9.1 times per hour (once every 7 minutes). In a similar observational study, Potter *et al.* (2004) found one nurse was interrupted mid-task 43 times during a 10 hour shift in order to source materials, equipment or personnel.

Perhaps the most immediately obvious and direct impact of interruptions and one cited often during interviews is the likelihood that distraction during a cognitively complex task such as the administration of medication will lead to errors and mistakes. There is considerable evidence within cognitive psychology supporting a relationship between task distractions and errors. Pani and Chariker (2004) argue that when an individual is interrupted the contents of their working memory are immediately discarded to allow them to attend to the distraction. When attempting to resume the original activity individuals often believe they are farther along in the process and will omit parts of the procedure or will repeat steps already taken (Mandler, 1982; Rudolph & Repenning, 2002). Pani and Chariker (2004) suggest that this is particularly likely in the administration of drugs, a task which over time may become automated and require little conscious attention due to constant repetition. However, despite abundant evidence which has reported the *perception* that interruptions precede medication errors (Conklin *et al.,*1990; Scholz, 1990; Walters, 1992; Davis, 1994; Segatore *et al.,*1994; Williams, 1996), there has been very little health care research conducted to measure whether interruptions actually *cause* medication errors (O'Shea, 2001).

The following tables (4.4.1.b - j) will summarise the remaining nine higher order themes with definitions and secondary and tertiary theme descendents illustrated with relevant excerpts from interviews.

Theme 2: Team communication

Definition: Aspects of an intra- or inter-departmental team or communication channels which prohibit effective communication between individuals or departments.

Table 4.4.1.b: Interview excerpts reflective of the theme team communication

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Team communication	**Written** (Gladstone, 1995; Howell, 1996; Osbourne *et al,* 1999; Phillips *et al.,* 2001)	**Paperwork tracking**	…there is so much information in these paper records that we have difficulty storing them, transferring them, keeping tabs on them
		Handwriting legibility	…I think it's an absolute scandal that we have to rely on doctor's handwriting that's often completely illegible.
		Document design	…on the front of drug charts, the majority of people have got [their allergies] documented, but you open the drug chart up and it's not there. A lot of them miss off documenting the name, the unit number and the allergies inside the drug chart.
	Verbal (Dean *et al.,* 2002; Wakefield *et al.,* 2008)	**Absence of formal information sharing**	…doctors do things and write things up in their own notes and expect it to be done without informing anyone what they've done. They expect you to look in the patients notes every 5 minutes to check.
		Over-reliance on informal communication	…the patient had been incorrectly written up for Magnapen on a verbal order by the Consultant, the house officer had written it up without looking at the notes. I think we're over-reliant on verbal orders and information handed over verbally.
	Team size (West, 2000; Temple, 2004)		…the more people are involved, the more communication there is, the more time is needed and I think it takes more time doesn't it and you've got lots of people that know lots of little bits but you haven't got one person that knows everything so I think things slip through the net sometimes.
	Multi-cultural issues (Bola *et al.,* 2003; Yi & Jezewski, 2000)		…a lot of errors are caused by communication problems since a lot of the nurses are from the Philippines and they often misinterpret instructions they have been given.

Theme 3: Routine procedures

Definition: Procedures routinely carried out by nursing staff in the course of a patient's stay in hospital regardless of the patient's condition.

Table 4.4.1.c: Interview excerpts reflective of the theme routine procedures

Higher order theme	Secondary theme*	Excerpt
Routine procedures[28]	**Checking procedures** (Pani & chariker, 2004)	…they get complacent they're used to doing the procedure or they know the patient and they just decide that they wont check numbers, check drug charts, they'll just give people medicines without going through the right checking procedures.
	Patient admission procedure (Cornish *et al.*, 2005)	…quite a lot of the time on here, people are admitted and you'll come on the next morning, be in ward round and there will be the name of the patient that was in the bed before them. I don`t think we have a proper admissions procedure on here, I think it`s very laissez-faire. There is a standard of what should happen (in a patient admission) but it is implied, there is nothing written down.
	Patient handover (Cook *et al.*, 2000; Coxon *et al.*, 2003; Pothier *et al.*, 2005)	…the handover is essential, that's where so many of these things go wrong and particularly if the handover is multidisciplinary, between different groups of staff, this is much more error prone.
	Patient discharge procedure	…I would have thought that the time when you're most likely to make mistakes, any mistakes with regards to patients, would be on patient discharge because that's when you've got to get everything together, all the information together, all the systems working.

* No tertiary sub-themes extracted

[28] The routine procedure of medication administration is notably absent. This is because after inter-rater reliability task (section 4.4.3.) issues such as ward drug stock levels, drug round timing and interruptions were reassigned to other themes (LWC, workload & ward climate respectively).

Theme 4: Workload

Definition: Facets of nursing care which place significant physical and/or mental demands upon nursing staff which could affect their ability to care for patients effectively.

Table 4.4.1.d: Interview excerpts reflective of the theme workload

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Workload	**Volume of work** (Conklin *et al.*, 1990; Pani & chariker, 2004; Wolf *et al.*, 2006)	**Task demand**	… you've got x amount of drugs to do and so many IV's to do and you know you've got the doctors coming on and you know that you've got to speak to a load of relatives…when you're that busy you forget things because you just don't have time to do everything that you should be doing.
		Task timing	…first thing in a morning is desperate time... it's the biggest drug round, you've got doctors coming doing rounds at the same time…you've got Physio's who come on and get going, social workers phoning up, the OT's come on saying "excuse me, so and so's not in her bed…"
	Cognitive workload (Beaudoin & Edgar, 2003; Pani & Chariker, 2004; Wolf *et al.*, 2006; Wakefield *et al.*, 2008)	**Over-burdened memory**	…you have a piece of paper in your pocket and you write down some of the jobs you need to do but you don't write every single thing down that you're ever likely to need to do…with all the best will in the world when you're tired and you're at the end of your shift and you're handing over you can't always remember everything.
		Information overload	…people do think that you know everything and you'll be able to take all this information on board. It doesn't matter how much they give to you, you're just gona take it and take it and that's it…you don't have that opportunity where you might have 10 minutes where you've not got a hundred things in your head to actually be able to think a little bit because you just don't have that mental capacity sometimes.

Table 4.4.1.d: continued.

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Workload	**Workload Planning** (Wolf *et al.*, 2006; Lundgren & Segesten, 2001; Tucker & Spear, 2006)	**Shift prioritising/ unpredictable environment**	…there are lots of patients being admitted, or very sick patients in need of a lot of attention, dealing with relatives and visitors. It can fluctuate throughout the day. It's very unpredictable. You might have a patient that just suddenly becomes ill and it can change your whole plans for that shift.
		Ability/time to plan ahead	…you are sort of fire-fighting and running round all the time and I suppose a lot of nursing is very reactive and some of that`s because you're doing that all of the time. I suppose it affects your ability to plan ahead because you're not given enough opportunity to do it.

Theme 5: Human resources

Definition: Aspects of the provision of health care personnel including the number of available permanent qualified staff, their respective skill-base and the employment of contingent workers.

Table 4.4.1.e: Interview excerpts reflective of the theme human resources

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Human resources	**Temporary/contingent workers** (Roseman & Booker, 1995; Anderson *et al.*, 1996; Rousseau & Libuser, 1997)		…you've got to be with them (agency staff) all the time, showing them what to do, teaching them, keeping an eye on them, checking everything they've done rather than being able to get on with your own job, your own role for that day.
	Staffing levels (Roseman & Booker, 1995; Blegen *et al.*, 1998; Aiken *et al.*, 2002; Needleman *et al.*, 2002)	**Medical Support**	…on weekends, it's just so awful trying to get medical staff onto the ward… after five o clock on a week day or all day at the weekend it's just an on-call assistant doctor so there's very few doctors about and they tend to be over worked.
		Nurses	…when there's not enough of you, you cut corners, do things quickly, rush things, don't communicate as well… we just haven't got enough staff to cover the nights and the days as well.
	Skill mix (Leape *et al.*, 1995; Lankshear *et al.*, 2005; Seago *et al.*, 2006)		…I suppose whether an error is likely to occur would depend on the skill mix on the ward at the time…written into the shift rota is other things like 'well, she's ok, she's a D grade, she's fine, level headed, she'll be able to manage anything and then that person will want to change [shifts] with somebody else who's not quite the same…on paper they are the same grade, but they're still not the same. And I think all that probably contributes towards a good shift and a bad shift.

Theme 6: Local working conditions

***Definition**:* Aspects of the individual or the immediate working environment such as work patterns and physical working conditions which hinder the provision of safe patient care and encourage the performance of unsafe acts.

Table 4.4.1.f: Interview excerpts reflective of the theme local working conditions

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Local working conditions	**Patient** (Barber, 2002; Dean *et al.*, 2002)	**Non-compliance**	…every drugs round that I do, there will be at least one patient that's got their own medicine in their locker and they've already taken them despite denying on admission they had any.
		Illness complexity/medication volume	…a lot of the patients are on a lot of medications… some people are on 2 or 3 drug charts full of tablets When you do your medication round it takes a long time, some people might be on 20 tablets not including nebulisers and things.
		Familiarity	…when you know the patients and you know who they are, you don't always check wristbands because you already know it's them.
	Ward design (Hendrich & Lee, 2003; Marck *et al.*, 2006)		…having to go to the other end of the ward to the clinic room for medication not on the drugs trolley does increase your potential for making errors…you get so many interruptions, the patient`s don't care which team you're working for so they will stop you and ask for things…you end up forgetting what you set out for.
	Personal issues (Combs & Taylor, 1952; Helmreich & Merritt, 1998; Sandal, 1999; Sexton *et al.*, 2000)		…I know in my heart of hearts that I can't expect her to function as she normally would because of what`s going on at home…what goes on in their personal life is bound to affect performance at work.

Table 4.4.1.f: continued.

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Local working conditions	**Fatigue and body rhythms**	**Shift patterns** (Narumi *et al.*, 1999; Gold *et al.*, 1992; Lowden *et al.*, 1998)	…night shifts are long and tiring, trying to keep busy and awake. By your 4th night you might handover and miss something which could be quite important.
		Inadequate staff breaks (Hawkins *et al,* 1985; Rogers *et al.*, 2004; Dean *et al.*, 2006)	…sometimes you might work 11 hours without a break at all because you've been so busy, that does happen quite frequently.
	Ward Noise levels (Morrison *et al.*, 2003; Topf, 2000; Walsh-Sukys *et al.*, 2001)		…it`s noisy on the ward, that`s a really big factor. You`ve got 'patient-line` TV now at every bed…sometimes you'll think "god, I cant' concentrate, it's so noisy.
	Equipment design and availability (Tucker, 2004)		…we often struggle to get hold of IV pumps and have to ring round loads of other wards to borrow them. Some wards have this equipment and won't share which is a problem.
	Pharmacy & dispensing issues	**Expired stock** (Lisby *et al.*, 2005)	…the expiration dates on the boxes are embossed and so you can't always read what the date is. It's really easy to give it when it's actually gone out of date because you just didn't notice.
		Labelling and packaging (Cohen, 2000)	…there's quite a lot of packaging for drugs that's very similar…very often you'll have a cupboard full of pastel or white boxes which are confusing.
		Ward stock levels and ordering	…sometimes drugs you`ve ordered doesn`t arrive until the following day, 36 hours sometimes. Pharmacy is an absolute nightmare for that.
		Drug names (Carothers, 1999; Phillips *et al.*, 2001)	...there are a number of drugs which have trade names which are virtually identical.

Theme 7: Bed management

Definition: Organisational procedures to manage either the number of available in-patient beds or the ways in which patients are allocated appropriate beds.

Table 4.4.1.g.: Interview excerpts reflective of the theme bed management

Higher order theme	Secondary Theme*	Excerpt
Bed management†	**Transfers and 'sleepouts'**	…when MAU or bed managers want us to make room they'll say sleep people out so we have to send our patient`s to other wards. We transfer patients to hospital X during the middle of the night, the lights are off and we're packing people up in the dark. The potential for things going wrong when there isn't as many staff on you can`t fathom. You end up sleeping patients out to inappropriate wards which is a terrifying potential for mistakes.
	Patient throughput	…we haven`t got fewer people in the country or we've got less ill people, it's just that you are under constraints to put them through faster…there's pressure on the beds and wanting to get people out as soon as possible and [consultants] are the ones who are trying to discharge people really really quickly even though, they might be ready medically bit they`re not ready what we call 'socially'.
	A&E breach rule	…patients are rushed through to any ward that they can get a bed on because they've got this silly 4 hour waiting thing on A&E. Everything's done in a big rush in order to free up a bed up here, so MAU can free up a bed for the person on the trolley in A&E…this isn't done in a nice planned way, it's done in a horrible rush taking no heed of what's good and nice for the patient, what's safe.
	Bed availability	…there's only a limited amount of beds and the beds have been reduced over the last couple of years dramatically, wards have closed and there's less beds now but just as many poorly people.

* No tertiary themes extracted

† No literature was found exploring the indirect relationship between inappropriate bed management on patient safety

Theme 8: Supervision and leadership

Definition: Aspects of immediate line management which impacts upon the ability of subordinates to provide or be motivated to provide timely, coordinated and safe patient care.

Table 4.4.1.h: Interview excerpts reflective of the theme supervision and leadership

Higher order theme	Secondary theme*	Excerpt
Supervision and leadership	**Task delegation** (Walston & Kimberly, 1997; Miller, 2005)	…I do think it is part of control as well because there are leaders which are good leaders who can organise and sort things out quite well and delegate quite well. There are others that don't delegate who try and take on too much then which cause further problems… if particular people are on, you do find a senior sister I know of for example can end up increasing the workload further by trying to take on too much, not letting others help then there ends up being problems, things get missed and don't get done and then it kind of passes on to the following day adding more and more to their workload.
	Leadership style (Firth-cozens & Mowbray, 2001; Flin & Yule, 2004)	…if the ward sister's perception on that unit is that to be efficient you have to have the medicines finished by half past 9, that's her perception of what efficiency is all about. If you get a different sister on there whose perception of efficiency is getting things done properly, not necessarily in the most timely way then that's a different perception. I think most things that happen at a ward level are driven by the sister and their individual`s styles and their style of leadership. To a large degree they become the role model for the whole ward. You get a lot of students on the ward and you'll see them doing a drug round and the senior nurse or whoever's doing it with them will say "right, we'll do it properly this time, this is not the way I normally do it but we'll do it properly".

* No tertiary themes extracted

Theme 9: Training

Definition: The availability, appropriateness and process of delivery of training to newly qualified and existing nursing staff.

Table 4.4.1.i: Interview excerpts reflective of the theme training

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Training	**Induction and preceptorship** (Bain, 1996; Hautala *et al.*, 2007)		…you're meant to get a preceptorship package when you start where you get loads of information and things, I never received one; I was just thrown in at the deep end. Some people do start work before they've started their induction program and when a new nurse comes to the ward, there's nobody saying to them 'right, this is how we do this, this is how we do this, this is how we do this and this is how it must be done'.
	Ongoing training	**Training in procedures**	…a lot of nurses have been using equipment for years by have never actually been taught how to use it properly, they've just been shown by someone else. So you're relying on them to know what they are doing but they might not.
		Time allocated to training	…because of the constraints of time and the fact that we don't have any extra time, if we've got a new nurse starting they could end up running one side of the ward by themselves on their first day. To have the ability to teach junior members of staff is not there like it was, you just have too many patients.
		Training resources	…places are available for training but it's whether at a senior level they will allow staff to go on study leave. A lot of areas are quite stringent. The hospital is busy and short staffed so they won't let X number of nurses go off the ward to be trained.

Theme 10: Written policies and procedures

Definition: Aspects of the development and dissemination process of explicit written policies, guidelines and procedures which impact upon the knowledge of and subsequent utilisation by nursing staff.

Table 4.4.1.j: Interview excerpts reflective of the theme written policies and procedures

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Written policies and procedures	**Policy knowledge** (Feder *et al.*, 1999)	**No. of policies**	…you couldn't know what's in every policy, there's that many. I don't think anybody, even the people who write them probably know them all, there's that many... because there's so many, it's knowing and trying to keep all that information in your head.
		Policy communication (incl. ineffective dissemination strategy and time and ability to disseminate) (Woolf *et al.*, 1999)	…some wards give individual policies to the individual staff in like pigeon holes and ask them to read them. Now then it becomes the onus on the individual to update themselves. Some areas have signing sheets to say "I have read the above policy", so that they know but you can't guarantee that that nurse has actually sat and gone through it. Most of the time they just come in the post and there's no explanation. It could just appear somewhere and it would be put on the shelf.
		Updates and revisions	…It's difficult because there are so many policies and to keep remembering what the new version is. Many a time I'll look at the new version but I don't know what's different on this one, what have they changed. Nothing's highlighted so you can quickly look to see what's changed. All these different complex policies are changing all the time.
		Time allocated to read	…we have protocols and procedures for everything you can think of. It's a good thing but having time to read them all is a different thing altogether… to actually go and start reading through policies you would probably have to stay behind your shift or on your break or if there was something that you just didn't know how to do or were unsure about, to get on the computer might be difficult.

Table 4.4.1.j: continued.

Higher order theme	Secondary theme	Tertiary theme	Excerpt
Written policies and procedures	**Policy development** (Cooper, 1995; Feder *et al.*, 1999)	**User-friendly** (Marck *et al.*, 2006)	… [policies] are not user friendly, they're too long winded which is probably why people don't approach them and look at them as often as they should.
		Workability	…if you absolutely followed every policy to the letter you wouldn't have time to do your job. They're written in a way of how things should happen but sometimes realistically they don't happen that way. If you were to give out the medications exactly according to all the guidelines and protocols half the patients on the ward wouldn't get their medications.
		Blunt-sharp end translation	…it's quite an hierarchical structure in the NHS and people that are quite senior are detached from the shop floor and they see a policy and they can't understand why people haven't adhered to that policy because they are detached from what's happening.

4.4.2. Content analysis

Content analysis was conducted to calculate the number of excerpts relevant to each higher order theme to compare frequencies for all three sources of qualitative data sources and also for manager and nurse interviews.

4.4.2.1. Comparison of data sources

Excerpts were considered endorsements of themes and frequencies and percentages for each theme are presented in Table 4.4.2.1. Themes appear in descending order of their number of total endorsements.

Table 4.4.2.1: Endorsement frequency by type of data source

Theme	Endorsement frequency (% of total endorsements)			
	Interview	**Observation**	**Incident**	**Total**
Ward climate	139 (79%)	37 (21%)	1 (0.6%)	**177**
Local working conditions	64 (61%)	15 (14%)	26 (25%)	**105**
Workload	61 (79%)	12 (16%)	4 (5%)	**77**
Human resources	70 (91%)	4 (5%)	3 (4%)	**77**
Team communication	40 (57%)	16 (23%)	14 (20%)	**70**
Routine procedures	49 (72%)	3 (4%)	16 (24%)	**68**
Bed management	44 (100%)	0	0	**44**
Written policies and procedures	38 (100%)	0	0	**38**
Supervision and leadership	17 (68%)	8 (32%)	0	**25**
Training	18 (82%)	0	4 (18%)	**22**
TOTAL	**540 (77%*)**	**95 (14%*)**	**68 (9%*)**	**703**

* Percentage of all excerpts extracted

Although each theme and their respective endorsements will not be discussed individually, particularly salient findings will be highlighted here with a view to illustrating the relative contribution of each data collection method.

Methodological strengths and weaknesses

A total of 703 endorsements of the 10 higher order themes were made during interviews, observations and incident reports. The majority of these endorsements were made during interviews and reflects the benefit of this method of data collection over observations and incident

report analysis. The comparatively lower number of theme endorsements made during observations is likely to be partly due to the researcher's lack of clinical training which may have limited the identification of certain themes. It is also likely that some themes or components of higher-order themes manifest implicitly in the nursing environment and as such would be particularly difficult to observe. This is particularly true of the *ward climate* theme. For example, during interviews nurses discussed the ways in which the absence of a challenging climate on the ward might lead to nurses allowing doctors mistakes to go unchallenged. The *absence* of a certain behaviour would be difficult to observe without first being aware that a mistake had been made but had gone unchallenged which is likely to require significant training in drug chart review. Other aspects of ward climate were easier to observe and required no clinical training, such as professional regard and interruptions. Notably, 37 per cent of the total number of ward climate endorsements related to mid-task interruptions. However, only one of these endorsements was extracted from incident reports which stated that a nurse had been interrupted during medication administration which had led to a dosage error. This was the *only* endorsement of the theme ward climate from incident reports. It is possible that this lack of endorsement in incident reports was due to a nursing perception that although interruptions (or other ward climate problems) pose a significant risk to the safety of patients, they are a common occurrence so ingrained in hospital culture there is little point in citing it as the cause of error when reporting incidents.

Table 4.4.2.1 shows that only 9 per cent of the total number of theme endorsements were identified from incident reports. This reflects the general paucity of detail provided on incident forms on which staff are requested to provide only factual detail of the incident and not surmise causation. Where causation themes were present on incident reports, these tended to focus on local working conditions such as problems with pharmacy (e.g. expired stock and delivery issues), checking procedures (e.g. checking patient wristbands) and written communication problems such as unclear prescriptions. It is worth pointing out that although causation themes were extracted from incident

reports, causes were often implied rather than an explicit description. For example, one incident read '...patient prescribed insulin on insulin chart, 81 units prescribed and nearly given. Should have been 8.1 units...'. Although it is reasonable to assume this was a prescribing error and the decimal point was either not present ot not clear due to illegible handwriting, this is only implied on the incident report.

Another interesting finding is the relative lack of supervision and leadership endorsements overall but more importantly the prevalence of such endorsements within interviews compared with those during observations and incident reports. It is hypothesised that supervision and leadership issues would be particularly unlikely to be presented as a cause of error on an incident form since these reports would initially be read by immediate nursing managers and reporting is not anonymous. In one particular observation a nursing manager was observed shouting at two junior nurses to speed up a patients discharge procedure due to the imminent arrival of a new patient who needed the bed. As a result the previous patient's details remained written on the board above the bed for the duration of the shift at which point the new patient had been on the ward for more than five hours. Discussions with staff members during observations uncovered this nursing manager was 'always like this' and nurses suggested they would never ask her for advice or help or disclose any errors they had made.

There was a notable absence of any bed management or written policies and procedures endorsements during observations and incident reports. It is hypothesised that these particular themes represent 'organisational' factors which affect local working conditions. For example, it was postulated by nurses that constant pressure from bed managers and other departments to discharge patients quickly due to a back-log of potential 'breach' patients in A&E[29] and lack of

[29] This government scheme was initiated in 2004 and aimed to reduce waiting times in Accident and Emergency departments. The 4-hour rule states that 98 per cent of patients entering Accident & Emergency

available space in the Medical Admissions Unit often involved making significant compromises to the safety of existing patients. Several nurses suggested that this pressure frequently resulted in transferring patients 'in the middle of the night, in the dark' or 'in a real rush'; something nurses said was an 'accident waiting to happen'. The organisational effects of bed management on patient safety were viewed as indirect effects on local working conditions such as increased workload, interruptions, admission and discharge and checking procedures which were rushed or omitted altogether. The lack of observational endorsements of written policies and procedures was to be expected since no policies were developed, updated or disseminated during observational periods. Furthermore, it was considered highly unlikely that nurses would cite the workability or dissemination strategy of a policy as a cause for not adhering to it on an incident report. This is partly due to the fact that policies are created predominantly by managers and so, as with supervision and leadership, nurses would be extremely unlikely to formally assign responsibility for an error or violation on the workability of a policy or the way in which it was disseminated.

One limitation of the interviews was that due to the purposive method of sampling, it was not possible to interview any foreign nurses. Matrons did not recommend any of these nurses to be interviewed and suggested that these nurses had only worked on these wards for 12 months or less and as such were not senior or knowledgeable about ward practices. However, by conducting observations it was possible to gain insight into the issues surrounding foreign nursing (e.g. language issues, cultural differences) and raise these issues in interviews, which may not have been possible otherwise. Additionally, this study was conducted in only one NHS Trust which may to some extent compromise the generaliseability of findings. However, it was not possible at this stage of the research to recruit further hospitals due to ethical and time constraints. By conducting research in general medicine it was anticipated that the themes and issues observed extracted and discussed observations, incident review and interviews were common to multiple areas of care and

---

should be seen, treated and admitted or discharged within 4 hours of arrival. Hospitals who achieve this are awarded an incentive of £100,000.

also to secondary care environments in general. The fact that there was considerable research support for the themes and relative descendents generated in this study suggests that findings are generalisable.

It is apparent that data obtained during interviews was significantly richer and offered a more varied perspective of medication error causation than both incident and observational data. Interview data allowed for the identification of themes which were not evident during observations and from incident review. There were no higher order themes identified from observations and/or incident reports which were not proposed during interviews. It is possible that the significantly larger number of endorsements obtained during interviews was due to the size of data sources which may not have been equal. However, many of the themes which arose during interviews were not always directly observable (e.g. ward climate) and so conducting more observations would have been unlikely to yield more endorsements. Furthermore, the paucity of detail contained within incident reports and the bias with which events are reported suggests that by increasing the number of reports reviewed, the number of thematic endorsements would not have increased. This finding highlights the difference between triangulating data sources for the purposes of confirmation or completeness discussed in section 4.3.1. As predicted, the disparity between data sources prevented using one source to validate another. However, it is proposed that by conducting interviews, observations and incident report review, a more complete perspective of medication administration error causation was achieved.

4.4.2.2. Comparison of interviewee

Thematic endorsement frequency was also calculated on interview data to allow comparison between management and nurses' responses. Table 4.4.2.2 summarises these frequencies.

Table 4.4.2.2: Endorsement frequency by interviewee type

Theme	Endorsement frequency		
	Management (n=8)	Nurses (n=11)	
		Actual	Weighted*
Ward climate	35	104	76
Local working conditions	8	56	41
Workload	16	45	33
Human resources	23	47	34
Team communication	8	32	23
Routine procedures	17	32	23
Bed management	1	43	31
Written policies and procedures	24	14	10
Supervision and leadership	8	9	7
Training	9	9	7
TOTAL	**149**	**391**	**285**

*Weighted endorsements for nursing group calculated by dividing 'actual' endorsements by 11 (nursing N) multiplied by 8 (management N) to ensure management and nurse group sizes are comparable.

The purpose of interviewing both managers and nurses was to obtain a holistic perspective of the causes of medication errors from the blunt and sharp end of the organisation (Reason, 1997). As Table 4.4.2.2 shows, nurses suggested significantly more causes of medication error than managers. Even after weighting nurses' endorsements to ensure group sizes were comparable, the endorsement of all themes by nurses was almost double the number of endorsements made by managers. The data indicates that managers were more likely to cite two or three potential causes of medication errors, namely ward climate, human resources and written policies and procedures. By comparison, nurse endorsement of themes was considerably broader and more frequent.

Of particular interest was the bed management theme. This theme was endorsed during interviews on 43 occasions by nurses yet was mentioned as a potential cause of error only once by a senior manager. Added to the finding that bed management was *never* endorsed within incident reports, lack of managerial endorsement suggests that there could be a communication gap between the

sharp and blunt end of the organisation which prevents organisational learning. Although nurses perceived bed management to pose a significant patient safety risk, by not citing this as a potential cause of errors on incident forms (mainly because they are not asked to provide such details), senior managers seemed unaware of the albeit indirect implications of the management of beds. It is also possible that that managers are less likely to cite causal factors which are temporally removed from adverse events (Reason, 1997), such as bed management as potential causes of error due to their distance from the 'shop-floor'. As evidence of this, managers were more likely in interviews to cite causes which were much closer (in time or space) to unsafe acts such as nurses commitment to care and inadequate checking procedures. Interestingly, neither of these secondary themes was discussed during interviews by nurses.

4.4.3. Inter-rater reliability

In order to test the reliability of the 10 proposed themes and decrease subjectivity interpretation, inter-rater comparison were conducted to test whether selected excerpts (from interviews only) were reflective of the themes to which they had been assigned. Although there are more rigorous tests of reliability, such as asking raters to analyse raw data transcripts and extract themes in a similar way to the initial analysis, the complexity of thematic analysis and time constraints did not allow for this depth of testing. Since some themes were significantly more complex than others it was decided that secondary themes would also be tested to assess the extent to which other raters would place them in their assigned higher order theme. One clinical rater (former senior nurse) and a non-clinical rater (senior lecturer in health psychology) were recruited for this task but did not undertake it at the same time. In the first part of the task, raters were presented with a 25 per cent random sample of interview excerpts printed on small slips of paper (n=135) from each of the secondary themes. Secondary themes and corresponding definitions (n=37) were printed on A4 paper and were spread out on a large table. Raters were asked to read each excerpt and assign it to a theme

which they perceived best represented the excerpt. This task took approximately 1 hour to complete. Excerpts which were either placed in a theme which was not the one it had originally been assigned to during thematic analysis or those which raters struggled to categorise altogether were discussed to reach a consensus. The mean inter-rater agreement for the first task was 83 per cent.

In the second stage of the task, raters were presented with the 10 higher-order themes printed on A4 paper and were asked to assign each of the secondary themes to the one which they felt represented it the most. This task took approximately 30 minutes to complete. The mean inter-rater agreement for this task was 89 per cent. The level of agreement between the original placement of excerpts into secondary themes and secondary themes into higher-order themes and subsequent rater judgements was considered excellent given the recommendation that 70 per cent inter-rater reliability is considered adequate (Miles & Huberman, 1984).

4.5. Conclusions

Ten overarching preconditions of medication administration error were extracted from three sources of qualitative data. Overall, interview data comprised the richest data source and while observational data was useful to some degree as a complementary data source, incident reports offered very little insight into the causes of error. Nevertheless, review of incident reports did highlight a potential organisational learning gap through which opportunities for identifying error causation may be lost. Nurses offered significantly more explanations than managers during interviews as to why errors occur but are not asked to disclose these causes in current NHS reporting systems. Furthermore, poor reporting culture means that some if not most errors are not reported which further prevents organisational learning. It is proposed that unless a more open, honest, encouraging and less punitive stance is taken to error management, incident reports will

continue to be overly-brief and defensive, providing reduced opportunities for organisational learning.

Research evidence was found which supported the 10 themes and their sub-theme descendents identified within the qualitative data. However, the majority of this research reported *perceptions* of error causation amongst various health care professionals but very few had empirically measured whether these causes *predict* errors. Only by developing a valid and reliable measure of these preconditions and testing their relationship with medication or any other type of medical errors will appropriate interventions to reduce error be effective.

CHAPTER 5
DEVELOPMENT OF THE ORGANISATIONAL SAFETY QUESTIONNAIRE

5.1. Introduction

The primary objective of this thesis was to explore the organisational and workplace preconditions of drug administration errors. The secondary objective was to conceptualize this knowledge within an instrument which could effectively measure the presence (and hence absence) of safety in a secondary care environment. Using the qualitative data obtained in study 1 - study 2 explored the ways in which each organisational failure would manifest on a hospital ward. For example, what do poor communication or training issues look like on a particularly good or bad working day. This study involved content analysis of data obtained in study 1 to generate 'safety indicators' followed by two pilot studies to test the face validity of indicators in order to develop the OSQ (Organisational Safety Questionnaire).

5.2. Method

Although it was the original intention to use focus group sessions, similar to those employed by Groeneweg (1992; see chapter 2 for details) to elicit safety indicators of each causal factor there were a number of problems with this approach in a busy clinical setting. Firstly, initial discussions with nursing staff indicated that they would be unwilling to attend focus groups outside the hospital unless they were being paid for their time and expenses. Secondly, nurses said they would only participate in focus groups if it was during a shift that they were already working. Nursing managers said they would only be able to release two nurses at a time to attend a focus group for a period of 20 minutes at which point these nurses would have to return to the ward and two more nurses would be sent in. This approach was not feasible since remaining in the focus group for the

duration of the session is essential to the dynamics of this type of group work. It was decided that this would not promote the depth of discussion necessary.

As an alternative to collecting new data from focus groups, the data from the interviews was re-examined to see whether any useful indicators of each theme could be extracted. Using similar methods to those used to generate causal themes (described in full in chapter 4, section 4.3.3), all 11 nurse interview transcripts were re-read with a view to extracting explicit indicators of each causal theme[30]. Often, original excerpts which had led to the development of themes in study 1 were relevant indicators of poor safety. For example, one nurse stated '…I think it's an absolute scandal that we have to rely on doctors' handwriting that's often completely illegible... '. In study 1 this excerpt was used as an endorsement of the theme *team communication/written communication/handwriting legibility* (see Table 4.4.1.b) but was also an explicit example of the way that this theme could lead to errors and was translated into an objective indicator. Table 5.2. presents examples of translation of excerpts into indicators for each theme.

In line with error management tools developed in other industries (e.g. Tripod-DELTA, REVIEW, MESH) items needed to be objective, tangible indicators of safety failures which could be answered with either a yes or no (Groeneweg, 1992). As discussed in chapter 2, section 2.3.1.1, indicators were developed for the Tripod-DELTA measure during several studies which included detailed observations of employees during high-risk activities followed by several focus group sessions to discuss observed practices. A similar methodology was employed in this study using observations of nursing practices to inform subsequent in-depth interviews. Only indicators which were considered generalisable to multiple areas of secondary care were extracted. There was no limit imposed on the number of indicators generated since it was the intention to test the face validity of

[30] Manager's interview data was inappropriate for this stage of the analysis since these interviewees rarely gave specific examples of the ways in which latent variables could manifest at a ward level in the same way nurses did.

all indicators with a group of expert nurses to ensure that only those which were the best reflection of each theme were included on the final tool. However, due to the breadth of the data, more indicators were generated for themes which were more frequently endorsed (i.e. ward climate, local working conditions). Indicators were developed which reflected the thematic structure of each higher-order theme. For example, 14 items were developed for the complex higher-order theme local working conditions which comprised 5 fatigue and body rhythm items (3 shift patterns and 2 staff breaks items), 1 equipment availability item, 2 ward noise items and 6 pharmaceutical items relating to expired stock, ward stock levels and labelling and packaging. Comparatively, the simpler theme supervision and leadership comprised only 10 items which related to either task delegation or leadership style (see chapter 4, tables 4.4.1.a-j for breakdown of each latent theme).

Indicators developed for the Tripod-DELTA tool (Groeneweg, 1992) were framed to reflect the occurrence of problems across a *standard* time period (e.g. in the last month/accounting period). However, it was evident from interviews in this study that there was a large variation in the frequency of occurrence of each causal theme. For example, nurses suggested that being interrupted (an indicator for the theme ward climate) was a daily occurrence. Conversely, problems which related to the dissemination of policies and procedures were significantly less frequent, although equally problematic, and occurred only a few times each year. Asking participants whether they had been interrupted or received a policy they did not understand in the last four weeks would undoubtedly have resulted in ceiling and floor effects respectively. It was not practical therefore to frame all items in exactly the same way. To account for this, each indicator was framed in terms of its relative frequency as indicated during interviews. This was particularly difficult due to the wide variety of indicators. However, a pilot study was conducted to examine the validity of these time frames (see section 5.3). All items were negatively framed since there is evidence which suggests negatively worded items can reduce response bias (Nunnally & Bernstein, 1994). Although it is common in test construction to develop both positive and negative items, Schmitt and Stults (1985)

showed that this can create problems in subsequent factor analysis when there are a large number of items (as was the case in this study) and participants have failed to recognise the reversal of wording. Therefore, all items (except four which were neutral) were negatively framed as safety *failures*.

Table 5.2: Examples of the translation of raw excerpts into safety indicators

Causal theme	Qualitative raw extract	Translated indicator
Ward climate	*'...when you come on to a shift you like to think that you're starting with a clean sheet and if that clean sheets blotted from the start cos' you've to finish chores that the previous staff should have done it doesn't always go down very well so people rush round trying to get their chores done...'*	In the last 2 months, I have rushed through outstanding jobs towards the end of my shift so the next shift did not have to do them
Local Working Conditions	*'...half the time I haven't even had anything to eat all shift and then it becomes difficult to concentrate even on minor tasks...'*	In the last 2 months, I have worked 6 or more hours into my shift without anything to eat or drink
Workload	*'...lots of patients being admitted, or very sick patients in need of a lot of attention, dealing with relatives and visitors. It can fluctuate throughout the day. It's difficult to plan a shift as it's so unpredictable what's going to happen at what time....'*	In the last 2 months, I have found it difficult to prioritize duties for the day because all tasks seemed equally important
Human Resources	*'...there's very little medical support on nights. You've got to practically beg. You need a really good reason as to why you need a doctor up here at night...'*	In the last 4 weeks, I have experienced difficulty getting an on-call doctor to come to this ward out of hours
Team Communication	*'...one potential systems error on every ward I've ever worked on is that drug charts tend to be free of the nursing or medical notes. They get lost and people have to try to remember what was on them and rewrite them. Doctors carry them round all day and charts often get tidied away in the wrong patient's notes.*	In the last 4 weeks, I have discovered a drug chart belonging to one patient in a different patient's nursing/medical notes/bedside locker/bed

Table 5.2: continued

Causal theme	Qualitative raw extract	Translated indicator
Routine Procedures	*'...people do mad things like put a patients identity bracelet and the allergy bracelet on different wrists so you pull the arm out of bed that you think you're looking for and see the identity band but not the allergy band...'*	In the last 2 months, I have encountered a patient on this ward who's allergy bracelet was on the opposite wrist to their ID wristband
Bed Management	*'...we sometimes have to transfer patients to hospital X in the middle of the night because we need their bed. The potential for things going wrong packing people up in the dark when there's not as many staff on you can't fathom...'*	In the last 3 months, this ward has transferred a patient to another ward during the night or early hours of the morning because their bed was needed for another patient
Written Policies and Procedures	*'...many a time I'll look at the new version [of an existing policy] but I've no idea what's different on this one, what have they changed?'*	In the last 12 months, I have received an updated version of a written policy, protocol or guideline without knowing which part of the original had changed
Supervision and Leadership	*'...there are leaders who are good and can organise and sort things out and delegate quite well. There are others who don't delegate, who take on too much, don't let others help and things get missed, jobs don't get done which adds to the workload...'*	In the last 2 months, there has been an important task outstanding at the end of the shift which I didn't do because I thought another nurse had done it
Training	*'...you're supposed to get a preceptorship package and induction when you start where you get loads of information and things. I never did, I was thrown in at the deep end. People do sometimes start work before they've started their induction...'*	In the last 12 months, **all** newly qualified nurses starting on this ward/unit/dept have received the standard induction training programme required

A total of 145 indicators were developed across all 10 themes which could be answered with either yes or no. It was especially difficult to develop 'tangible' indicators for the ward climate theme since many of these excerpts referred to a general perception or implicit feeling of 'the way things are' on the ward. Developing tangible indicators necessitates removal of emotive words such as think, believe or feel to avoid subjective interpretation by respondents. For instance, nurses suggested during interviews that there were some doctors they did not 'get along with' which was assigned to the secondary theme professional regard. This would have been difficult to translate into an observable indicator without knowing how not getting along with another professional is manifested.

5.3. Pilot study 1

An initial pilot study was conducted to assess the face validity of all 145 generated safety indicators. The collaborating partners were keen to produce a tool which was considerably shorter than 145 items based on concerns for the utility and response rates from excessively long instruments. The aim of the pilot therefore was to also highlight indicators which were the most useful or represented the most risk to patient safety for inclusion in the final version.

5.3.1. Method

Participants

Twenty-four qualified nurses who worked on four non-specialist wards and who had not previously been involved in interviews (e.g. general medicine, medical admissions, and elderly medicine) were invited by letter to participate in the initial pilot study. Participants were advised that the study was testing the workability of a new questionnaire which had been developed by researchers at the University of Leeds to measure aspects of their working environment which might promote risks to

patient safety. In agreement with the hospital's Research and Development office it was agreed that because the pilot study was expected to take nurses approximately two hours to complete they could be provided with lunch and offered £10 each to participate. Ten nurses agreed, comprising one matron, three sisters, four senior staff nurses and two staff nurses.

Questionnaire and task design

Two tasks were designed to measure face validity of the indicators and questionnaire as a whole. Participants were first asked to simply complete the 145 item questionnaire which comprised indicators presented under their parental themes in tables on double-sided A4 paper in a booklet format (Dillman, 2000; Sudman & Bradburn, 1982; Bourque & Fielder, 1995). The name of each theme and its definition (see chapter 4, section 4.4.1 for definitions) was presented at the top of each table. Participants were instructed to read each question carefully and respond by circling either 'yes' or 'no', the event had or had not happened on their ward in the stated time scale or circle 'not applicable' if they did not know. Table 5.3.1 gives an example of this presentation. The final version of this questionnaire is included in the appendices (appendix XII).

Table 5.3.1: Questionnaire presentation with corresponding response options

TEAM COMMUNICATION **Aspects of the ways in which ward teams and intra-professional teams communicate. Includes written (i.e. document design, tracking and legibility) and verbal (i.e. inter and intra-departmental information sharing) communication, team size and multi-cultural communication issues.**		**Yes / No / Not Applicable**		
Q1	In the last 4 weeks, have you sought assistance from a colleague to interpret the handwriting on a patient's drug chart?	**Y**	**N**	**NA**

In line with face validity studies conducted to develop the Tripod-DELTA error management tool (Groeneweg, 1992), the second task invited participants to make comments on any questions they considered problematic. Participants were given an answer sheet and asked to provide details on problematic questions under the following four headings;

- **Clarity** – whether the item was clearly worded and understandable using appropriate terminology
- **Time frame** – whether the stated time frame was appropriate (i.e. 'In the last 4 weeks... ')
- **Patient safety impact** – whether the item would impact on the safety of patients
- **Inappropriate category** – whether the item was not reflective of the theme to which it had been assigned or more related to another theme

Additional space on the answer sheet was provided for participants to comment on any other changes they thought would improve the questionnaire or any aspects of patient safety they thought were not addressed within the items.

Procedure

Once nurses had agreed to participate, a meeting was scheduled with the researcher to discuss the purpose of the study and what was required of participants. This was considered particularly important since these nurses had never met the researcher before, having not been involved in the previous study. It was hoped that this meeting would enable participants to feel more at ease giving feedback on the questionnaire, something they may have been reluctant to do if they thought the questionnaire had been designed to assess performance. Participants were asked to complete the questionnaire first, followed by the second task and were given two weeks to complete both. After this deadline, participants were revisited to collect questionnaires and to discuss any issues they wished to raise.

5.3.2. Results

Once questionnaires had been completed, a debriefing meeting was scheduled to reassure participants their data would be confidentially treated and for participants to provide any feedback regarding the questionnaire in general. Overall, nurses were positive about the questionnaire. Items were removed if five or more participants (50 per cent) agreed the item was unlikely to impact on patient safety (n=9)[31]. Items were also moved to recommended categories if five or more respondents agreed on a more appropriate placement (n=3). Items judged unclear by *any* respondent were amended to improve its understanding. Several participants commented that some items were too similar and seemed to be measuring the same thing. Again, when five or more respondents agreed this was problematic for a particular item, the item was either removed or consolidated with its similar counterpart (n=7). A further 15 items were removed because they had all been answered 'not applicable'.

Of particular interest was the relative lack of variability in responding whereby most participants answered the majority of items with a yes response. This was encouraging to some extent in that it suggested the indicators represented commonly occurring safety failures and were less likely to be a subjective interpretation of qualitative data. However, in terms of the final tool this lack of variability would present statistical problems in discriminating between wards and determining the predictive validity of the items. It is possible that this lack of variability in responding was due in part to the item time frames since there were a number of comments made by participants that these were generally underestimated. For example, one indicator asked participants whether in the last 4 weeks they had been interrupted during the drug round by a patient, visitor or other health professional. The vast majority of participants responded yes to this question. It was possible that by reducing the time frame to 'in the last week' the sensitivity of such items would be increased. However, observations, interviews and pilot study responses revealed that frequent interruptions

[31] Lawshe (1975, pp. 568) recommended that if more than half the raters in a pilot study indicate an item is essential or relevant to the construct being measured, that item has face/content validity.

during the drug round by patients, visitors and other health professionals are a daily occurrence. Even by reducing the time frame to 'in the last shift', ceiling effects on using dichotomous response options would always occur. This posed a significant problem for items which represented frequent and significant risks to patient safety compared with those which were equally problematic but not frequent (e.g. poor knowledge of current policy driven by ineffective dissemination of written policies and procedures).

Development of the Tripod-DELTA tool (Groeneweg, 1992; see chapter 2 of this thesis) principally aimed to detect latent failures in an organisation with a low rate of adverse incidents compared with those occurring in health care. Research in general has shown a much lower incident rate and an 'almost complete absence of catastrophic failures over many years' in many high-risk industries (Reason, 2000, pp. 769). The much higher rate of errors and other adverse incidents in health care suggests the prevalence of latent failures is likely to be much greater than in other high-risk industries which is supported by the ceiling effect responses observed in this pilot study. In the development of other error management tools in aircraft maintenance (MESH) and railway operations (REVIEW), Reason (1997) notes that ceiling and floor effects led to researchers employing frequency rating scales as opposed to more restrictive 'yes' and 'no' options. Although rating scales are generally viewed as being more subjective, Reason emphasises that individuals are extremely competent at judging the frequencies of different events. Indeed studies have shown correlations between estimates of the frequency of events and the *actual* number of these events to be +0.9 or above (Hasher & Zacks, 1984). Since it was not possible to explore further the relative frequency of each indicator, response options were changed to a Likert frequency scale to enable comparisons between wards (see pilot study 2). Colla, Bracken, Kinney and Weeks (2005) reviewed nine measures of aspects of organisational safety in health care settings and revealed all had employed a five-point Likert scale and so five response options were developed for this tool from 'not at all' to 'nearly all the time'.

After removal of items, the questionnaire was reduced from 145 to 114 items. It was the original research aim to develop a tool which respondents would be happy to complete every 6-12 months and which would take approximately 15 minutes to complete. The collaborating institute had concerns that the 114 item revised questionnaire was still too long for this purpose based on previous experiences of poor responding with longer questionnaires. A further concern was that some of the items might not be generalisable beyond general medicine or elderly care. To address these concerns, a second pilot study was conducted.

5.4. Pilot study 2

Pilot study 2 was conducted to examine whether the tool had face validity beyond elderly care and general medicine wards, whether it could be completed in less than 20 minutes and whether nursing staff would be prepared to complete the tool every 6 months.

5.4.1. Method

Participants

Fifty qualified nurses who worked on eight wards not involved in earlier phases which comprised surgical, neonatal, paediatrics and gynaecology wards and a large medical admissions unit (MAU) were invited by letter to participate in this pilot study. As with pilot study 1, participants were advised that the study was testing the workability of a new questionnaire which had been developed by researchers at the University of Leeds to measure aspects of their working environment which might promote risks to patient safety. Participants were offered paid lunch and a £10 incentive to participate. Twenty nurses agreed comprising six sisters, ten senior staff nurses and four staff nurses from two surgical and two paediatric wards and the MAU.

Questionnaire design

As with pilot study 1, the revised 114-item questionnaire was printed in a double-sided A4 booklet with indicators presented under each of their parental themes in separate tables. The name of each theme and its definition was presented at the top of each table. Beneath each indicator was a Likert-type frequency scale. As with pilot study 1, participants were instructed to read each question carefully and respond by circling the frequency with which each event had happened on their ward in the stated time scale or circle 'not applicable' if they did not know. Response options were not at all, occasionally, quite often, frequently and nearly all the time.

Procedure

Meetings were held between the researcher and each participant to discuss the purpose of the research and to describe the tasks they were required to undertake. Essentially, this pilot study was identical to pilot study 1, asking participants to complete the questionnaire and make comments on problem items for clarity, time frames, impact on patient safety and category 'belongingness'. In addition, participants were also asked to rate the questionnaire on overall presentation, length and ease of use as being 'poor', 'fair', 'good' or 'excellent'. They were also asked whether they would be prepared to complete the questionnaire every 6 months and if not, to provide details of their main reason for not doing so. A final question asked respondents how long it took them to complete the questionnaire.

5.4.2. Results

Face validity

In terms of item clarity, very few comments suggesting any items were unclear were made. Items were removed if 10 or more respondents (50 per cent) suggested they would be unlikely to affect the safety of patients (n=4). A further three items were removed which were judged too similar to other items. No comments or recommendations were made by participants on the belongingness of any item to its respective category.

Variability

Analysis revealed that even with the revised five point response scale, five items had no variability in responses and a further eight had extremely limited variability of responses and were subsequently removed. Interestingly, four of these removed items related to reporting culture on the ward and asked participants whether factors such as not being directly involved in the event or fear of disciplinary repercussions had prevented them from reporting an error they had made or witnessed. All 20 participants responded to these questions with 'not at all'.

Notably, three of the eight reduced variance items referred to the impact of supervision and leadership on patient safety, in particular the leadership style of immediate supervisors. These items asked respondents whether they had been advised by a supervisor in the last two months to 'cut corners' or go against standard practice in order to complete an aspect of patient care. Although cited frequently and openly during interviews, most nurses in this pilot study responded 'not at all'. It is possible that reassuring participants that their questionnaire responses were anonymous, and would not be seen by immediate nursing managers was not effective and nurses were reluctant to make what could be perceived as disparaging comments about their supervisor. Additionally, these items were more likely to be construed by respondents as admitting they had committed violations of standard procedures which they may be unwilling to disclose.

Alternatively, it is also possible that for this group of respondents, there were fewer problems with their supervisors than in the group of nurses who were interviewed.

Questionnaire design

All 20 respondents rated the overall presentation and layout of the questionnaire and the ease of use as either 'good' or 'excellent'. Only one respondent rated the length as being 'poor', six rated it 'fair' and the remaining thirteen respondents rated it as 'good'. When asked whether they would be happy to complete the questionnaire once every six months seventeen participants said yes and three said no but did not provide details why not. On average respondents stated the questionnaire had taken approximately 20 minutes to complete but several nurses commented this was difficult to judge because they were constantly interrupted.

5.5. Methodological concerns

After analysis of pilot study 2 data, the revised questionnaire comprised 94 items which it was estimated would reduce the completion time to 15 minutes. It was the intention to administer this questionnaire to a large sample of qualified nurses in study 4 (chapter 7) in order to measure multiple aspects of validity and reliability. This study would have enabled revision of the number of items using factor analysis to produce the final error management tool. However, the collaborating institute suggested that previous attempts at administering questionnaires of this length at this hospital had resulted in poor response rates. A review of nine similar measures of patient safety climate by Colla, Bracken, Kinney and Weeks (2005) revealed that the number of items to be completed in these measures was (with one exception) less than 71 with an average test length of 43 items. Several options were discussed in order to reach a compromise between producing a tool which would meet the pragmatic requirements of usability in busy clinical settings and which would also be valid and reliable and proactively predict medication errors.

It was considered essential that each latent theme was represented by as many indicators as possible which would enable redundant items to be removed and retention of only items with the highest factor loadings and to achieve the highest alpha levels. Firstly, it was proposed that the questionnaire could remain as 94 items[32] providing the available sample size was large enough to accommodate poor response rates. It was predicted that approximately 300+ participants would be required for factor analysis with this number of items (Kass & Tinsley, 1979; Tabachnick & Fidell, 2001)[33]. Accounting for the poor response rates of approximately 20 per cent that are common within nursing research would mean that approximately 1500 qualified nurses would need to be sampled which would involve recruitment of at least one more hospital. Although two other hospitals were contacted and showed initial interest in being involved, conflicting research commitments at one hospital and an inability to commit to the study within the necessary time scale of the PhD at the other hospital meant that recruiting nurses beyond the collaborating hospital was not possible.

The second possible option to meet the needs of the hospital and the researcher might have been to reduce the number of latent variables measured. It was proposed that the tool could have been a measure of three to five latent preconditions of MAEs (as opposed to one which measured ten), each with approximately ten indicators. However, this would involve making a judgement on the themes which were most or least important predictors of medication errors. It was difficult to ascertain from interviews exactly which latent factors were the most predictive of error. Additionally, this was the purpose of the final study. Since it was a main aim of the research to develop a tool which was a broad reflection of multiple latent causes of MAEs, reducing the tool would have produced a much narrower perspective of latent failures and would be less likely to

---

[32] Although not discussed in this chapter, it became apparent at this point that additional items would need to be developed in order to measure the tool's predictive validity due to inherent reporting bias in formally reported incidents. See chapter 6 for full discussion of this study.

[33] Kass & Tinsley (1979) recommend having between 5 and 10 participants per variable up to a total of 300 at which point parameters tend to be fairly stable regardless of participant to variable ratio. Similarly Tabachnick & Fidell (2001) agree it is 'comforting to have at least 300 cases for factor analysis'.

represent overall organisational safety. Reason (1997) argues that errors are more likely to occur as a result of a culmination and complex interaction of multiple latent failures. A tool which measures only a limited number of these failures is unlikely to capture the complexity of this relationship.

The third option discussed to reduce the length of the questionnaire was to further reduce the number of items within each of the 10 latent factors. The disadvantage of this was that individual items were considered as representing scales measuring each underlying latent theme. The premise of scale and indeed questionnaire development is to include more items than one would expect to measure each construct, thus allowing for the removal of weak and unrelated items during factor analysis. Clark and Watson (1995) advise that researchers should 'err on the side of over-inclusiveness' and ensure that each construct is well represented in the initial item pool. However, the problem of having too many items was complicated by an additional need in study 4 to add another 20-25 items which would measure the both the predictive and concurrent validity of the tool (see chapter 6). Since inter-rater agreement that excerpts were reflective of each theme had been high and face validity of each translated indicator had yielded positive results, it was decided that reducing the number of items per latent factor was the best option of the three proposed. All three sources of qualitative data were revisited in order to remove items (n=12) which were less frequently endorsed overall (from interviews, observations and incident review) within each higher-order theme and also those less frequently endorsed within descendent themes (see chapter 4, section 4.4.2). Items which were frequently endorsed but which would make intervention difficult were also removed. Table 5.5.a illustrates how this was done for the theme 'local working conditions'.

Table 5.5.a: Content analysis and removal of indicators in the theme local working conditions

Secondary theme	Tertiary theme	Definition	Indicator	Total no. of endorsements	Removed or retained
Pharmaceutical issues	Ward stock levels and ordering procedures	**Inadequate levels of stock drug items and procedures of ordering replenishment stock**	*In the last **2 months**, a patient on this ward has missed 3 or more doses of their prescribed medication because it had been ordered but not sent from pharmacy*	**13**	**Retained**
	Labelling and packaging	**Outer labels or packaging which do not correspond to the drug contained within**	*In the last **2 months**, I have encountered boxed medication which did not correspond with the outer package labelling*	**13**	**Retained**
	Expired stock	**Medication stored in drug trolleys and clinic rooms provided by Pharmacy which has passed expiration dates stated on packaging**	*In the last **3 months**, I have encountered medication (e.g. boxed or IV) in the drug trolley or clinic room which had expired by more than a month*	**5**	**Retained**
	Drug names	**Drug names which are difficult to remember or too similar to other drug names**	*In the last 4 weeks, I have encountered medication with a name almost identical to another drug in the drug trolley or clinic room*	**3**	**Removed** Least frequently endorsed within secondary category 'pharmaceutical issues'

Table 5.5.a: continued

Secondary theme	Tertiary theme	Definition	Indicator	Total no. of endorsements	Removed or retained
Fatigue and body rhythms	Shift patterns	**Working multiple and/or consecutive extended shift lengths and working during the night shift**	*In the last **3 months**, I have worked 3 or more 'long shifts' (7am – 9pm) one after another*	**13**	**Retained**
	Rest breaks	**Inadequate opportunities to rest, drink or eat during the course of a shift**	*In the last **2 months**, I have worked 6 or more hours of my shift without anything to eat or drink*	**9**	**Retained**
Patient issues	Illness complexity and medication volume	**The complexity and/or co-morbidity of patient illnesses which leads to an increase in the number of required medications**	*In the last 4 weeks, I have encountered a patient who was prescribed more than 10 medications to be administered at the same time*	**7**	**Removed** Considerably difficult to intervene
	Patient familiarity	**An over-reliance on familiarity with patients which reduces reliance on formal checking procedures**	*In the last 2 months, I have administered a medication to a patient without checking their details first because they have been on the ward for a long time and I know them well*	**6**	**Removed** Considerably difficult to intervene
	Non-compliance	**Patients who refuse to take medication or cannot be found at the time of administration**	*In the last 4 weeks, I have omitted a patient's medication because I couldn't find them at the time of administration*	**5**	**Removed** Considerably difficult to intervene

Table 5.5.a: continued

Secondary theme	Definition	Indicator	Total no. of endorsements	Removed or retained
Ward design	**Aspects of the way the ward is designed and beds are positioned which promote inadequate checking and interruptions**	*In the last 2 weeks, having to go to the clinic room for medication not in the drug trolley has considerably slowed my progress during the drug round*	**9**	**Removed** Although frequently endorsed the design of wards and positioning of clinic rooms and patient beds is unlikely to be targeted for intervention on the basis of findings from this questionnaire
Equipment availability and design	**Equipment which is either difficult to locate or obtain at the time of need or designed in a non-user-friendly way**	*In the last **2 months**, I have been unable to get hold of IV medication equipment because I could not find it in the hospital or it was all in use on other wards*	**9**	**Retained**
Ward noise	**Increase levels of noise during complex tasks which promote**	*In the last 2 weeks, I have been unable to concentrate during the drugs round because the ward was so noisy*	**7**	**Removed** It was considered that ward noise was due to other factors such as timing of other health professionals tasks which were best represented in other items
Personal issues	**Aspects of nurses' personal life which affect their ability to work to their usual standard**	*In the last 4 weeks I have found it difficult to perform my job to the best of my ability due to personal problems*	**3**	**Removed** Less frequently endorsed in higher order theme and a difficult target for intervention

This final review led to a further 12 items being removed. Table 5.5.b shows the remaining 82 items which comprised the Organisational Safety Questionnaire.

Table 5.5.b: Organisational Safety Questionnaire items (n=82)

Item code*	TEAM COMMUNICATION
TC1	In the last **4 weeks**, a patient's clinical investigation paperwork (e.g. blood test, e.c.g report etc.) has been sent to this ward/unit/dept and subsequently lost
TC2	In the last **4 weeks**, I have discovered a drug chart belonging to one patient in a different patient's nursing/ medical notes/bedside locker/bed
TC3	I have attended a ward/unit/dept meeting in the last **3 months**
TC4	In the last **4 weeks**, I have encountered a patient's allergy status written on their main drug chart but not on their PRN chart
TC5	In the last **4 weeks**, a patient's drug chart has gone missing from this ward/unit/dept and had to be rewritten
TC6	In the last **4 weeks**, I have encountered a drug chart where the details of a particular drug (e.g. dosage, frequency/route of administration) had been crossed out and altered 2 or more times
TC7	In the last **4 weeks**, I have sought assistance from a colleague to interpret the handwriting on a patient's drug chart
TC8	In the last **4 weeks**, a doctor or other health professional has written instructions in a patients notes for the nursing team to carry out without verbally advising them they had done so
TC9	There has been an incident debriefing or team meeting to discuss the *last* 'significant' clinical incident to happen on this ward/unit/dept within one month of its occurrence
TC10	In the last **4 weeks**, I have encountered a patient's drug chart which had been changed by a doctor since the previous drug round without my knowledge
	ROUTINE PROCEDURES
RP1	In the last **2 months**, I have encountered details on a patient's ID wristband which did not correspond with their drug chart
RP2	This ward regularly *uses* a standardised list (**written document**) of tasks which should be carried out for each new admission
RP3	This ward *uses* a standardized list (**written document**) of information which *must* be passed on to colleagues during patient handover
RP4	In the last **2 months**, I have encountered a patient on this ward whose allergy wristband was on the opposite wrist to their ID wristband
RP5	In the last **2 months**, I have had to telephone the ward after my shift had ended because I remembered information I had previously forgotten to pass on during handover

RP6	In the last **2 months**, I have experienced difficulty writing all the information passed on to me during patient handover because I didn't hear or understand everything or there wasn't time to write everything
RP7	In the last **2 months**, I have encountered a patient who had been on the ward for more than 24 hours who did not have *this ward's* ID wristband on
RP8	In the last **2 months**, I have encountered a patient on this ward wearing an allergy wristband whose allergy was not documented on their drug chart or vice versa
RP9	In the last **2 months**, I have come across details written on the board above a patient's bed for the patient who had *previously* occupied that bed
	WORKLOAD
WL1	In the last **2 months**, I have been unable to coordinate the ward/unit/dept because of my patient allocation
WL2	In the last **4 weeks**, I have sped up during a drug round to avoid a backlog of jobs later in the shift
WL3	In the last **4 weeks**, I have sped up during a drug round to avoid making the next drug round late
WL4	In the last **4 weeks**, the number of doctors arriving for morning ward round has slowed my drug round progress
WL5	In the last **2 months**, I have found it difficult to prioritize duties for the day because all tasks seemed equally important
WL6	In the last **2 months**, 2 or more new admissions have arrived within minutes of each other on this ward
WL7	In the last **4 weeks**, an *unexpected* new admission on this ward has meant that other important tasks were incomplete by the end of the shift
	HUMAN RESOURCES
HR1	In the last **4 weeks**, I have worked a shift when an agency nurse was requested but did not arrive, leaving the ward/unit/dept understaffed
HR2	In the last **2 months**, as a qualified nurse I have had to leave the ward to accompany a patient to another department leaving the ward short of qualified nurses
HR3	In the last **2 months**, a doctor has refused to carry out something I needed them to do because the patient concerned was not theirs
HR4	In the last **2 months**, I have worked with an agency nurse who did not know how to carry out a skill or procedure required of their nursing grade
HR5	In the last **2 months**, an agency nurse has been sent to this ward/unit/dept who was less qualified than the grade of nurse we requested
HR6	In the last **6 months**, there has been a deterioration in the level of experience within the team on this ward/unit/dept
HR7	In the last **4 weeks**, I have bleeped a doctor 2 times or more without receiving a response
HR8	In the last **6 months**, there has been an increase in senior nursing turnover on this ward/unit/dept

HR9	In the last **4 weeks**, I have experienced difficulty getting an on-call doctor to come to the ward out of hours
	LOCAL WORKING CONDITIONS
LWC1	In the last **3 months**, I have worked 3 or more 'long shifts' (7am – 9pm) one after another
LWC2	In the last **2 months**, I have worked 6 or more hours of my shift without anything to eat or drink
LWC3	In the last **2 months**, I have encountered boxed medication which did not correspond with the outer package labelling
LWC4	In the last **2 months**, I have been unable to get hold of IV medication equipment because I could not find it in the hospital or it was all in use on other wards
LWC5	In the last **2 months**, a patient on this ward has missed 3 or more doses of their prescribed medication because it had been ordered but not sent from pharmacy
LWC6	In the last **2 months**, a patient from this ward/unit/dept has been discharged without taking home their medication because it had been ordered but not been sent by pharmacy
LWC7	In the last **2 months**, I have worked an entire shift without a rest break
LWC8	In the last **3 months**, I have encountered medication (e.g. boxed or IV) in the drug trolley or clinic room which had expired by more than a month
LWC9	In the last **4 weeks**, I have encountered drugs packaging for two different types of medication which was virtually identical
	WARD CLIMATE
CLT1	Senior nurses on this ward/unit/dept prefer *all* patient queries to come through them in the first instance – even if there are other nurses on shift who would know the answer
CLT2	In the last **2 months**, a senior nurse has complained that jobs from the previous shift were left unfinished
CLT3	In the last **2 months**, I have refrained from bleeping a particular doctor because I wasn't confident they had the necessary skills to help me
CLT4	In the last **2 months**, I have been asked by a consultant to carry out a task which was really the job of a junior doctor rather than a nurse
CLT5	In the last **2 months**, I have rushed through outstanding jobs towards the end of my shift so the next shift did not have to do them
CLT6	In the last **4 weeks**, I have had to break off from a drug round because I was interrupted by another health professional (e.g. doctor, nurse, physio)
CLT7	In the last **2 months**, I have refrained from bleeping a particular doctor again because I knew they would be annoyed I had bleeped them several times that shift
CLT8	In the last **3 months**, I have made or witnessed an error which I did not 'formally' report because the patient wasn't harmed by it in any way
CLT9	In the last **6 months**, I have formally reported a safety concern on this ward/unit/dept without receiving any feedback
CLT10	In the last **3 months**, I have followed instructions against my better judgement because a

	doctor or more senior nurse told me it was what I should do
CLT11	In the last **2 months**, a doctor on this ward has refused to wait until I was finished with a particular task before conducting their ward round without me
CLT12	In the last **3 months**, I have experienced difficulty challenging a doctor about a prescription I thought they had written incorrectly
	BED MANAGEMENT
BED1	In the last **3 months**, a patient *from this ward* has been 'slept out' elsewhere but had to return the next day because they were unwell
BED2	In the last **3 months**, I have been pressured into arranging to discharge a patient quicker than had been originally planned
BED3	In the last **3 months**, this ward has transferred a patient to another ward during the night or early hours of the morning because their bed was needed for another patient
BED4	In the last **3 months**, this ward has had an 'inappropriate admission' because A&E did not want to breach the 4-hour waiting rule
BED5	In the last **3 months**, this ward has had to rush an existing patient's discharge to admit a new patient from MAU or A&E to avoid a patient breaching the 4 hour rule in A&E
BED6	In the last **3 months**, a patient has been discharged from this ward before all the necessary agencies needed to care for them at home (e.g. community nurse, social worker) had been arranged
BED7	In the last **3 months**, a patient has been 'slept out' *to this ward* who needed specialised care I thought this ward could not effectively deliver
	SUPERVISION AND LEADERSHIP
SP1	*All* senior nurses on this ward/unit/dept demonstrate a safe example of patient care for staff to follow
SP2	*All* of the senior nurses on this ward/unit/dept are approachable if a nurse needs help or advice with a work issue
SP3	In the last **2 months**, a senior nurse has advised me to cut corners in order to complete all jobs planned before the end of the shift
SP4	In the last **2 months**, I have been confused as to what jobs were supposed to be my responsibility during the course of a shift
SP5	In the last **2 months**, senior nurses have been rushing around the ward/unit/dept completing most tasks while other nurses have less to do
SP6	In the last **2 months**, there has been an important task outstanding at the end of the shift which I didn't do because I thought another nurse had done it
	TRAINING
TR1	In the last **6 months**, I have performed a clinical procedure or used a piece of equipment for which I was not formally trained
TR2	In the last **6 months**, I have been taught a clinical procedure or how to use a piece of equipment by a nurse who had not received the relevant 'formal' training

TR3	In the last **12 months**, I have mentored a newly qualified nurse *before* I have attended mentorship and teaching and assessment programs
TR4	In the last **12 months**, *all* newly qualified nurses starting on this ward/unit/dept have been given and have worked through a preceptorship pack
TR5	In the last **6 months**, I have booked a training course or study day which was cancelled due to budgetary constraints
TR6	In the last **6 months**, I have booked a training course or study day which was cancelled due to a clash with work requirements or staff shortages
TR7	In the last **12 months**, *all* newly qualified nurses starting on this ward/unit/dept have received the standard induction training programme required
	WRITTEN POLICIES AND PROCEDURES
WPP1	In the last **12 months**, there has been discussion of a policy or guideline that I was not aware existed until *after* an incident had occurred
WPP2	In the last **12 months**, I have signed to say I have read and understood a new policy, procedure or guideline when I haven't either read it or understood it
WPP3	In the last **12 months**, I have received an updated version of a written policy, protocol or guideline without knowing which part of the original version had changed
WPP4	In the last **12 months**, I have been asked to read a written policy, protocol or guideline but only skimmed it to get the gist because it was too long
WPP5	In the last **12 months**, I have encountered a written policy or procedure that I thought would be impossible to follow in the course of my job
WPP6	In the last **12 months**, all the reasons have been made clear why a new procedure or policy has been introduced

*** Item codes provided for the purposes of cross-reference in chapter 7**

5.6. Conclusions

The aim of this study was to conceptualize latent preconditions of medication administration error identified in chapter 4 within an organisational safety tool. Employing qualitative data collected in study 1, this study explored the ways in which each organisational failure would manifest on a hospital ward. Using content analysis of qualitative data obtained in study 1, 145 safety indicators were generated based on these manifestations of poor safety. Pilot studies to test the face validity of indicators and subsequent content analysis to remove less commonly endorsed items led to refinement of the tool to develop the 82 item Organisational Safety Questionnaire (OSQ).

CHAPTER 6
DEVELOPMENT OF THE DRUG ROUND BEHAVIOUR QUESTIONNAIRE

6.1. Introduction

In order to show predictive validity, the proposed Organisational Safety Questionnaire (OSQ; see chapter 5) should be able to predict medication administration errors. However, this depends on there being a readily available valid measure of these types of error. This chapter describes current NHS incident report data which are commonly employed within health care research as error outcome measures. Multiple drawbacks and biases inherent within such formal reporting systems are discussed to emphasise the need to develop an independent error measure which is not reliant upon such organisational reporting systems. To inform the development of an independent measure of medication errors, medical and psychological literature were reviewed and relevant evidence will be presented here which has attempted to understand and categorise everyday errors, medical and medication errors. Study 3 will then be described which employed the qualitative data obtained in study 1 to explore the types of medication administration errors which can arise in secondary care as a direct or indirect result of the ten latent preconditions identified in chapter 4. The ultimate aim of this study was to develop a measure of drug administration errors which was not reliant upon adverse patient outcome which current systems of formal reporting would be unlikely to identify.

6.2. NHS incident reporting: Problems and biases

As was noted in chapter 4, the system of reporting incidents in the NHS relies upon staff documenting 'significant' incidents such as medication-related events, patient complaints, slips and falls, issues with medical records, delayed diagnoses and equipment problems. Staff members are required to submit their own details and those of any other staff members involved in the incident.

Completed incident forms are reviewed by immediate nursing managers and can either be dealt with at a ward level if the incident is relatively minor or will be forwarded to risk management for more serious events. There are several fundamental problems with such systems of reporting errors which are well documented within the medical literature and are outlined below.

Underreporting

One particular shortcoming of NHS reporting systems is that 'most medical errors are not reported' (Cullen *et al.*, 1995). It is generally recognised that in the current UK NHS system of error reporting, only those errors which have resulted in negative patient outcomes (judged on a spectrum of severity and impact) are commonly reported by clinical team members. While clinical risk teams might encourage staff members to report near misses (those errors which had the potential to result in an adverse patient outcome but did not), most patient safety researchers agree that this is currently not happening (Classen *et al.* , 1991; Cullen *et al.*, 1995; Leape, 2000). In a study investigating the nature and causes of error in intensive care units, Donchin *et al.* (1995) found that over a period of 4 months, whilst more than 1,000 *actual* errors were observed (by clinically trained observers), only 476 of these were formally reported. In the USA it is estimated that as few as 1.5 per cent of all adverse events and only 6 per cent of adverse drug events are reported (O'Neil *et al.*, 1993). Cullen *et al.* (1995) argue that using incident reports as a proxy measure of errors might capture only 5 per cent of actual errors occurring in health care practice. Evans *et al.* (1992) revealed a 40-fold increase in detected errors (which had resulted in a negative patient outcome) after a computerized medication ordering/transcribing/dispensing system was introduced in one hospital. This system automatically detected errors in the medication process from ordering to dispensing compared to previous reliance on self-reported medication errors. This is especially important since this increase in detected errors involved only those which had resulted in an actual outcome and not those near misses which had been averted, suggesting that even adverse incidents are under-reported.

This level of underreporting is likely to be the result of a culmination of several factors. One such explanation relates to a general lack of clarity amongst health care providers about what exactly constitutes an incident, adverse event or error and whose responsibility it is to report them. The fact that the NHS reporting system is known as *incident* reporting without a separate system of reporting errors combined with a lack of error training for health professionals is likely to lead to ambiguity as to what staff are supposed to report. In a questionnaire study by Evans *et al.,* (2006) both doctors and nurses suggested that patient falls were the most important incidents to report while drug error near misses were the least important. Staff also suggested they were frequently unsure who was responsible for reporting errors and mistakes. Evans *et al.,* (2006) found that doctor's decisions not to report errors was governed by a lack of organisational feedback on previously submitted reports, the excessive time taken to complete forms and a perception that some errors were too 'trivial' to be reported. By comparison, nurses reported an overall belief that near misses should not be reported and suggested that incident reports were often not completed if the ward was busy.

Many researchers purport that despite recent plans to improve reporting systems by government and senior NHS management bodies, there remains a palpable fear of blame and penalty amongst health care professionals reporting errors which may explain the considerable level of underreporting. In a qualitative study which explored perceptions of medication error reporting, Osbourne *et al.,* (1999) found that this fear of reprisal was particularly evident within the nursing community. In the study by Evans *et al.,* (2006) nurses were more likely to cite fear of disciplinary action as a barrier to incident reporting than doctors.

In the Department of Health's document *Building a safer NHS for patients* (2001), recommendations are made for NHS organisations to encourage staff reporting of near misses in order to improve patient safety and quality of care measures. Among other recommendations, the

DoH suggests that organisations need to '...recognise that it is weak systems that create the conditions for and the inevitability of error', and that research should '...establish agreed definitions of adverse events and near misses for the purposes of logging and reporting them' (DoH, 2001, pp. 3). Furthermore, they encourage a 'reporting culture' within the NHS organisation as a whole. However, it is clear that at present, reported errors may not represent a useful outcome measure for studies of medical error.

Assumed safety in 'closed' reporting cultures

Where underreporting of errors exists there will be a level of *assumed* safety which is largely non-representative of *actual* safety. For example, wards or departments which are more susceptible to motivational and attititudinal barriers to reporting (closed reporting culture) will appear as though they are relatively safe compared with wards less susceptible to these barriers which report a larger proportion of their errors and mistakes ('open' reporting culture). To show validity the Organisational Safety Questionnaire (OSQ) should predict medication administration errors. However, the literature described above suggests that in fact wards with the poorest levels of safety would also be those which reported the fewest errors being those with a closed reporting culture. Therefore, such wards may appear safer in terms of the number of errors they report. Conversely, wards with more open reporting cultures which report a higher proportion of their errors and mistakes could actually represent the safest wards with fewer latent problems. In support of this, Cohen *et al.,* (2005) report an intervention to improve reporting culture in a community hospital resulted in an increase in reported drug errors from 35 per 1000 patient days in 2001 to 132 per 1000 patient days in 2003. Similarly, Leape (2000) revealed a 20-fold increase in the number of reported medication errors when nurses were assured they would be immune from disciplinary action.

With these obstacles in mind, an independent measure of drug administration errors was developed with a view to testing the predictive validity of the OSQ. It was the objective of this study to develop a measure which was reflective of everyday drug administration errors in secondary care informed by medical and psychological literature on known types of error.

6.3. The psychology of human error

A vast array of cognitive failures such as slips of the tongue and pen, lapses of consciousness and perceptual and memory illusions such as false recollection have been observed and documented from as far back as the late 19th - early 20th century (Sully, 1881; Paul, 1880; Meringer & Mayer, 1895; Jastrow, 1905). Over the next 50 years or more, an equally vast range of cognitive theories were also proposed to explain these cognitive 'anomalies' such as information processing 'bottlenecks' (Broadbent, 1958; Norman, 1968) and divided attention theories (Knowles, 1963; Kahneman, 1973). Progressively more sophisticated and comprehensive cognitive models of explanation were developed towards the end of last century which attempted to incorporate both correct performance and error forms within the same conceptual framework (Newell & Simon, 1972; Norman & Shallice, 1980).

Perhaps the most notable psychological theory of human error was proposed by Rasmussen and Jensen (1974), originating from case studies of technicians in industrial installations. Their *skill-rule-knowledge* framework distinguished between three levels of information processing; skill-, rule-, and knowledge-based, which relate to the degree of conscious control an individual can exert over any given activity. When faced with a situation which is unfamiliar and novel, an individual must perform the task in a wholly conscious manner applying an analytical approach using stored knowledge. This mode of operation is known as the *knowledge-based* level of performance and requires considerable mental effort to assess, respond and monitor actions to achieve optimum performance and avoid errors. Errors made at this level are likely to occur as a result of incomplete

or inaccurate knowledge or 'resource limitations' (Reason, 1990). At the opposing end of functioning is the *Skill-based* level of performance. This behaviour occurs when individuals are highly practiced in a task which requires little conscious effort to execute and will subsequently be performed in an automated manner. Because of the automated unconsciously monitored nature of such tasks, errors at this level often occur as a result of lapses of attention or distractions.

In an extension of Rasmussen's skill-rule-knowledge framework (1983), and based on extensive evidence obtained using diary studies (Reason, 1979), Reason (1987) proposed three basic error types: skill-based slips and lapses, rule-based mistakes and knowledge-based mistakes conceptually known as the Generic Error Modelling System (GEMS). As well as incorporating Norman's slip/mistake dichotomy (Norman, 1981), GEMS also acknowledges the existence of rule violations as a distinct form of unsafe act identified as particularly important in high risk industries but previously ignored by other error theories (Lawton, 1998).

Slips, lapses, mistakes and violations

Reason (1990) describes an error as; '…a planned sequence of mental or physical activities which fails to achieve its intended outcome…when these failures cannot be attributed to the intervention of some chance agency'. His conceptual Generic Error Modelling System (GEMS) distinguished three basic error types; slips and lapses, rule-based mistakes and knowledge-based mistakes which can be fundamentally dichotomised as actions which do not go as planned (slips and lapses) and plans which are inadequate (mistakes). Reason categorises slips and lapses as common when individuals execute highly familiar routine tasks with little conscious effort at a skill-based level of performance (Rasmussen, 1983). He defines these types of errors as those which result from 'failure in the execution of an action sequence, regardless of whether or not the plan which guided them was adequate to achieve it's intended outcome'. In this way, slips can be observed as 'actions-not-as-planned' (e.g. slips of the tongue) and lapses as more *covert* failures (e.g. lapses of

memory) which may not necessarily result in observable behaviour or negative outcome by anyone other than the protagonist.

By comparison, Reason defines mistakes as 'deficiencies or failures in the judgemental and/or inferential processes involved in the selection of an objective or in the specification of the means to achieve it'. In other words, mistakes are errors which represent a failure in the formation of an accurate plan of action and the means by which to achieve this goal. Reason acknowledges that this form of error often goes undetected since individuals perceive their action sequences to be accurate in achieving their desired goal and as such do not always result in a negative outcome. As an extrapolation of Rasmussen's framework, Reason's distinguishes between *rule-based* and *knowledge-based mistakes* which operate respectively at the rule-based and knowledge-based performance levels of Rasmussen's model. Unlike slips and lapses which occur at an automated level of sensory information processing, rule-based and knowledge based mistakes are more likely to occur when individuals are involved in some conscious problem solving activity. Knowledge-based mistakes occur when individuals have inappropriate or insufficient information upon which to formulate their decision or action plan. By comparison, rule-based mistakes occur when individuals either do not know or understand the rules which apply to the current circumstance, or they misapply the rule due to lack of experience or knowledge.

It is important at this stage to emphasise that although GEMS comprised three basic error types, Reason also makes the distinction between these 'cognitive aberrations' and more deliberate violations of practices deemed by the organisation to be necessary for the safety of the system (Reason, 1993). He acknowledges that while violations are intentional deviations from practice they are less likely to involve deliberate sabotage of the system, although this can occur in rare cases. Moreover violations are likely to be intentional but 'nonmalevolent infringements' of safe practice. Reason makes the further distinction between *routine violations*, governed by a human

tendency to take the 'path of least effort' in an 'indifferent environment' which neither monitors nor punishes violations and *reasoned violations* (which can be situational, exceptional or optimising) which individuals believe they have good reason for making. Time constraints, unusual circumstances and policies which do not seem to fit the current circumstances are often cited as reasons for these types of violation. Later work by Reason, Parker and Lawton (1998) distinguished between 10 types of rule-related behaviour governed by the presence or absence and correct or incorrect application of appropriate rules or procedures.

It would be useful here to place this theory into relevant context for this research by providing health care specific examples of each error type. A slip or lapse in the medication delivery process could manifest as a physician unintentionally prescribing 5mg instead of 0.5mg, a pharmacist forgetting to add a solution to an IV fluid preparation order due to an interruption during the process or a nurse picking up the wrong IV fluid which has similar packaging to the one required. Mistakes on the other hand are notoriously difficult to detect. This is particularly true of rule-based mistakes which could be alternatively construed as deliberate violations depending on the protagonist's intention. A rule based mistake could occur if a paediatric doctor follows a treatment regime relevant for a 10 to 15 year old child when the patient is 10 years old but very small for her age, potentially leading to overdose. A doctor who prescribes and a nurse who subsequently administers Heliclear without knowing the medication contains penicillin to a patient with a known penicillin allergy are both examples of knowledge-based mistakes.

Violations from safe practice are commonplace in everyday life (e.g. driving 35mph in a 30mph speed limit area, driving through an almost red traffic light because we are late for an appointment etc.). Although violations from safe practices in health care are also likely due to the sheer proliferation of safety rules which must be adhered to, they are unlikely to be formally reported, due in part to the fact that the outcome of violations is not always negative (Lawton & Parker, 2002).

For this reason, such deviations from rules and procedures can become tacitly accepted practice over time either at an individual, team or organisational level (Reason *et al.,* 1998). For example, cross-checking patient's identity wristbands with their drug charts before administering their medication is a well-documented example of a safety rule often broken for a variety of reasons (Murphy *et al.,* 2007; Howanitz, Renner & Walsh, 2002; Skibinski *et al.,* 2007), yet would be unlikely to be reported as a safety concern unless a significant and negative patient outcome had occurred as a result.

6.3.1. Understanding medical error

In the medical field, Pani and Chariker (2004) have drawn upon Reasons GEMS model to some extent in their attempt to explain errors which occur in medicine. They present examples of cognitive functioning and malfunctioning in health care which are likely to lead to slips, lapses, mistakes and violations. For example, they suggest that many procedures in health care become so routinely executed that individuals perform them with little conscious attention, often on 'automatic pilot' which can lead to slips and lapses of attention, particularly when individuals attempt to multi-task. Pani and Chariker refer also to the unpredictable nature of medical care which can actively prevent effective memory aids commonly employed such as chunking[34]. They argue that when workload is beyond the capacity limits of working memory, important aspects of a task necessary for problem solving and decision making will be pushed out of working memory. Pani and Chariker note that sequencing tasks in times of high workload is an essential skill in health care, yet constant interruptions prevent this ability, placing further demands upon working memory. They suggest that interruptions can lead to a belief that an individual was further along in the process than they were, thus skipping components of a task, or else repeating elements they had already completed. They add that it will be infinitely more difficult for an individual to remember where they were in a process when it was interrupted if the task is familiar or routine and emphasise that

[34] A cognitive strategy for making more efficient use of short-term memory by recoding information.

the activity of medication administration is particularly susceptible to the effects of interruptions, excessive workload (which in turn leads to high levels of stress) and automatic pilot.

In a more overt attempt to utilise psychological error theory, Zhang *et al.,* (2004) propose an 'action-based cognitive taxonomy' of medical errors based predominantly upon Reason's GEMS model and Norman's seven-stage action theory (Norman, 1988). Focussing on slips and mistakes, Zhang *et al.* propose that errors can occur at any of the seven stages of action[35]. By extending Reason's definition of errors which focussed upon sequences of action, to include processes of evaluation following actions (e.g. misinterpreting feedback from medical equipment), they suggest both medical slips and mistakes can be conceptualised as either errors of execution or evaluation. However, there has been no further published evidence to support this taxonomy or to suggest it has been applied to measure errors in health care.

6.3.2. Patient safety taxonomies

Alongside research to understand the nature of human error, there have been significant attempts to classify and measure a wide variety of human errors. Many of these attempts have been to categorise errors specific to particular safety-critical industries such as aviation (Funk, 1991; Shappell & Wiegmann, 1997), nuclear energy (Smidts, Song-Hua & Mosleh, 1995) and railway networks (Gibson, Megaw, Young & Lowe, 2006). Of particular interest to this study are the multiple efforts to categorise medical errors or other aspects of patient safety. In order to develop a measure of medication administration errors, it would be useful to know the ways in which these errors have been categorised in the patient safety literature. Table 6.3.2.a provides a brief summary of four international patient safety classification systems which are most commonly utilised in a number of countries to classify medical errors.

---

[35] Establishing the goal, forming the intention, specifying the action specification, executing the action, perceiving the system state, interpreting the state, evaluating the system state with respect to goals and intentions (Norman, 1988)

Table 6.3.2.a: Patient safety and medical error classification systems employed in international health care organisations

Patient Safety Event Taxonomy Chang *et al.*, (2005)
• Developed by Chang *et al.*, (2005) on behalf of the Joint Commission on Accreditation of Healthcare Organisations based on review of 11 formal classification and reporting systems and endorsed by the US National Quality Forum as the US national standard in August 2005. • Designed to generate standardized patient safety data which can be used to understand the variables and relationship between variables in adverse events and near misses. • Theoretically based upon Reason's model of error causation and Rasmussen's theories on human-system interactions. • Proposes three types of medical error; 1. **Communication** - between patient and provider, provider and non-medical staff and among providers 2. **Patient management** - improper delegation, failure in tracking or follow-up, wrong referral or consultation 3. **Clinical performance** - the 'full range of failures that could lead to iatrogenic events during the pre-intervention, intervention, and post-intervention phases of care' (Chang *et al.*, 2005)
The Australian Incident Monitoring System (AIMS) and the Generic Reference Model (GRM) Runciman, (2002); Spigelman & Swan, (2005); Runciman *et al.*, (2006)
• AIMS developed by the Australian Patient Safety Foundation (APSF) in 1987 to monitor reported actual or potential adverse events. • GRM theoretically based on Reason's theory of error causation and Rasmussen's theory of complex system failure designed as an incident classification system to support AIMS. • GRM categorizes incidents according to generic Healthcare Incident Types ™ (HITs); 1. **Therapeutic agents** – e.g. medication and IV fluids (subdivided into administration/prescribing/dispensing/delivery/storage/reaction) 2. **Equipment and infrastructure** – e.g. medical devices; equipment and property; buildings, fittings, fixtures and surrounds 3. **Physical harm and occupational health and safety** – e.g. falls; pressure ulcers; accidents and occupational health and safety 4. **Clinical management, investigation and documentation** – e.g. clinical management; pathology/laboratory; healthcare associated infections 5. **Corporate management and security** – e.g. organization, management and services; security 6. **Behaviour, human performance and aggression** – e.g. behaviour and performance; aggression

Table 6.3.2.a: continued

The Eindhoven Classification Model and The Prevention and Recovery Information System for Monitoring and Analysis (PRISMA) Boxwala *et al*, (2004); Battles, Kaplan, Van der Schaaf & Schea, (1998); Van der Schaaf, (2005)
• PRISMA is a patient safety surveillance system developed by the Faculty of Technology Management at Eindhoven University of Technology • ECM is one component of PRISMA, originally designed for the chemical industry and expanded in 1997 to be applicable to health care. • ECM theoretically based on theories of Reason and Rasmussen. • Errors classified as **latent /technical factors** (equipment, design, construction, materials, or external factors); **organizational factors** (transfer of knowledge issues, protocols, procedures, management priorities, or cultural issues); **active errors** (knowledge-based, rule-based, and skill-based errors) and **other factors** (patient related factors or unclassified issues). • **Active errors** subdivided into 9 categories; **external** (beyond control of individual), **knowledge-based**, **qualifications** (lack of training for task), **coordination** (lack of task coordination within team), **verification** (checking problem), **intervention** (faulty task planning and execution), **monitoring** (of patient condition), **slips** (performance failure) and **trips** (physical failure).
The National Reporting and Learning System (NRLS) and the NRLS Classification NPSA (2005)
• NRLS is a voluntary reporting system introduced by the UK National Patient Safety Agency (NPSA) in 2004 to complement local reporting systems. • Intended for NHS professionals in England and Wales to anonymously report patient safety incidents, including near-misses that they are involved in or witness. • Information provided to the NPSA is analysed to identify national patterns, to identify national patient safety priorities and to develop solutions. • The customised NRLS Classification is embedded within the NRLS to categorise information into 6 nodes; **incident details** (e.g., type, location, specialty area, outcome); **patient details** (i.e., demographic and pertinent clinical information); **medication; medical devices; staff details; contributing factors.**

It is immediately apparent from Table 6.3.2.a that there is considerable variation in the methods employed by the four international classification systems to categorise incidents and errors. Furthermore, variability in defining errors and incidents within these taxonomies make it difficult, if not impossible to know what types of errors or incidents occur in health care environments. Furthermore, several of the models purport to capture 'events', 'errors' and 'incidents' yet published articles did not make it clear how and at what point in analyzing reported problems this distinction was made. For example, Van der Schaaf (2005) suggests that the Eindhoven model captures active and latent *errors*. However, this suggests that latent *conditions* are an outcome rather than the consequence. Although Reason (1990, pp. 20) acknowledges that latent conditions were originally referred to as latent errors, these types of problems should be referred to as latent *failures* since they do not necessarily involve error or failure but represent conditions under which errors are more likely to occur. Furthermore, this author could find no published data on how active errors were subcategorized by this model or any other taxonomy reviewed. In terms of the development of a measure of drug administration errors, these taxonomies provided little value.

There is however, a particularly notable taxonomy of medication administration errors which is employed widely in the US. This taxonomy was developed by the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP), formed in 1995 from 24 US national health care organisations. A primary objective of this alliance was '…to develop standardization or classification systems for the collection of medication error reports so that databases will reflect reports and grading systems' (NCC MERP, 1995). The NCC MERP group developed an index for categorising medication errors (since revised in 2001) based on nine criteria including error type, severity and patient outcomes. Of interest to this research is their category 'error type' which is subdivided into 12 types of medication error and presented in Table 6.3.2.b. Although adopted by many national organisations and health care institutions in the US, the NCC MERP taxonomy has come under criticism for its lack of empirical reliability and validity. Attempts to address this by

Forrey *et al.,* (2007) revealed 'high to near-perfect' inter-rater reliability kappa coefficients using 27 medication error scenarios across all 9 index categories. However, whether the NCC MERP taxonomy is a valid index of medication errors remains to be seen since no published data could be found which has attempted to examine its validity.

However, individual components of this taxonomy have been supported across a number of clinical areas in secondary and tertiary care hospitals. There is considerable evidence for example in the field of paediatrics;

- **Improper dose** (Raju *et al.,* 1989; Kaushal *et al.*, 2001; Paton & Wallace, 1997)
- **Wrong drug** (Selbst *et al.,* 1999; Wilson *et al.,* 1998; Ross *et al.*, 2000)
- **Wrong route of administration** (Kaushal *et al.,* 2001; Kozer *et al.,* 2002; Prot *et al.*, 2005)
- **Wrong frequency of administration** (Raju *et al.*, 1989; Vincer *et al.*, 1989; Bordun & Butt, 1992)
- **Dose omission** (Aneja *et al.,* 1992; Nixon & Dhillon, 1996; Schneider *et al.,* 1998; Herout & Erstad, 2004)
- **Wrong patient** (Paton & Wallace, 1997; Selbst *et al.,* 1999)
- **Wrong rate of IV drug administration** (Vincer *et al.,* 1989; Paton & Wallace, 1997; Blum *et al.,* 1988).

Several other studies have attempted to measure the incidence of multiple types of MAEs. For example, Bates *et al.,* (1995) carried out the Adverse Drug Events[36] Prevention Study to examine the frequency, type and preventability of medication errors associated with adverse outcomes on medical and surgical wards. All ADEs (n=247) occurring over a 6 month period in 11 inpatient units in two US hospitals, formally reported by health care professionals and recorded by independent observers and chart reviewers were examined. This study identified 8 of the 12 NCC MERP MAE categories. Analysis revealed the most common type of MAE was 'wrong dose'

[36] '...an injury resulting from a medical intervention related to a drug' (Bates *et al.,* 1995)

(classified as under- and over-dose) which accounted for 27 per cent of all medication errors. In a similar study, Malyn-Haw, Dickens & Stubbs, (2005) employed the NCC MERP taxonomy as a basis for quantifying MAEs occurring in one UK mental health hospital over a 29 month period. This study revealed support for 7 of the 12 NCC MERP MAE categories. In a US study of MAEs occurring over a 3 month period in 5 intensive care units, Calbrese *et al.,* (2005) found support for 6 of the 12 categories. A summary of results for each MAE types relative frequency for each of the three studies are presented in Table 6.3.2.b.

Table 6.3.2.b: Support for NCC MERP medication error taxonomy

NCC MERP medication error type		**No. of medication administration errors**		
		Medical/surgical (Bates *et al.,* 1995)*	**Mental health** (Malyn-haw *et al.,* (2005)†	**Intensive Care** (Calbrese *et al.,* 2001)Δ
Dose omission		10 (8%)	19 (17%)	27 (14%)
Improper dose	Over-dose	34 (27%)	35 (31%)	22 (12%)
	Under-dose			
	Extra-dose	3 (2%)		
Wrong strength		**0**	**0**	4 (2%)
Wrong drug		15 (12%)	23 (21%)	1 (<1%)
Wrong dosage form		5 (4%)	2 (2%)	**0**
Wrong technique		18 (14%)	0	**0**
Wrong route of administration		3 (2%)	2 (2%)	**0**
Wrong rate	Too fast	**0**	**0**	75 (40%)
	Too slow			
Wrong duration		**0**	**0**	**0**
Wrong time		9 (7%)	6 (5%)	26 (14%)
Wrong patient		**0**	2 (2%)	**0**
Monitoring error		1 (<1%)	**0**	**0**

*Percentages calculated from 126 total medication administration errors over 6 month period

† Percentages calculated from 112 total medication administration errors over 29 month period

ΔPercentages calculated from 187 errors measured over 3 month period

As this table shows, several of the NCC MERP MAE categories were not supported by some (e.g. wrong patient/wrong rate) or all three (e.g. wrong duration) of these studies. However, MAE categories overall were reasonably represented in three distinct fields of health care and so were considered an appropriate taxonomy on which to base the development of the drug administration error tool in study 3.

6.4. Drug Round Behaviour Questionnaire

It was necessary to consider several important criteria in the development of the DRBQ:

- **Broad spectrum of errors** – it was anticipated that by using the NCC MERP taxonomy of medication errors that this would be achieved.

- **Tangible items** – items were translated from interview and incident report excerpts which were considered directly observable and phrased as such. It was hoped that this would also ensure items were objective and not subject to interpretation.

- **Concise scale** – although a broad spectrum of medication errors was necessary to reflect nursing practice it was also necessary to ensure the scale was as short as possible to improve response. It was the intention to generate items to represent each of the 12 NCC MERP error types (see Table 6.3.2.b).

- **Not reliant upon adverse patient outcomes** - it was acknowledged that there were considerable biases within the current system of NHS incident reporting whereby only events which result in negative patient outcome are generally reported which may lead to underreporting (see section 6.2.). For this reason, care was taken to phrase items which did not refer to patient outcome which could also encapsulate near misses. It was hoped that this would foster respondents' trust, increase the likelihood they would respond honestly, therefore providing a more accurate reflection of actual errors and near misses than currently captured by formal incident reporting systems.

6.4.1. Method

Using the NCC MERP taxonomy as a template, interview and incident report data collected during study 1 (see chapter 4) were reviewed with a view to extracting examples of each type of medication administration error[37]. Participants were not directly asked about any errors they or others had made during interviews since it was not the original aim to produce a drug error tool (see appendix VIII for interview schedule). Interviewees were however invited to use their own experiences of error to illustrate any points they made and so interview data was considered an appropriate source of data.

Interview transcripts and incident reports were reviewed using the same method of translating excerpts into items described in chapter 5 (section 5.2), with a view to developing examples of each of the 12 NCC MERP medication administration error types.

6.4.2. Results

Examples of each error type were extracted (where evident) and translated into DRBQ items. Table 6.4.2.a provides examples of each error type[38], qualitative excerpts and their translated items. All excerpts provided in this table were taken from incident reports since this proved the best data source of actual drug errors. For details of qualitative data sources see chapter 4, sections 4.2.2 and 4.2.3.

---

[37] Observational data was not used in study 3 since it was not possible to identify errors during observations
[38] Not all translated items are shown since many were removed during subsequent pilot studies.

Table 6.4.2.a: Translation of raw excerpts into DRBQ items using NCC MERP categories

NCC MERP error type	Qualitative raw extract*	Translated DRBQ item
Dose omission	*'...incorrectly withheld dose of morphine slow release during am drug round'*	In the last **4 weeks**, I have withheld or *almost* withheld a drug which should have been administered
Improper dose	*'...Mrs. * given 500mgs of frusemide instead of 250mgs in error'*	In the last **4 weeks,** I have given or *almost* given a patient the wrong dose of the correct medication
	'...patient prescribed 500mgs IV erythromycin, but infused 1gram in error'	In the last **4 weeks**, I have infused or *almost* infused an *incorrect volume* of IV fluid to a patient (e.g. 75mls instead of 50mls)
	'...patient prescribed insulin on insulin chart, 81 units prescribed and nearly given. Should have been 8.1 units'	In the last **4 weeks,** I have given or *almost* given a patient a dose of a drug which was 10 times more *or* less than it should have been (e.g. 25mgs instead of 2.5mgs)
Wrong strength	Combined with 'improper dose' above. NCC MERP do not clearly distinguish the difference between 'strength' and 'dose'	
Wrong drug	*'...over 2 days the patient was given the wrong medication, 4 doses of prazocin have been given instead of pizotifen'*	In the last **4 weeks**, I have given or *almost* given a patient a different drug to the one they should have received?
Wrong dosage form	*'...gave a child of 13 years 15**mls** instead of prescribed 15**mgs** of oromorph'*	In the last **4 weeks**, I have given or *almost* given a patient the wrong unit of a drug (e.g. mcgs instead of mgs)?

*All excerpts from incident report review data

Table 6.4.2.a: continued

NCC MERP error type	Qualitative raw extract*	Translated DRBQ item
Wrong technique	No examples in data	In the last **4 weeks**, I have given or *almost* given a patient medication using the wrong technique (e.g. wrong technique when crushing tablets, omitting necessary aspects of drug preparation)
Wrong route of administration	*'...I gave prescribed dihydrocodeine IV rather than IM'* *'...mistakenly instilled ear drops into patient's eye'*	In the last **4 weeks**, I have given or *almost* given a patient medication via one route (e.g. IV) when it should have been given by another (e.g. orally)?
Wrong rate	*'...commenced IV saline with 27mmols of KCl to be infused over 8 hours but given over 1 hour'*	In the last **4 weeks**, I have infused or *almost* infused IV fluid over an incorrect time period (e.g. over 8 hours instead of 16)
Wrong duration	Combined with 'wrong rate' above. NCC MERP does not clearly distinguish between 'rate' and 'duration'.	
Wrong time	*'...accidental second dose of paracetamol given to patient in same meds round'*	In the last **4 weeks,** I have given or *almost* given a patient the same medication twice in the same medication round
	'...patient administered IC antibiotics 3 hours before it was due'	In the last **4 weeks**, I have given or *almost* given medication to a patient too soon after their last dose (e.g. only 4 hours after previous dose instead of 8 hours)
	'...drug discontinued on the 14th but had been signed as being administered from the 14th to the 19th by two nurses'	In the last **4 weeks**, I have given or *almost* given a patient a drug which I later discovered should have been withheld or discontinued

*All excerpts from incident report review data

Table 6.4.2.a: continued

NCC MERP error type	Qualitative raw extract*	Translated DRBQ item
Wrong patient	*'...I was doing a medicine round at 21:00 and I gave medication prescribed for patient x to patient y'*	In the last **4 weeks**, I have given or *almost* given a patient medication which was intended for a different patient
	'...patient given IV meropenem labelled for another patient as part of his take home medicine'	In the last **4 weeks**, I have given or *almost* given a patient someone else's medication to take home upon discharge
Monitoring error	*'...methadone liquid 1mg in 1ml found to have expired on the 24th. Patient concerned has been given 20mls on 4 different occasions by 7 members of staff'*	In the last **4 weeks**, I have given or *almost* given a patient medication which had expired
	'...patient states allergic to penicillin. Allergy band in situ but I failed to check before administering Heliclear. No adverse reaction. Monitored for 48 hours'	In the last **4 weeks,** I have given or *almost* given a patient medication without checking their allergy/ID wristband
	'...patient allergic to codeine. Documented in admission notes, allergy band in situ and documented on drug chart in red ink. cocodamol given on 29th after being prescribed on prn side chart'	In the last **4 weeks**, I have given or *almost* given a patient a drug that they were allergic to (e.g. Heliclear with penicillin allergy)

* All excerpts from incident report review data

Following test construction recommendations to develop twice the number of items required (Clark & Watson, 1995; Kline, 2002), a total of 27 items were generated from qualitative data excerpts which represented 10 of the 12 NCC MERP error types. Error types 'wrong duration' and 'wrong

rate' were collapsed as were 'wrong strength' and 'improper dose' since NCC MERP definitions made little distinction between these types and it was not always clear from the reported incidents which type of error had been involved. To account for this, multiple items were developed to represent wrong duration/rate and wrong strength/improper dose collapsed categories. The error type 'wrong technique' defined by the NCC MERP taxonomy as; '...using an inappropriate technique to administer medication...' (e.g. crushing tablets which should be given whole) was not reflected in the incident data. To overcome this, one item was developed to represent this error type, adapted from published examples on the preparation and administration of medication by the American Society of Consultant Pharmacists (ASCP, 1993). It is important to emphasise that interview data yielded very few examples of MAEs, perhaps because at the time of conducting these interviews, this was not an aim of the study and so participants were not directly asked to provide examples. Interviews may also have been less appropriate for gathering data in view of reservations about disclosure. To illustrate this point, Table 6.4.2.b indicates the frequency with which each MAE category was endorsed overall and between the two data sources.

Table 6.4.2.b: Endorsements of each NCC MERP error type by data source

NCC MERP error type	Data source		
	Incident reports	Interviews	Total
Improper dose/strength	122	3	**125***
Dose omission	84	7	**91**
Wrong time	75	1	**76***
Wrong drug	69	0	**69***
Monitoring error	55	5	**60***
Wrong patient	33	0	**33***
Wrong route of admin	28	1	**29**
Wrong rate/duration	16	0	**16**
Wrong dosage form	3	0	**3**
Wrong technique	-	-	**-**

* Error types with two or more translated items

In line with previous research (Bates *et al.,* 1995; Malyn-Haw *et al.,* 2005), Table 6.4.2.b shows that administering an improper dose to a patient was the most frequently endorsed error type followed closely by dose omission. NCC MERP error types were well represented in the incident report data. Two or more items were generated for 4 of the error types due to the volume and complexity of the data since it was unclear which of the items would *best* represent the MAE category. For example, the category 'wrong time' was translated into 3 items as the incident data suggested that a patient could be administered a medication at the wrong time in a number of ways which would not have been encapsulated within a simple 'have you administered a medication to a patient at the wrong time' question. Furthermore, the ways in which medication could be administered at an incorrect time were also likely to be affected by distinct causes and so multiple items were needed to reflect this. To illustrate this point, incident data revealed three main ways a patient could receive a medication at the wrong time;

1. Twice in the same medication round. This could be due to a lapse in attention due to interruptions or poor communication.
2. Administering medication too soon after the last dose. This could be due to an administrative error (e.g. illegible handwriting) or poor communication (e.g. believing a previous dose had not been administered).
3. Administering withheld drugs. This which could be due to poor design of documentation, illegible handwriting, a lapse in concentration due to interruptions or poor communication.

6.5. Pilot study

This study was conducted to assess the face validity of developed DRBQ items with the ultimate aim of removing items which nurses thought were rare or had poor validity. It was the aim of this study to reduce the overall questionnaire length of the DRBQ to approximately 10-15 items, based on recommendations to increase response rates (Edwards *et al.,* 2002).

6.5.1. Method

Participants and recruitment

Thirty senior nurses who worked on four wards which comprised surgical, paediatrics and a large medical admissions unit (MAU) from the collaborating hospital were invited by letter to participate in this pilot study. Since the final study would involve hospital-wide dissemination of the final questionnaire to this hospital it was important that the same participants were not repeatedly asked to complete questionnaires to reduce response fatigue. Participants were therefore advised they should not have already participated in previous pilot studies associated with this research. As with pilot studies conducted for study 2, participants were advised that this study was testing the workability of a new questionnaire which had been developed by researchers at the University of Leeds to measure the frequency of 'clinical events and near misses associated with medication administration' which did not harm the patient. It was hoped that using non-judgemental language framed as near misses would encourage uptake. Participants were advised their participation would involve completing a short questionnaire (approximately 5-10 minutes) and a short task (approximately 10 minutes). As with the pilot for study 2, participants were offered paid lunch and a £10 incentive to participate. Eighteen nurses agreed and consisted of three sisters, eleven senior staff nurses and four junior staff nurses from the MAU and surgical wards.

Design

<u>DRBQ</u>

The DRBQ was presented in a folded booklet format in a similar way to the OSQ in study 2 (see chapter 5, section 5.4). Both university and collaborating hospital logos were printed on the front page which comprised a participant information sheet (appendix XIII) followed by DRBQ items in the booklet centre. Only the final DRBQ items are presented in this chapter (see Table 6.5.2.2). As with the OSQ, participants were instructed to respond to each question on a 5-point Likert scale from 'not at all' (1) to 'nearly all the time' (5) with an additional 'not applicable' option. All items

were phrased as 'given or *almost* given' medication, to incorporate near misses. Participants were assured their responses were completely anonymous and they were under no obligation to respond to any questions they felt uncomfortable answering.

Face validity task

This task was designed to measure the face validity of the DRBQ. A booklet format was used for participants to complete this task (appendix XIV) which instructed them to read each of the 27 DRBQ items and respond to *all questions* as either yes or no in terms of;

1. **Comprehension** - did the question make sense? Participants asked to provide further comment if the answer was no.
2. **Frequency** – were participants able to judge the item by its frequency of occurrence? (i.e. is the problem particularly common or rare). It was anticipated that even if respondents had said the drug error had not happened to them in the last 4 weeks it may have been that 4 weeks was an inappropriate time scale for a particular item.
3. **Examples** - were respondents able to think of any examples of when such an event might have occurred either to them *or to someone they know*? It was anticipated that even if participants had said an item had not occurred to them in the last 4 weeks they may have been likely to admit they knew it had happened to someone else. Items which respondents said they could not think of a single example could therefore be judged as an anomaly and removed.

At the end of this task sheet, participants were offered the opportunity to make any other comments regarding the DRBQ or describe any other types of drug administration errors they thought were significant but which had not been included.

Procedure

Ten minute meetings were scheduled with recruited participants to describe the nature of the research and the purpose for completing the validity task (since participants had not been involved

with previous stages of the research). Meetings were also intended to encourage and improve the likelihood that responses would not be affected by fear of reprisal. Due to the particularly sensitive nature of this pilot study, it was emphasised that the DRBQ was not an audit tool but a simple assessment of the validity of a general measure of drug round behaviour and that data would not be fed back to the organisation. Participants were assured their responses were completely anonymous and confidential available only to the principal researcher at the University of Leeds. While participants were given the opportunity to withdraw from the study at any time, no individual chose to do so. Participants were given 'task packs' which contained a participant information sheet, a copy of the DRBQ, the face validity task and a freepost envelope with which to return completed documents anonymously to the principal researcher. All 18 participants returned completed packs.

6.5.2. Results

Results will be discussed separately in terms of the face validity and also response variability.

6.5.2.1. Face validity

This task was designed to measure the face validity of the DRBQ in terms of comprehension, likelihood of occurrence and ability of participants to think of examples. Overall, participants responded that they had no problems with any of the 27 items in terms of thinking of examples or estimating the frequency each event occurred. This indicated that all items were relevant, current and frequent examples of drug administration errors. Several participants indicated they did not understand one or two of the items. However, no participant indicated the same item as being problematic or provided further comment on why the question was not understandable or how it could be changed and so no items were removed at this stage. Lack of negative response on this task could indicate that there were no problems with any of the items. However, there was some degree of missing data and acquiescent responding (e.g. broadly circling the yes response for all

items down the page rather than circling yes for each item) which suggested that participants may not have taken the task as seriously as was necessary. It is possible that asking participants to complete this task face-to-face would have improved this type of responding. However, when invited to make any further comment about the questionnaire, most respondents made comments which were generally positive. In view of the lack of useful data yielded by this task and a need to remove redundant items, a meeting was set up with two of the matrons originally recruited for study 1 (see chapter 4) but who had not participated in this pilot study, with a view to discussing any items which they thought had poor face validity. Matrons said they had no problems with the face validity of any item but suggested that several items (n=12) appeared to be asking the same thing. Since both matrons independent of one another suggested changing or removing the same items, these 12 items were reduced to 6 by collapsing items or removing items which were virtually identical to another; leaving 21 items.

6.5.2.2. Response variability

Since it was an aim of this study to produce a measure which was approximately 10-15 items in length, variability in responding was analysed with a view to removing items which yielded reduced variability. Results revealed 8 of the 21 items had floor effects and were consistently answered 'not at all' by all 18 respondents (these errors had not happened in the last 4 weeks) and so were removed. It was possible that the time frame was inappropriate for these items. However, time constraints for this study meant that it was not possible to re-examine the time frames further (i.e. it was not known what would have been more appropriate time frames for these items) and conduct another pilot study before study 4. Also, the main purpose of developing twice as many items as the number needed was to allow for the removal of items such as these. Table 6.5.2.2. shows responding frequencies for the remaining 13 items.

The majority of items showed a reasonable distribution of responses considering the sensitive nature of the questions, with very few 'not applicable' responses. However, data did indicate that there remained a positively skewed distribution in favour of 'not at all' responses for most items. Of particular interest are items 4 and 11 to which the majority of participants responded 'not at all'. Given the possible consequences both to patients and to nurses committing errors reflected in these particular items (e.g. administering 10x the prescribed dose and administering medication to which a patient is allergic), this skewed response was to be expected. It is likely that these errors were considered by respondents as potentially more serious. This skewed response could also indicate respondents were not convinced their information would be treated confidentially. In this case, respondents may have attenuated the actual frequency of these behaviours. Alternatively the response bias may represent an accurate frequency of this behaviour. Since the final validation study (chapter 7) would administer these items to approximately 200+ participants, it was anticipated that this analysis would enable the removal of any further redundant items and so it was therefore decided that all 13 items would be retained.

Table 6.5.2.2: Pilot study DRBQ response frequency

#	DRBQ item	Error type	Response (n=18)*					
			NAA	OCC	QO	FR	NAT	NA
1	In the last **4 weeks,** I have given or *almost* given a patient medication I was not sure had been prescribed correctly	**Wrong drug**	9	4	2	1	1	1
2	In the last **4 weeks,** I have given or *almost* given a patient the wrong dose of the correct medication	**Wrong dose**	4	7	2	2	3	0
3	In the last **4 weeks**, I have infused or *almost* infused an *incorrect volume* of IV fluid to a patient (e.g. 75mls instead of 50mls)	**Wrong dose**	7	4	5	1	1	0
4	In the last **4 weeks,** I have given or *almost* given a patient a dose of a drug which was 10 times more *or* less than it should have been (e.g. 25mgs instead of 2.5mgs)	**Wrong dose**	15	3	0	0	0	0
5	In the last **4 weeks**, I have given or *almost* given a patient someone else's medication to take home upon discharge	**Wrong patient**	6	3	3	3	2	1
6	In the last **4 weeks**, I have given or *almost* given a patient medication which was intended for a different patient	**Wrong patient**	10	3	2	2	1	0
7	In the last **4 weeks,** I have given or *almost* given a patient the same medication twice in the same medication round	**Wrong time**	3	7	5	2	1	0
8	In the last **4 weeks**, I have given or *almost* given medication to a patient too soon after their last dose (e.g. only 4 hours after previous dose instead of 8 hours)	**Wrong time**	7	3	4	2	2	0
9	In the last **4 weeks**, I have given or *almost* given a patient a drug which I later discovered should have been withheld or discontinued	**Wrong time**	1	2	8	3	4	0
10	In the last **4 weeks**, I have given or *almost* given a patient medication which had expired	**Monitoring error**	9	3	3	2	1	0
11	In the last **4 weeks**, I have given or *almost* given a patient a drug that they were allergic to (e.g. Heliclear with penicillin allergy)	**Monitoring error**	16	2	0	0	0	0
12	In the last **4 weeks**, I have infused or *almost* infused IV fluid over an incorrect time period (e.g. over 8 hours instead of 16)	**Wrong rate/duration**	3	4	6	2	1	2
13	In the last **4 weeks**, I have withheld or *almost* withheld a drug which should have been administered	**Dose omission**	2	3	8	3	2	0

*NAA – Not at all; OCC – Occasionally; QO – Quite often; FR – Frequently; NAT – Nearly all the time; NA – Not applicable

6.6. Conclusions

This study aimed to develop a measure of medication administration errors which was broad yet concise and not reliant upon adverse patient outcome with which to test the predictive validity of the Organisational Safety Questionnaire (OSQ). The 27-item Drug Round Behaviour Questionnaire (DRBQ) was developed using the NCC MERP taxonomy of 10 medication error types as a basis for item generation from qualitative data obtained in study 1. A pilot study and subsequent analysis of responding variability was conducted to measure the face validity of the measure and response patterns to items, resulting in the removal of 14 items. The remaining 13 which had good face validity and a reasonable response distribution items measure 7 types of drug administration error; wrong drug, wrong dose, wrong patient, wrong time, dose omission, wrong rate/duration and monitoring error. The 13-item DRBQ will be embedded within the 82 item OSQ for the final validation study (study 4 described in chapter 7). Further analysis on the scale reliability of the DRBQ and its concurrent validity with formally reported medication errors will be reported in chapter 7.

CHAPTER 7
VALIDATION OF THE ORGANISATIONAL SAFETY QUESTIONNAIRE

7.1. Introduction

Study 4 was the final stage of this research and was conducted to measure the validity and reliability of the Organisational Safety Questionnaire (OSQ) according to the following criteria;

- **Internal reliability** – whether the overall OSQ scale and 10 component scales reveal good levels of internal consistency of responding.
- **Concurrent validity** – whether responses on the OSQ are correlated with a previously validated measure of organisational safety.
- **Criterion validity** – whether scores on the OSQ predict scores on an independent criterion measure (e.g. measure of drug administration errors)

7.2. Method

Materials

The OSQ comprised 82 items which were considered to reflect a broad range of latent conditions (see chapter 4 for descriptions). Items were arranged in random sequence with no definition or reference to any of the 10 latent variables reflected in the items (as there was with the pilot studies – see chapter 5). Drug Round Behaviour Questionnaire (DRBQ) items were randomly placed amongst OSQ items to attenuate the risk of participants responding in a socially desirable or consistent manner. It was expected that had DRBQ items been presented separately participants may have anticipated the purpose of the study and responded accordingly. Research suggests embedding items which comprise an instrument intended to support the predictive validity of another instrument (as opposed to administering the measures separately) will reduce *criterion*

contamination and lessen the chance that participants are aware of the expected outcome of the measures (Clark-Canter, 2004). Three additional validated items were embedded within the final questionnaire from the *Hospital Survey on Patient Safety Culture* (Sorra & Nieva, 2004) regarding the likelihood of reporting errors. Since research has indicated errors are generally underreported (see chapter 6, section 6.2), it was hypothesised that these items would provide useful information on the formal reporting habits of nurses in general and highlight any differences between 'safe' and 'unsafe' wards (i.e. those with high and low OSQ scores respectively). These items were considered to capture the frequency of reporting adverse events in a way not represented by other OSQ items;

1. When a mistake is made that *could* harm the patient, but does not, how often is this reported on your ward/unit/dept?
2. When a mistake is made, but is *caught and corrected before affecting the patient*, how often is this reported on your ward/unit/dept?
3. When a mistake is made, but has *no potential to harm the patient*, how often is this reported on your ward/unit/dept?

Participants were instructed to respond to each question from 1 (not at all) to 5 (nearly all the time) or answer 'not applicable' if they felt the question was not relevant in their role. The questionnaire was printed in a booklet format with a participant information sheet on the front page which gave a brief description of the purpose of the research, instructions on how to complete the questionnaire and assurances of confidentiality and anonymity. Participants were also given the contact details of the principal investigator in case they wished to find out more about the study before participating. OSQ/DRBQ items were presented in the middle pages of the booklet with demographic questions detailed on the back page. Demographic items were taken from the *Hospital Survey on Patient Safety Culture* (Sorra & Nieva, 2004) and asked participants for details such as length of nursing experience, number of hours worked per week and staff position (see appendix XII for full questionnaire). One further item was taken from this survey which asked respondents to grade their

ward/unit on patient safety (from failing to excellent). Participants were also given a large space in which to write any additional comments (Dillman, 2000).

Several recruitment strategies highlighted in a recent systematic review of methods to improve response rates to postal questionnaires (Edwards *et al.,* 2002) were employed. Edwards *et al.* reviewed 75 strategies for improving postal questionnaire uptake from 292 randomised controlled trials. Those strategies which were significantly successful in improving response rates and appropriate for this study are described here;

- **Short questionnaires** – Attempts had been made to keep the questionnaire as short as possible and answerable within 15 minutes allowing for the number of variables being measured (see chapter 5). The questionnaire was also printed in a double sided booklet to make it look shorter.
- **Personalised questionnaires and letters** – Access was given to a Human Resources database of all the names of every qualified and unqualified nurse working in the collaborating hospital. One questionnaire was sent to every nurse with a letter of introduction from the principal researcher addressed to each nurse by first name (see appendix XV). Each cover letter was signed by hand (Dodd & Markwiese, 1987).
- **Monetary incentive** – The collaborating hospital expressed concern that providing monetary incentive would discourage participants from participating in future in-house non-incentivised research such as clinical audits, and so participation in this stage of the research was not incentivised.
- **Use of coloured ink** – All questionnaires and cover letters were printed in coloured ink.
- **Enclosing stamped return envelopes** – Participants were provided with a freepost envelope addressed to the principal investigator at the University of Leeds with which to return completed questionnaires. It was hoped that this would also encourage participants to feel confident their individual responses would be treated confidentially which would in turn improve response rates.
- **Contacting participants before sending questionnaires** – A short email introducing the research project and advising nurses that they would be sent a questionnaire in the following 2 weeks was written by the principal investigator and forwarded to all nurses by the Chief Nurse who also added her endorsement of the project.

- **Providing non-respondents with second copy of the questionnaire** – Questionnaires were administered twice in 3 months. It was not possible to follow up non-respondents since all questionnaires were anonymous and so questionnaires were sent to all qualified nurses twice.
- **Questionnaires originating from universities more likely to be returned than from other sources** (e.g. such as commercial organisations) – All questionnaires and cover letters bore the University of Leeds logo and reference to the principal investigator being employed by the University of Leeds. The collaborating hospitals logo was also provided on this paperwork.

In order to measure the concurrent validity of the OSQ, 9 items from a recent validated behavioural measure of localised safety culture (*The Safety Organising Scale*; Vogus & Sutcliffe, 2007) were also added to the questionnaire. Because the response scale for these items differed from the OSQ, these items were presented separately from OSQ items. The 9-item Safety Organising Scale is a unidimensional behavioural measure of safety culture in secondary care nursing units with high internal reliability and good criterion validity with medication errors and other patient safety outcome measures. Participants are required to respond to 9 positively framed items on a Likert-type scale from 1 (not at all) to 7 (to a very great extent) (see appendix XII for items).

Participants

Although the OSQ was developed from information provided by qualified nurses working predominantly in the field of general medicine and elderly care, it was an objective of this stage of the research to determine to what extent the tool was generalisable beyond this sample. It was decided that all grades of qualified and unqualified (e.g. nursing assistants, student nurses, health care assistants) nursing staff working in two hospitals at the collaborating NHS Trust would be sampled (n=2010). The number and type of nursing staff who responded to the questionnaire will be presented in section 7.3.1 with a discussion of the generaliseability of the tool.

Procedure

An email was forwarded to all potential participants which briefly described the purpose of the research (to explore aspects of the ward and hospital environment which make errors more likely) and encouraged nurses to complete the OSQ. All participants were then posted the OSQ with freepost envelopes with which to return completed questionnaires to the University of Leeds. Each questionnaire was sent with a personalised hand signed letter to each participant. All questionnaires were batch coded in order to determine the unit or ward each respondent was from but not the identification of individual participants (to maintain anonymity). Coding questionnaires in this way resulted in a clearer picture of the departments for whom the questionnaire was and was not applicable. Participants were asked to complete and return questionnaires within four weeks of receipt. A further email was sent two weeks after receipt to encourage nurses to complete their questionnaires, endorse by the Chief Nurse. Three weeks after the initial deadline had passed, the OSQ was resent to all participants (since questionnaires were anonymous and so non-responders were not known) to improve response rates.

7.3. Results

The OSQ was administered in two recruitment rounds to qualified and unqualified nurses (n=2010) from two Bradford hospitals in a three month period. All collected data was entered into the statistical package SPSS and analysed for generaliseability, validity and reliability. Results are presented in this section under each component heading.

7.3.1. OSQ generaliseability

After administering the questionnaire to 2010 nurses, 642 questionnaires were returned which represents an overall response rate of 32 per cent. After initial examination of completed questionnaires it became evident that a considerable number of OSQ items were largely not

applicable (NA) for many respondents (as indicated by a high degree of NA responding). Since it was a broad objective of developing the OSQ that items should be generalisable to as many different types of nursing staff in as many specialities of care as possible (to obtain the most holistic picture of ward and hospital activity) analysis was carried out to examine the type of nurses and departments for which the OSQ was most appropriate. It was possible that NA responding was entirely random; certain individuals in a variety of departments simply struggled to answer some questions (particularly since the items were developed with qualified nursing staff in general medicine only). It was also considered a possibility that particular grades of nurse (e.g. unqualified) could be skewing these responses particularly since the questionnaire was developed in collaboration with qualified nurses only. In order to examine this, questionnaire data was separated by self reported nursing grade and analysed for NA responding to OSQ items for each group. Table 7.3.1.a summarises these findings. It should be noted that all data reported in this section relates to *OSQ items only* (n=82), since DRBQ items (n=13) were included in this study solely to test the criterion validity of the OSQ. These findings will be discussed separately in section 7.3.5.

Table 7.3.1.a: Not applicable responding by nursing grade (n=642)

Nursing grade*	N	Mean % NA (Std)	Median % NA	Min-Max % NA
Matron	9	9.93 (13.62)	3.53	0 – 42.35
Sister	112	20.04 (19.73)	13.53	0 – 76.47
Senior staff nurse	141	14.57 (18.02)	5.88	0 – 71.76
Staff nurse	196	16.72 (19.93)	8.24	0 – 84.71
Student nurse	8	24.56 (21.97)	19.41	0 – 57.65
Midwife	42	26.44 (20.51)	20.00	0 – 71.76
Health Care Assistant†	83	**45.24 (17.42)**	**43.53**	**17.75 – 98.82**
Other†	51	**40.55 (24.86)**	**37.65**	**4.71 – 95.29**

*Sorted in order of seniority
† Nursing grades with worst levels of NA responses which were removed after analysis

Not applicable responses were converted into percentages of OSQ responses overall. This was done to take missing data into account which for some participants was considerable. For example, if a participant had responded to only 60 of the 82 OSQ items, leaving 15 items blank their NA response rate was calculated as a percentage of the items they responded to (60) rather than all 82 items. Mean and median % NA scores by type of nursing grade were calculated and findings revealed that OSQ items were generally not applicable to health care assistants (unqualified) and respondents describing themselves in the 'other' category (e.g. nurse researchers, theatre technicians etc.). This was generally anticipated since these types of nurses were not involved in any stage of developing the OSQ. Additionally, OSQ items were those thought to cause errors in the administration of medication which is an activity neither group would be involved in. Therefore these individuals would be less familiar with the likelihood and frequency of these indicators of safety failure. Responses from these participants were excluded from further analysis (n=134). A further 13 cases were removed due to missing data levels which exceeded more than 35 per cent of respondents data set for OSQ items.

Ranges and standard deviations for the remaining 6 nursing groups indicated there was still a high degree of within-group NA responding and median scores suggested that OSQ items were still generally not applicable to some individuals. It was possible that the questionnaire was not relevant for particular specialties and so NA responses were examined by ward/department. Table 7.3.1.b shows the wards/departments for whom the OSQ was *least* applicable based on median NA responding and respective ranges. Data from these departments (n=14) were removed from further analysis.

Table 7.3.1.b: OSQ least applicable wards/departments

Department	N	Mean % NA (Std)	Median % NA*	Min-Max
Theatres – Modular	11	35.08 (8.36)	36.47	23.53 – 51.76
Antenatal	5	42.82 (5.56)	42.35	34.12 – 48.24
Theatres – ENT	2	42.35 (14.98)	42.35	31.76 – 52.94
ENT Outpatients	12	44.71 (12.90)	45.89	21.18 – 65.88
Theatres – general	2	46.48 (2.50)	46.48	44.71 – 48.24
Radiology	5	46.82 (10.17)	48.24	34.12 – 56.47
Orthopaedic Outpatients	6	52.94 (11.21)	50.00	42.35 – 71.76
Diabetology	3	34.51 (29.88)	51.76	0 – 51.76
Dermatology	6	56.86 (18.57)	54.71	32.94 – 84.71
Community midwifery	15	43.62 (24.74)	55.29	3.53 – 71.76
Pre-assessment Ward	7	60.00 (8.56)	55.29	52.94 – 72.94
Dept of infect. disease & sex. health	5	60.24 (12.21)	57.65	44.71 – 75.29
Paediatric OPD	3	54.90 (10.87)	61.18	42.35 – 61.18
Pain therapy	2	73.53 (4.16)	73.53	70.59 – 76.47

It was important to distinguish between departments with outlier scores affecting the mean and range and those units for whom the OSQ items were generally inapplicable. Therefore mean scores, standard deviations, median and score ranges were obtained. Results revealed departments with the highest mean and median % NA responding levels and relatively small standard deviations were indicative of those for whom OSQ items were generally not applicable. These departments (n=14) were removed from further analysis as it was clear that for these departments the OSQ was not relevant. Additionally, data from these cases (n=128) would not be useful for subsequent factor analysis. Even though one aim of factor analysis would be to remove redundant items (thus reducing the overall % of NA responses), it was felt that departments with a mean NA response rate of 25 per cent or more with little within-group variance were unsuitable cases. Notably, departments for whom the OSQ was not relevant were generally outpatients departments, day units, non-ward based units (e.g. community midwifery) or highly specialised areas which were very

distinct from the sample of wards the OSQ was developed with (e.g. department of infectious diseases and sexual health). Therefore it is unsurprising that these wards had such high rates of NA responding. A total of 128 cases across 20 departments were excluded from further analysis. An additional 39 cases were also removed with more than 25 per cent NA response. It was considered that even with the removal of items during factor analysis, the OSQ was generally not applicable to these individuals who came from a variety of departments and nursing groups. The data set now comprised 328 cases which represented 6 types of qualified nurse working in 34 clinical areas. Sample details will be presented in section 7.3.3. It should be acknowledged here that although preliminary screening of data in this way is not generally reported to this degree, it was an important objective of developing the OSQ that the tool was generalisable beyond the clinical speciality in which it was developed. Therefore, reporting details on the types of participants and multiple clinical areas for which the OSQ is applicable highlights that this objective was generally met and is invaluable information for future use of the tool.

Item variability

Responses to OSQ items were also examined to eliminate items which were redundant or exhibited no variability.

Redundant items

After removal of nursing groups and clinical specialties for which the OSQ was *generally* not applicable, some individual OSQ items still had high levels of NA responding. In order to make the OSQ as generalisable as possible to the remaining sample, NA responses for each item were calculated. The following items were deemed to have high levels of NA responding (>30 per cent of the remaining 'relevant' sample of 328 considered the item NA). Table 7.3.1.c shows items removed from further analysis on this basis sorted in descending order of NA response levels.

Table 7.3.1.c: Items with >30% NA responding (n=7)

Latent variable	Item	No. of NA responses (% of sample)
Bed management[3]	In the last **3 months**, this ward has had to rush an existing patient's discharge to admit a new patient from MAU or A&E to avoid a patient breaching the 4 hour rule in A&E	120 (37%)
Workload[1]	In the last **4 weeks**, the number of doctors arriving for morning ward round has slowed my drug round progress	119 (36%)
Workload[1]	In the last **4 weeks**, I have sped up during a drug round to avoid a backlog of jobs later in the shift	115 (35%)
Workload[1]	In the last **4 weeks**, I have sped up during a drug round to avoid making the next drug round late	108 (33%)
Bed Management[3]	In the last **3 months**, this ward has had an 'inappropriate admission' because A&E did not want to breach the 4-hour waiting rule	105 (32%)
Workload[2]	In the last **2 months**, I have been unable to coordinate the ward/unit/dept because of my patient allocation	102 (31%)
Ward climate[1]	In the last **4 weeks**, I have had to break off from a drug round because I was interrupted by another health professional (e.g. doctor, nurse, physio)	101 (31%)

It is unclear why these items in particular had such high levels of NA responding although several reasons are postulated for each item;

1. These items relate *specifically* to conducting a drug round. Although the nurses and departments remaining in the sample were considered to be those which routinely engage in drug rounds, interviews with nurses during study 1 revealed that not all qualified nurses carry out drug rounds on a regular basis.
2. This item relates to coordination of a ward, an activity which may be restricted to only the most senior nurses (e.g. matrons and sisters only) rather than *all* qualified nurses.
3. These items related to a specific safety concern which involves rushing existing patient's discharges and accepting inappropriate admissions from the Medical Admissions Unit due to

the A&E 'breach-rule'. It is possible that this is something which largely happens only in general medicine (where the item originated from) due to the nature of this clinical specialty.

NA response data was recalculated for each individual after removal of the above items. A final 17 cases were removed from further analysis due to having more than 25 per cent NA responses; leaving 311 cases and 75 items.

Poor variability

Although participants were given a 5-point Likert response scale from 1 (not at all) to 5 (nearly all the time), the overall variability in responding was very low and both histograms and K-S values (Kolmogorov-Smirnov) revealed every remaining OSQ item (n=75) was significantly positively skewed. Comparable evidence of skewed item responses when administering a similar health care measure of teamwork and safety climate was found by Hutchinson *et al.,* (2006). This study revealed weak discrimination between 'agree slightly' and 'agree strongly' responses.

To rectify this skew, a number of different transformations were carried out on the dataset including reciprocal, logistical and square root transformations in an attempt to attenuate the skewness of data and secure more normal distributions before factor analysis was conducted. However, distributions were so skewed that transforming the data in this way had little effect, and in some cases made the skew worse. It was hoped that collapsing the five response categories into three response groups (not at all; occasionally/quite often; frequently/nearly all the time) would reduce this skew. Unfortunately this was not the case and this method did little to improve response distribution. It was also considered that by transforming the data from a scale response to a dichotomous response (yes / no) the problem with skewed responding would disappear altogether (collapsing all 2-5 responses into 'yes' and leaving response 1 – currently 'not at all' as a 'no'). However, it was an

important objective of this study to measure the underlying structure of the OSQ which would ultimately involve factor analysis. Since factor analysis is not recommended for dichotomous data (Kim & Mueller, 1985; Streiner, 1994; Polit, 1996) because this type of data influences the correlation between constructs and potentially introduces significant amounts of measurement error, response categories were not collapsed in this way. Therefore the decision was made to retain the data in its current 5 point response scale and remove items which showed the worst variability (hence those which were most positively skewed and had the highest K-S values) in responding. The items which had the highest K-S values and lowest mean scores (n=12) were removed from further analysis and are reported in Table 7.3.1.d.

Although the remaining 63 items were significantly positively skewed (as indicated by statistically significant K-S values), histograms indicated these distributions were not as problematic as K-S values would suggest and so the remaining items were retained for further analysis[39]. Although it is acknowledged that it is not ideal to conduct factor analysis on data which is not normally distributed, it is widely accepted that questionnaire items which are particularly sensitive will inevitably yield responses which are skewed to some extent (Tourangeau & Smith, 1996; Giles & Field, 2006).

[39] It should also be noted that it is well documented that with large samples, even the slightest skew will result in significant high skewness statistics and large K-S values. Under these circumstances researchers are urged to use histograms with corresponding normality curves to observe skew (Field, 2005).

Table 7.3.1.d: Items with lowest mean scores and highest K-S values

LATENT VARIABLE	ITEM	% VALID RESPONSES					Mean response (Std)	K-S value†
		NAA	OCC	QO	FRE	NAT		
Local working conditions	In the last **2 months**, I have encountered boxed medication which did not correspond with the outer package labelling	80	16	3	<1	<1	1.24 (.54)	.47
Training	In the last **6 months**, I have been taught a clinical procedure or how to use a piece of equipment by a nurse who had not received the relevant 'formal' training	80	12	6	1	1	1.31 (.72)	.47
Training	In the last **12 months**, I have mentored a newly qualified nurse *before* I have attended mentorship and teaching and assessment programs	82	9	3	1	5	1.36 (.94)	.47
Supervision	In the last **2 months**, a senior nurse has advised me to cut corners in order to complete all jobs planned before the end of the shift	88	8	2	1	1	1.20 (.61)	.50
Ward climate	In the last **2 months**, I have refrained from bleeping a particular doctor again because I knew they would be annoyed I had bleeped them several times that shift	79	15	5	<1	<1	1.31 (.68)	.46
Ward climate	In the last **3 months**, I have made or witnessed an error which I did not 'formally' report because the patient wasn't harmed by it in any way	82	15	2	0	<1	1.22 (.53)	.48
Supervision	In the last **2 months**, I have been confused as to what jobs were supposed to be my responsibility during the course of a shift	80	15	3	2	<1	1.29 (.67)	.47
Local Working conditions	In the last **3 months**, I have encountered medication (e.g. boxed or IV) in the drug trolley or clinic room which had expired by more than a month	72	26	1	1	<1	1.34 (.60)	.43
Ward climate	In the last **3 months**, I have experienced difficulty challenging a doctor about a prescription I thought they had written incorrectly	76	20	3	<1	<1	1.31 (.65)	.44
Written Policies and Procedures	In the last **12 months**, I have signed to say I have read and understood a new policy, procedure or guideline when I haven't either read it or understood it	72	18	4	3	3	1.48 (.95)	.41
Bed Management	In the last **3 months**, a patient has been discharged from this ward before all the necessary agencies needed to care for them at home (e.g. community nurse, social worker) had been arranged	76	18	4	1	<1	1.31 (.64)	.45
Training	In the last **6 months**, I have booked a training course or study day which was cancelled due to budgetary constraints	73	17	5	4	1	1.44 (.87)	.43

† All Values significant at level of $p<.001$

In summary, preliminary screening of the data led to the removal of 331 cases and 7 items due to high levels of 'not applicable' responses. A further 12 items were removed due to reduced variability in responses. All analyses reported in subsequent sections are therefore based on a dataset of 311 participant's responses (from 34 clinical areas) to 63 OSQ items.

7.3.2. Underlying structure of the OSQ & internal reliability

In order to determine whether items were representative of their relative latent variables and whether these variables reflected one concept of 'organisational safety', factor analysis was carried out. It was anticipated that factor analysis would support a 10 factor solution. Since the test construction was still deemed to be at an exploratory stage of development, it was decided that exploratory factor analysis (EFA) would constitute the most appropriate method of analysing the data[40]. Although EFA is constrained by assumptions of normality, Costello and Osbourne (2005) recommend that in cases where data 'severely violates' assumptions of normality (as was the case in this study) principal axis factoring is a more appropriate method of analyzing data than either confirmatory factor analysis or the most commonly employed principal components analysis (PCA)[41].

Initially, the factorability of the 63 OSQ items was examined. The Kaiser-Meyer-Olkin measure of sampling adequacy was .85, well above the recommended value of .6 (Hutcheson & Sofroniou, 1999, pp. 224-5), and Bartlett's test of sphericity was also significant (χ^2 (1953) = 5526.65, $p <$ .001) (Field, 2005, pp. 640). With only 5 exceptions (items TC1, TC9, LWC1, WPP6 and TR1 were removed from further analysis; see table 5.5.b for item codings), communalities were all

---

[40] Had the sample been larger, it would have been useful to randomly divide the sample in two and conduct EFA on one half and CFA on the other (Hutchinson *et al.*, 2006).

[41] They suggest that unlike principal components analysis, principal axis factoring discriminates and 'partitions' the shared variance of each variable from both its unique variance and error variance to reveal an underlying factor structure and therefore not as susceptible to the effects of non-normally distributed data. They claim that PCA should be used where the main purpose is data reduction and not exploring any underlying structure of the data caused by latent variables which was the aim of this study.

above .3, further confirming that each item shared some common variance with other items (Field, 2005, pp. 640). Given these overall indicators, principle axis factoring was conducted with the remaining 58 items.

Principle axis factoring revealed a 17-factor solution (eigenvalues >1) which accounted for 48 per cent of the overall variance in scores. Factor 1 explained 20 per cent of this variance and comprised 40 items. Factors 2-17 accounted for between 3 and less than 1 per cent of the variance each but were generally uninterpretable and predominantly freestanding one item factors. Hutchinson *et al.*, (2006) suggest this can occur when trying to measure more than one underlying factor and recommends factor analyzing each proposed variable. In their study, teamwork and safety climate items were factor analysed respectively to reveal two- and three factor solutions. Unfortunately, in this study high levels of not applicable responding had led to the removal of a considerable number of items, which meant there were too few items remaining in some of the latent variables (e.g. bed management, workload etc.) to conduct separate factor analyses.

To avoid over-extraction of factors, Costello and Osbourne (2005) recommend examination of the scree plot for the natural bend in the data. This analysis suggested a one-factor solution. Since only one factor was extracted, rotation was not applied to subsequent analyses. Using an iterative process to remove items loading on factors 2-17, cross-loadings and then any items loading <.4 on factor 1 (recommended when participant to item ratio is <10:1; Field, 2005), the best solution was obtained from 28 items with factor loadings ranging from .41 to .70 (see appendix XVI for remaining items and individual item loadings). This factor reflected one construct named organisational safety which, due to the large range of items subsumed in this construct with good factor loadings, was considered to have good content validity (Clark-Canter, 2004). After redundant

items had been removed[42], the overall one factor solution accounted for 26 per cent of the variance in scores. This scale revealed high internal consistency (Cronbach`s alpha .90). No substantial increases in alpha could have been achieved by eliminating further items. Although factor analysis had not extracted the 10 latent variables proposed, reliability analysis was carried out on the items which remained of each latent variable to examine internal consistency of subscales and determine whether future analyses could be conducted using these subscale. Overall results are reported in Table 7.3.2.

Table 7.3.2: Alpha coefficients for 9 remaining latent variables

Latent cause	No. of items	Scale reliability
Team communication	6	.73
Routine procedures	2	.45
Workload	2	.62
Human resources	4	.63
Local working conditions	4	.54
Ward climate	4	.56
Bed management	2	.39
Supervision	2	.23
Written policies and procedures	2	.49
Overall	**28**	**.90**

Although the OSQ was successful in measuring organisational safety overall, factor analysis revealed that the 10 proposed subscales representing each latent variable did not emerge. Possible methodological reasons behind this will be discussed in section 7.4

[42] TC3, TC7, RP1, RP2, RP3, RP4, RP6, RP8, RP9, HR3, HR5, HR7, HR8, HR9, LWC2, LWC5, CLT1, CLT3, CLT4, CLT10, BED1, BED7, TR4, TR6, TR7, WPP3, WPP4, WL6, SP1, SP2 were all removed (see chapter 5, table 5.5.b for complete item listing with corresponding codes).

7.3.3. Descriptive statistics

Sample details

The OSQ was considered relevant for all *qualified* nurses working in 34 clinical areas (see section 7.3.1, Tables 7.3.1.a & b). Of the total number of participants for whom the OSQ was considered relevant (n=918), 367 nurses responded (before removal of all not applicable cases), representing a 40 per cent response rate. Comparable response rates were found in a similar study conducted in secondary care to measure the validity and reliability of the Teamwork and Safety Climate Survey (Hutchinson *et al.,* 2006). After removal of not applicable cases, the sample comprised of 8 matrons, 60 sisters, 94 senior staff nurses, 126 staff nurses, 4 student nurses and 19 midwives (total n=311). The majority of nurses stated they worked between 20-39 hours per week (83 per cent), with 5 per cent working fewer and 12 per cent working more hours per week than this. In terms of the number of years nurses had worked on their *current ward*, 8 per cent had worked there less than a year, 39 per cent between 1-5 years, 25 per cent between 6-10 years, 10 per cent between 11-15 years and 18 per cent said they had worked there for 16 years or more. To retain anonymity, no other personal details were asked of participants.

Patient safety culture and incident reporting

As mentioned previously, 4 items were added to the OSQ from the *Hospital Survey on Patient Safety Culture* (Sorra & Nieva, 2004). One item asked respondents to grade their ward/unit on patient safety from failing to excellent. It was anticipated that these grades would be associated with scores on the OSQ. Individuals with the poorest levels of safety (as indicated by high scores on the OSQ) would be more likely to assess their wards as poorer in terms of patient safety compared to individuals working on 'safer' wards. Results revealed 5 per cent of the sample rated it as 'failing' or 'poor', 33 per cent as 'acceptable', 47 per cent as 'very good' and 15 per cent as 'excellent'. The remaining three questions measured the frequency of reporting errors dependent on the outcome, from 1 (not at all) to 5 (nearly all the time). Mean responses are presented in Table

7.3.3. Notably, participants said they were more likely to report mistakes nearly all the time they occurred when the mistake could have harmed the patient (question 1) than when it had been caught and corrected (question 2) or had no potential to harm the patient (question 3). A one-way within-subjects ANOVA revealed this difference was statistically significant ($F_{(1.93,\ 540.8343)} = 31.05$, $p<.001$).

Table 7.3.3: Likelihood of reporting mistakes

Item	Mean (Std)
1. When a mistake is made that *could harm the patient*, but does not, how often is this reported on your ward?	3.64 (1.36)
2. When a mistake is made, but is *caught and corrected before affecting the patient*, how often is this reported on your ward?	3.03 (1.33)
3. When a mistake is made, but has *no potential to harm the patient,* how often is this reported on your ward?	3.20 (1.39)

Good internal consistency of items (Cronbach's alpha = .73) allowed a reporting likelihood scale score to be calculated for further analysis. The mean reporting likelihood score for the sample was 3.29 (std. 1.10).

The 9-item *Safety Organising Scale* (SOS) was also administered within the OSQ questionnaire (Vogus & Sutcliffe, 2007) in order to examine concurrent validity of the OSQ measure (see section 7.3.4). This scale involves responding to positively framed items from 1 (not at all) to 7 (to a very great extent). Descriptive statistics revealed an overall sample SOS mean of 4.61 (Std.1.25) and excellent consistency of items (Cronbach's alpha = .93).

---

[43] Since Mauchley's test of sphericity was significant – values reported using greenhouse-Geisser correction.

OSQ items

Since factor analysis did not extract the 10 latent variables proposed and reliability analysis revealed poor to moderate alpha coefficients for remaining latent variables, all statistics conducted in remaining sections will be based on the overall 28-item organisation safety construct (see appendix XVI for final questionnaire). Participants were required to respond to all items from 1 (not at all) to 5 (nearly all the time). It is important to emphasise here that high OSQ scores are reflective of poor organisational safety. To deal with missing data the following procedures were followed;

1. NA and missing responses were recoded to zero so they would not affect the overall sum
2. A worst possible OSQ outcome score was calculated for each participant based on the number of non-zero responses multiplied by 5 (the highest response option indicating the indicator occurred nearly all the time). This provided a value which was reflective of the highest OSQ score each participant could have had based on the number of items they had completed. For example, if a participant had responded to 23 out of the 28 possible items, their worst possible OSQ score would be 115 (23 x 5).
3. A sum of *all* responses was calculated (including zero responses which would not affect this sum) and divided by the 'worst possible OSQ outcome' score and multiplying by 100 to provide a percentage (Brace *et al.,* 2003, pp. 133). Hypothetically, a participant who had responded 5 (nearly all the time) to all 28 items would have an OSQ percentage score of 100, indicating the worst levels of safety on that ward.

The overall mean OSQ percentage score (referred to from this point forward as just OSQ score) for the sample (n=311) was 40.24 (Std. 11.19). In order to illustrate whether the OSQ could discriminate between different types of clinical areas, a two-way between-groups ANOVA was conducted on 5 wards[44] chosen as being of similar size with similar response rates. Results revealed

[44] Five wards were chosen simply to illustrate this point as opposed to testing the difference between *all wards* (which was also significant overall), the reporting of which (including post-hoc tests for each significant difference) would have been beyond the bounds of this section.

an overall significant difference in OSQ scores between a neonatal unit, the maternity delivery suite, an adult female surgical ward, a renal dialysis unit, and a large paediatric medical ward ($F_{(4,68)} = 17.98$, $p<.001$). Post hoc tests revealed that the paediatric medical ward had significantly higher OSQ scores (indicating the poorest levels of safety) than other 4 wards measured.

DRBQ items

Preliminary analysis of Drug Round Behaviour Questionnaire items revealed data was significantly positively skewed (as indicated by K-S values and histograms). Data was so skewed in fact that no single item had been answered 'nearly all the time' or 'frequently' (2 of the 5 response options). Although reduced variability was anticipated to some degree given the sensitive nature of these items, the pilot study carried out in study 3 suggested there would be a little more variability in the data (see chapter 6, Table 6.5.2.2.a). To enable useful analysis, responses were collapsed dichotomously into 'yes' (any response between 'occasionally' and 'nearly all the time') and 'no' (retaining 'not at all' response) categories. Because of this, factor analysis to explore the underlying structure of this measure was not possible. Nevertheless, reliability analysis revealed a Cronbach's alpha coefficient of .68 which was deemed fair considering the limited range of responses. It was not possible to simply calculate a sum of the number of medication administration errors (MAEs) to which each participant had responded 'yes' (they had executed in the last 4 weeks) since there were still not applicable and missing data within this dataset. Instead, a frequency percentage of all MAEs was calculated by counting the total number of 'yes' responses per participant and dividing by the number of valid (as opposed to non-valid missing or not applicable responses) DRBQ responses multiplied by a hundred. This value therefore represented the frequency with which MAEs had been committed as a percentage of the number of DRBQ items participants had answered. The mean frequency of MAEs in the 4 weeks preceding administration of the tool for the overall sample was 10.70 and ranged from 0 (no errors executed) to 91.67 (12 of the 13 MAEs executed).

In order to determine whether it was possible to distinguish between wards on DRBQ scores, a one-way between-groups ANOVA was conducted. Using the same 5 wards described earlier as an example, there was a statistically significant difference between DRBQ scores across the 5 wards ($F(4, 68) = 2.77$, $p<.05$). Post-hoc tests revealed that the paediatric medical ward had significantly higher DRBQ scores (indicating they had committed significantly more errors) than the other 4 wards measured. This was consistent with this ward having the highest OSQ scores.

7.3.4. Concurrent validity

An important objective of study 4 was to measure the concurrent validity of the OSQ – that is the extent to which it was associated with a similar measure. The 9-item *Safety Organising Scale* (SOS; Vogus & Sutcliffe, 2007) was considered a suitably comparable measure with which to achieve this since it has been found to be a statistically valid (concurrent with similar measures) and reliable self-report measure of nursing behaviour which encapsulates underlying safety culture. Pearson's correlation analysis revealed a statistically significant association between individual's OSQ and SOS scores ($r = -.40$, $p<.001$) and therefore good evidence of concurrent validity. This relationship was inverse because higher scores on the OSQ indicated poor levels of safety (since items were negatively framed) whilst higher levels on the SOS indicated a good safety culture (these items were positively framed).

The OSQ was developed as a ward-based tool (as opposed to a measure of individual perception) and it was anticipated that wards with the highest OSQ scores (based on their aggregate OSQ score) would also be those with the worst ward safety cultures (as measured by the SOS) compared to wards presenting lower safety risks. Pearson's correlation analysis revealed a significant association between *ward* OSQ and *ward* SOS scores ($r = -.69$, $p<.001$). Wards with higher OSQ scores were those with the lowest SOS scores supporting the validity of the OSQ as a ward-based measure.

7.3.5. Criterion validity

The main aim of developing the OSQ was to construct a *proactive* measure of organisational safety which could predict future MAEs. Therefore, scores on the OSQ should be able to predict scores on a measure of drug administration errors. However, as previously discussed, inherent biases with current system of NHS reporting cast considerable doubt on whether data from incident reports would be a useful or accurate reflection of error occurrence (see chapter 6, section 6.2). To account for this, the DRBQ was developed as a measure of specific drug administration errors which nurses had previously been willing to report formally (since many of these items were developed from incident report data; see chapter 6) but which were framed as potential near misses in an attempt to encourage 'open' reporting. It is acknowledged that predictive validity measures are ideally administered *after* the measure being validated (Clark-Canter, 2004), however time and sample accessibility constraints in this study meant that this was not possible and so all items were administered at the same time. Regression and correlational analyses were performed to test five research questions.

1. Does the OSQ predict medication administration errors?

Linear regression analysis was carried out to determine whether scores on the OSQ would predict the number of MAEs committed in the previous 4 weeks as measured by the DRBQ. Analysis revealed OSQ scores were significantly predictive of scores on the DRBQ ($F_{(1,309)} = 67.82$, $p<.001$) and accounted for 18% of the variance in MAE scores (Adjusted $R^2 = .18$; $\beta = .43$). Individuals who had high OSQ scores (indicating poor safety levels on their ward) admitted making significantly more MAEs in the previous 4 weeks than those with lower OSQ scores.

Since scores on the OSQ were shown to be significantly associated with scores on the SOS and previous research has shown the SOS to be predictive of formally reported medication errors (Vogus & Sutcliffe, 2007) multiple linear regression analysis using the enter method was conducted

to determine which of the two safety measures would be a *better* predictor of MAE frequency. Table 7.3.5.a summarises these findings.

Table 7.3.5.a: Multiple linear regression analysis: Comparison of the predictive validity of OSQ and SOS scores

Model	Beta	t	Adjusted R^2	R^2 change	F	Sig.
1: Only SOS scores	-.19	-3.30	.03	.03	10.91	.001
2: SOS & OSQ scores	.43	7.27	.18	.15	32.85	.000

Entering SOS scores as the first predictor variable, results revealed that model 1 was a significantly predictive of MAE scores ($F_{(1, 298)} = 10.91$, $p<.005$) and accounted for 3 per cent of the variance (Adjusted $R^2 = .03$; $\beta = -.19$). The inclusion of OSQ scores into the model in step 2 resulted in an *additional* 15 per cent of the variance in DRBQ scores being explained ($\Delta R^2 = .15$; $\beta = .43$) and as such was a significant predictor of DRBQ scores ($F_{(2, 297)} = 32.85$, $p<.001$). Although Vogus and Sutcliffe (2007) found the SOS to be significantly predictive of medication incidents, these results show the OSQ to be a *better* predictor of MAEs. This finding was to some extent anticipated since the OSQ was designed specifically to predict medication administration errors and comprised more items of a significantly broader range than the SOS to reflect this.

2. Does the OSQ predict formally reported incidents?

It was hypothesised that underreporting biases in formal NHS incident reporting systems would make these incidents an unsuitable outcome measure against which to test the predictive validity of the OSQ. Nevertheless, Vogus and Sutcliffe (2007) report that in their study the SOS successfully predicted 'formally reported medication errors' and 'patient falls'. To measure whether OSQ scores would be similarly predictive of formally reported errors, data was gathered from the collaborating

hospital on the number of Formally Reported Incidents[45] (FRIs) in the 3 months preceding the administration of the questionnaire[46]. Data was collected on 5 types of incident; patient slips and falls, near misses, medication incidents; blood transfusion incidents and fluid infusion incidents. Because this data was only available at a ward level, mean OSQ, DRBQ and SOS scores were also calculated for each ward (n=34). To account for the substantial difference in ward sizes and patient intake (and the number of staff working on each ward) the total number of incidents was calculated as a function of the number of qualified nurses working on each ward to obtain an average number of reported incidents (the total number divided by the number of staff on the ward). In doing this, it was hoped that a fairer comparison of incident frequency could be made. For example, the ICU is a very large department employing almost 70 qualified nurses and will undoubtedly report a significantly larger number of errors than the breast care unit which is comparatively small employing only 11 qualified nurses. Preliminary correlation analysis between OSQ scores and all five types of incident are presented in table 7.3.5.b.

Table 7.3.5.b: Correlations between FRIs and OSQ scores

Formally Reported Incidents	**OSQ score** *(Sig.)*
Patient slips and falls	.26 *(p = .15)*
Near misses	.04 *(p = .83)*
Medication incidents	.001 *(p = .99)*
Blood transfusion incidents	.32* *(p = .08)*
Fluid infusion incidents	-.10 *(p = .57)*
Mean number of incidents	.33* *(p = .07)*

* Correlations almost significant at the .05 level

---

[45] These events are referred to as 'incidents' rather than errors since it is not known whether the events had occurred as a direct result of error.

[46] Although this was almost 3 times longer than the frequency criteria on the DRBQ (the previous 4 weeks), comparatively small numbers of reported events meant that a longer time period was necessary.

As this table shows, there were no statistically significant associations between OSQ scores and any of the 5 types of FRI or the mean number of FRIs. However, since the relationship between the mean number of FRIs and OSQ scores was approaching significance (p=.07), linear regression analysis was conducted to explore whether OSQ scores were significantly predictive of the mean number of FRIs overall. Using the enter method, results revealed that OSQ scores were not significantly predictive of the mean number of FRIs ($F_{(1, 30)} = 3.63$, $p = .07$) and accounted for only 8 per cent of the variance (Adjusted $R^2 = .08$; $\beta = .33$).

As correlation analysis suggested, regression analysis revealed that OSQ scores did not predict any of the 5 independent types of FRIs alone, with medication incidents having the worst relationship with OSQ scores ($F_{(1, 30)} = .000$, $p = .99$; Adjusted $R^2 = -.03$; $\beta = .001$). It is important to note however that the beta values for the relationship between OSQ scores and blood transfusion incidents and mean incidents overall (.32 & .33 respectively) were quite high considering there were so few cases (e.g. 34 wards) entered into the model. It is proposed that by increasing the statistical power (i.e. obtaining data from more wards) it is likely that this relationship would have reached significance. Table 7.3.5.c summarises these findings.

Table 7.3.5.c: Summary of linear regression analyses for OSQ scores and 5 types of FRI

Regression	Beta	t	Adjusted R^2	F	Sig.
1: OSQ scores & patient slips/falls	.26	1.47	.04	2.16	.15
2. OSQ scores & near misses	.04	.21	-.03	.05	.83
3. OSQ scores & medication incidents	.00	.01	-.03	.00	.99
4. OSQ scores & blood transfusion incidents	.32	1.82	.07	3.32	.08
5. OSQ scores & fluid infusion incidents	-.10	-.57	-.02	.33	.57

Vogus and Sutcliffe (2007) found SOS scores to be predictive of formally reported medication incidents and patient slips and falls in a US sample. Correlational analysis was conducted to examine this relationship in this UK sample. Table 7.3.5.d summarises these findings.

Table 7.3.5.d: Correlations between FRIs and SOS scores

Formally Reported Incidents	**SOS score** *(Sig.)*
Patient slips and falls	**-.35** ***(p = .05)****
Near misses	.05 *(p = .77)*
Medication incidents	-.05 *(p = .77)*
Blood transfusion incidents	-.20 *(p = .28)*
Fluid infusion incidents	-.003 *(p = .99)*
Mean number of incidents	-.28 *(p = .12)*

* Correlation almost significant at the .05 level

In line to some extent with Vogus and Sutcliffe's findings, the relationship between SOS scores and the number of patient slips and falls approached statistical significance (p=.05). However as indicated above, there was no further association between SOS scores and any of the other 4 types of FRI or the mean number of reported incidents.

Results revealed that while OSQ and SOS scores were not statistically significant predictors of FRIs, including those categorised as medication-related, these scores did predict drug administration errors as measured by the DRBQ. Participants who scored high on the OSQ and low on the SOS (both indicating poor safety levels) reported they had committed more MAEs in the previous four weeks compared with individuals with better safety levels. However, it is important to emphasise that the power achieved in regression analysis using whole sample DRBQ data (n=311)

was significantly greater than that achieved using ward-based FRI data (n=34) which could explain why OSQ and SOS scores predicted DRBQ scores but not FRIs. It is likely that OSQ scores would have significantly predicted FRIs by increasing the number of wards sampled and increasing statistical power.

Statistical power issues aside, it is also possible that since OSQ and SOS scores significantly predicted errors measured by the DRBQ, formally reported incidents either capture more than just errors or else are not a measure of error at all. To examine this hypothesis, correlational analysis was conducted to examine the relationship between mean ward MAE scores as measured by the DRBQ and all five types of FRI and mean FRIs overall. Table 7.3.5.e summarises these findings. As this table shows, DRBQ scores were not significantly associated with any of the 5 types of FRI or mean FRIs overall.

Table 7.3.5.e: Correlations between DRBQ scores and FRIs

Formally Reported Incidents	**DRBQ score** *(Sig.)*
Patient slips and falls	.22 *(p =.23)*
Near misses	-.13 *(p =.47)*
Medication incidents	-.03 *(p =.89)*
Blood transfusion incidents	.21 *(p =.25)*
Fluid infusion incidents	-.01 *(p =.96)*
Mean number of incidents	.26 *(p =.15)*

3. Do 'safer' wards commit fewer errors?

Individuals with the highest OSQ scores (indicative of poor safety) were more likely to commit significantly more MAEs than individuals with lower OSQ scores. However, since the OSQ was developed as a ward-based tool it was important to measure the ability of the OSQ to distinguish between ward based scores. It was hypothesised that *wards* with the poorest levels of safety (high OSQ scores) would commit significantly more MAEs than *wards* with higher levels of ward safety. Pearson's correlation analysis revealed there was a significant association between OSQ scores and MAEs committed at a ward level ($r = .49$, $p<.01$), supporting this hypothesis.

To examine whether there was a particular OSQ score at which point wards were considerably more likely to commit MAEs, wards were split into 'high', 'medium' and 'low' safety-risk categories by percentiles. Low risk wards had mean OSQ scores between 0 - 38.32, medium risk between 38.33 - 44.14 and high risk wards scored 44.15 and above. A one-way between-groups ANOVA on the ward data (n=34) revealed there was a statistically significant effect of ward safety level on DRBQ score ($F_{(2, 31)} = 9.06$, $p<.01$). Bonferroni post-hoc tests indicated this difference was significant between low and medium (mean diff. = 7.08, $p<.01$), and low and high (mean diff. = 8.16, $p<.01$) but not between medium and high safety-risk wards. Therefore, high and medium-safety risk wards reported significantly more MAEs than low safety risk wards. For future error management purposes, wards with OSQ scores within the low safety-risk range (0-38.32) have better organisational safety and as such are less likely to make drug administration errors.

As already discussed, OSQ scores overall were not significantly predictive of formally reported incidents (FRIs). When wards were split into high, medium and low safety risk categories there remained a non-significant effect of ward safety level on the number of FRIs ($F_{(2,29)} = 1.30$, $p = .29$) and there was no significant effect of ward safety level on any of the five types of FRI, including medication incidents ($F_{(2, 29)} = .07$, $p = .93$).

4. Is the OSQ a good predictor of patient safety perception?

One item on the questionnaire asked participants to grade their ward in terms of patient safety as being either failing, poor, acceptable, very good or excellent. Linear regression analysis using the enter method revealed that OSQ score was a significant predictor of patient safety grade ($F_{(1,309)} = 57.80$, $p<.001$), accounting for 16 per cent of the variance (Adjusted $R^2 = .16$; $\beta = -.40$). Therefore, participants who scored highly on the OSQ (indicating poor levels of safety) would be more likely to give their ward a poor grade on overall patient safety. This finding lends support to the content validity of the OSQ since items were a good indicator of patient safety perception.

5. Are 'safer' wards more likely to report their future mistakes?

It was anticipated that wards with the highest OSQ scores (poorest levels of safety) were more likely to have relatively 'closed' reporting cultures and thus less likely to report errors formally (see chapter 6, section 6.2). Indeed OSQ scores were not as significantly predictive of formally reported incidents overall (FRIs) as they were of MAEs (measured by the DRBQ). To examine this issue more closely participants were asked to rate the likelihood they would formally report errors depending on the patient outcome, from 1 (not at all) to 5 (nearly all the time). Using ward-based data, Pearson correlation analysis revealed a significant association between OSQ scores and reporting likelihood ($r = .51$, $p<.01$).

Using the same method of dividing wards according to safety level described above, a one-way between-groups ANOVA revealed a statistically significant effect of ward safety level on the likelihood of reporting errors ($F_{(2, 31)} = 5.98$, $p<.01$). Post-hoc analysis revealed this difference was significant between low and high safety risk wards only (mean diff. .41, $p<.01$). Wards categorised as high safety risks were significantly less likely to formally report errors than low safety risk wards.

7.4. Discussion

This study was conducted to measure the validity and reliability of the Organisational Safety Questionnaire (OSQ). The original 107-item[47] questionnaire was administered in two rounds to both qualified and unqualified nurses working in 54 clinical areas across two hospitals and was found to be appropriate for all *qualified* nurses in 34 of these areas. Although it was proposed that items on the OSQ represented 10 latent causes of medication administration errors (MAEs), factor analysis revealed 28 items representing one construct named organisational safety which had good internal reliability. There were several key findings. Firstly, the OSQ had good levels of concurrent validity with a valid and reliable 9-item measure of ward safety culture (*Safety Organising Scale*, (SOS); Vogus & Sutcliffe, 2007) and a 1-item measure of patient safety perception (*Hospital Survey on Patient Safety Culture*; Sorra & Nieva, 2004). Individuals and also wards which had high OSQ scores (indicative of poor organisational safety) had worse perceptions of patient safety and poorer ward safety cultures as measured by the SOS.

To test its predictive validity, the OSQ was compared against six outcome measures; the total number of MAEs committed in the previous four weeks as measured by the DRBQ and five types of formally reported incident (FRIs). Scores on the OSQ were significantly more predictive than the SOS of DRBQ scores. Furthermore, high safety risk wards committed significantly more MAEs than wards which posed a lower safety risk. Neither the OSQ nor the SOS predicted FRIs overall or any of the five different types of incident, including those which were medication related. However, as mentioned previously the ability of the OSQ to predict FRIs was *almost* significant and insufficient statistical power achieved from data from only 34 wards should not be underestimated. It is recommended that future research should increase this statistical power by sampling more 'relevant' wards to explore other types of reported incident which the OSQ may be able to predict (e.g. patient complaints).

---

[47] Which comprised 82 OSQ items, 13 DRBQ items, 9 SOS items and 3 reporting likelihood items.

Scores on the DRBQ were not significantly associated with FRIs. It is therefore possible that FRIs constitute adverse events and not errors. Early stages of reviewing incident reports as part of study 1 (see chapter 4) revealed that many reported incidents related to abusive patients or visitors during drug rounds, theft of medication, absconding patients, slips and falls and other events which did not involve errors. These findings could suggest that while FRIs may be an appropriate indicator of organisational safety culture, they might not be an appropriate measure of errors and capture only adverse events which individuals consider appropriate or are willing to disclose. This finding is particular important since FRIs are the most commonly cited measure of errors in patient safety research and frequently used in the development of safety related measures.

In terms of reporting culture, it was predicted that wards with the poorest levels of organisational safety (i.e. those with the highest OSQ scores) would formally report fewer incidents or errors than safer wards due to 'closed' reporting cultures. If this were the case, wards categorised as high safety risks would have reported significantly lower levels of FRIs than lower risk wards. However, this was not found. There was no significant difference in the numbers of FRIs reported by high, medium and low safety risk wards.

The items administered from the *Hospital Survey on Patient Safety Culture* (Sorra & Nieva, 2004) to explore reporting likelihood may shed some light on these findings. Wards categorised as high safety-risks stated they were significantly less likely to *formally* report their errors than lower risk wards. It is proposed that poorer safety cultures mean that high safety-risk wards feel less inclined to report their mistakes and so formally report a comparatively smaller percentage of the errors which *actually occur* compared to lower risk wards. Interestingly, though there was no difference between the groups on the number of FRIs reported, high safety risk wards admitted making significantly more MAEs on the DRBQ than wards presenting lower safety risks. With this in

mind, it is possible that the DRBQ was a more sensitive measure of the *actual* number of drug administration errors occurring on wards.

One of the limitations of this study was the failure to support the existence of the 10 latent preconditions of medication administration errors identified by qualitative methods in study 1 (see chapter 4). This does not mean to say however that the 10 variables identified are not predictors of MAEs, only that they were not well represented in this tool. It may have been the case that after the removal of cases for which the OSQ was generally not applicable, the sample size was simply too small for the number of variables being measured. However, it is widely accepted that beyond 300 participants, parameters tend to be fairly stable[48]. It is possible that by phrasing some items as representing the respondent (i.e. "**I** have worked with an agency nurse who was less qualified....") and others more generically (e.g. "There has been a deterioration of the skill level on this ward...") may have resulted in different responding patterns for each type of question. Often referred to as item characteristic effects (Podsakoff *et al.,* 2003), this form of bias refers to the influence specific characteristics of questionnaire items may have over responding tendencies. However, in this study early attempts to factor analyse 'I' items (n=58) and 'non-I' items (n=24) separately did not improve either the factorability of items or the interpretation of a factor solution (i.e. items still fractured into multiple uninterpretable factors). Indeed sample mean and median scores for 'I' and 'generic' items were virtually identical[49]. It is proposed that because OSQ items were neither attitudinal nor behavioural (items rarely asked respondents to comment on their own behaviour), the difference in wording did not unduly affect responding variance. Nevertheless, it may be useful in future versions of the questionnaire to standardise such wording to further reduce the risk of introducing bias in responding.

---

[48] Kass and Tinsley (1979) recommend having between 5 and 10 participants per variable up to a total of 300 at which point parameters tend to be fairly stable regardless of participant to variable ratio. Similarly Tabachnick and Fidell (2001) agree it is 'comforting to have at least 300 cases for factor analysis'.

[49] These descriptive statistics are not reported within this thesis since they not relevant to further statistical analysis since items were not analysed in this way.

Alternatively, it is possible that findings were confounded to some extent by common method variance. For example, research has found that in an attempt to appear consistent and rational in their responses, questionnaire respondents will search for similarities within questions and frame their responses accordingly (Johns, 1994; Schmitt, 1994). In this way, findings can represent over-inflated relationships between measured constructs. It is possible that in this study, presenting OSQ, SOS and DRBQ items within the same questionnaire may have increased the likelihood of this consistency effect and thus over-inflated the correlation between OSQ, SOS and DRBQ scores. However, in a review of various types of common method variance, Podsakoff *et al.*, (2003) acknowledge that this bias effect in particular is largely problematic in questionnaires which measure multiple constructs of respondents' *attitudes* and *perceptions*. Although SOS items were attitudinal (and as such were presented separately on the questionnaire), DRBQ and OSQ items were not attitudinal, but were developed as behavioural items and observable indicators of systems problems respectively. As such, these items were less likely to result in consistency bias (Podsakoff *et al.*, 2003). Furthermore, Kline (2000) recommends that where 'different' constructs are being measured within the same questionnaire, 'intermixing' items (as was done in this questionnaire) should reduce common method variance, and in particular social desirability bias since respondents are less likely to 'work out' the constructs and respond desirably. However, there is also evidence which suggests that intermixing items in this way can also lead to reduced reliability since respondents are also less likely to see similarities in items measuring the same construct. This could explain the findings from factor analyses which did not support a 10 factor model. Nevertheless, it should be acknowledged that DRBQ items *were* naturally more likely to be affected by social desirability bias given the sensitive nature of these items compared with OSQ items. Findings from study 1 (e.g. observations and interviews) highlighted a poor reporting culture within the nursing population and a fear of blame and reprisal. For this reason, particular attention was given to the phrasing of DRBQ items, attenuating any wording which could have been perceived as potentially judgemental or 'blameworthy'. These items were also framed as near misses to increase

honest responding and reduce socially desirable responding. Despite this, response variability was extremely low for these items which could either suggest that respondents were responding in a socially desirable way or were simply responding honestly to items with poorly referenced time frames (e.g. 'In the last 4 weeks…').

In addition to problems associated with administering multiple constructs simultaneously Podsakoff *et al.,* (2003) emphasise the problem of administering both criterion (DRBQ) and predictor (OSQ) measures at the same time. Although administering these measures separately is the preferred method, this was not possible. In this study, constraints upon sample availability (due to multiple research projects being conducted in this hospital at the same time) and time constraints meant that it was not possible to administer the three measures independently which would have reduced the potential for this and other aforementioned types of bias. However, in this type of applied research setting it is important to find a balance between reducing potential bias effects which may confound findings and working within the parameters of such an organisation.

Factor analysis may not have resulted in the extraction of the ten identified latent preconditions due to a pragmatic need for keeping the tool concise and which lead to a relatively small number of items representing each variable. It is commonly suggested that within instrument development it is essential to have many more items than required for the final tool. Kline (2002) recommends developing twice the number needed, particularly when multiple factors are being addressed as was the case in this study. However, because of the high complexity of the data and the pragmatic needs of keeping the tool concise (i.e. 82 items representing 36 subcategories reflective of 10 latent variables) it was not always possible to have multiple items to represent subcategories and in some cases even though there were multiple items reflecting the overall variable, some of the component subcategories were underrepresented. Furthermore, once items had been removed which had high levels of not applicable responding some of the predictors had only 3 or 4 items remaining (e.g.

workload, bed management) each of which measuring a different aspect of that construct. The proposed structure of variables may therefore have been over-complex without sufficient items with which to test these sub-structures at this stage of research. In this way, the tool may have been trying to measure too many aspects of safety. As discussed previously, this may have been rectified by using more items to represent fewer themes in one measure. However, since it was not known which of the 10 latent variables were *most* predictive of medication errors, and as such should be included in the tool, this was not possible. An alternative method to explore the underlying structure of the longer version of the OSQ may have been to administer different subscales of the OSQ in several stages using multiple samples. In each stage 2-3 latent variables, each with 15-20 items could be administered to a sample to examine the item cohesion and aid in the removal of redundant items until all 10 variables had been administered. Unfortunately as with any type of applied research, particularly within health care, constraints on time and sample availability meant that such a stratified approach was not possible.

A second limitation of the study was the relatively low level of response variability which ultimately affected subsequent analysis choices. It is possible that this was due to problems with the time frames, which may have been too short (e.g. 'In the last 2 weeks/4 weeks/2 months...' etc.). However, previous pilot studies had not highlighted any such problems with corresponding time frames. It is proposed that the nature of the data may have simply been too sensitive in this sample of nurses who are unlikely to have been asked for this type of information before. Since data responses were not only of limited range, but were extremely positively skewed (i.e. responses generally clustered around the lower end of the scale) this is the most likely explanation. It has been recommended that to improve reporting of sensitive information attempts should be made by researchers to familiarise themselves with the sample before administration of questionnaires (Edwards *et al.*, 2002). While this was largely achieved during in study 1 and subsequent pilot studies and which undoubtedly contributed to the honesty with which participants disclosed

information, this was not possible in the large-scale final study which involved a potential sample of more than 2000 nurses.

7.5. Conclusions

Originally developed as 10 scales representing 10 latent conditions identified during qualitative research to be predictive of MAEs, the final 28-item scale comprised one overall construct named organisational safety. While the final scale did not consist of individual subscales, the 28 remaining items did reflect 8 of the 10 proposed preconditions of MAE and so does to some extent support the role of these preconditions in the occurrence of MAE. The final 28-item OSQ had good internal consistency, was correlated with an independent measure of organisational safety and was predictive of medication administration errors. The OSQ was not significantly predictive of formally reported incidents which may in part be due to the nature of the events individuals are willing to report and which may not accurately reflect the frequency of errors. Alternatively, this may have been an artefact of low statistical power which could have been improved with a larger sample. Finally, high safety risk wards said they were less likely to *formally* report their errors than lower safety risk wards. However, high, medium and low safety risk wards actually formally reported a similar number of incidents. It is possible that high risk wards report a comparatively smaller percentage of the errors which *actually occur* compared to lower risk wards due to poorer safety cultures. Interestingly, high safety risk wards admitted making significantly more MAEs on the DRBQ than wards presenting lower safety risks suggesting the DRBQ may be a more sensitive outcome measure of the *actual* number of drug administration errors occurring.

CHAPTER 8
CONCLUSIONS

8.1. Introduction

The overall objective of this research was to explore latent causes of medication administration errors in secondary care and to develop a valid and reliable proactive measure of these error causes similar to those already available in other safety-critical industries. This chapter will present a brief overview of key findings. Methodological limitations will then be considered before discussing the practical implications of these findings and recommendations for future research.

8.2. Summary of findings

Chapter 2 of this thesis discussed the 'systems' approach to understanding human error whereby latent conditions develop over time within an organisational system from decisions made by top-level management, designers and procedure writers. Reason's Swiss cheese model was presented as a framework for understanding the interplay between these latent factors at the 'blunt-end' of the organisation and unsafe acts at the 'sharp-end' (Reason, 1990). It was noted that although well-utilised in other high-risk industries, there has been few attempts to measure the validity or reliability of this framework or make use of it to explain errors which occur in medicine. This finding was supported by the systematic review presented in chapter 3 which identified studies that had employed systems interventions to reduce the occurrence of medication error. Almost all of these studies gave no indication that *any* theory had been used as a basis for their intervention and most provided sparse empirical evidence to support their chosen interventions. Overall, studies which had carried out multiple interventions provided insufficient methodological details to enable replication and had not reported whether independent components of those interventions had been successful. More than half of the studies reviewed had implemented computerised systems to

reduce medication errors driven largely by 'evidence-light' government policy. This atheoretical faith in technological interventions (Kremsdorf, 2005) could mean that only a limited number of latent failures are currently being addressed in health care. It was concluded in chapter 3 that a tool which could measure a wide range of latent failures and monitor them over time would be extremely valuable in health care for the development of appropriate interventions to reduce errors.

With a view to developing such a tool, study 1 was conducted to explore possible latent causes of medication administration errors in secondary care. Semi structured interviews, clinical observations and a review of a number of incident reports were conducted. While observational data was useful to some extent, the lack of clinical training of the observer meant that this data was not as detailed as interviews which represented the richest and most useful source of information. Formally reported incidents offered little insight into the causes of error and due to the general paucity of detail provided comprised the least useful source of qualitative data. Nevertheless, combined analysis of the three data sources identified ten potential latent causes of medication administration errors (MAEs). Using research conducted in the oil industry to develop proactive measures of latent failure as a basis for test construction, items were developed to represent observable indicators of each of the ten latent causes of MAEs. Pilot studies revealed that after the removal of items with poor response variability and the replacement of dichotomous response options with a Likert frequency scale to improve this variability, the 82-item Organisational Safety Questionnaire (OSQ) had good face validity.

It was a key aim of this research to produce a tool which, unlike similar proactive measures developed in other high-risk industries, had good predictive validity. Research has suggested that most medical errors are not 'officially' reported and that wards with the poorest levels of organisational safety would also be those who are less likely to report their errors due to 'closed' reporting cultures (Cullen *et al.*, 1995). Therefore, formally reported incidents may not have represented an appropriate outcome measure against which to test the predictive validity of the

OSQ. To overcome this, an independent measure of MAEs was developed. It was important for items on this measure to be objective (not open to interpretation) and not reliant upon adverse patient outcomes in order to encapsulate near misses. Items were developed from interview and incident report data collecting during study 1 and represented ten types of MAE as proposed by a well-known taxonomy of error. A pilot study revealed that after the removal of items with poor variability, the 13-item Drug Round Behaviour Questionnaire (DRBQ) had good face validity and measured seven types of MAE.

Chapter 7 of this thesis described study 4 which measured the concurrent and predictive validity, internal reliability and the generaliseability of the OSQ. Findings showed that the OSQ was relevant for qualified nurses working across a wide range of clinical specialties. The questionnaire had good concurrent validity with another valid and reliable measure of local safety culture (SOS; Vogus & Sutcliffe, 2007) and was significantly more predictive of MAEs than this independent measure. Problems with skewed responses which lead to the removal of large number of items meant that factor analysis did not support the OSQ as ten individual scales as had been predicted. Nevertheless, the remaining 28 items reflected eight of the ten proposed latent conditions, a finding which supports to some extent the role of these conditions in the occurrence of MAEs.

Further analysis revealed that wards with high levels of latent failure (measured by the OSQ) had poor organisational safety cultures (measured by the SOS) and admitted to making significantly more MAEs than 'safer' wards. While scores on the OSQ and the SOS were not *significantly* predictive of formally reported incidents, this may have been an artefact of the low statistical power obtained using ward-based data which may have been improved with a larger sample. Nevertheless, it is possible that formally reported errors represent more than just errors or are not an appropriate measure of error at all, particularly since their frequency was not associated with scores on the DRBQ. This was further supported by initial reviews of incident reports during study 1

which revealed many 'incidents' did not necessarily constitute errors. This finding is particularly significant because formally reported incidents are the most commonly cited measure of errors in patient safety research and are often used in the development of safety related measures.

Finally, findings revealed that wards classified as 'high-safety risks' formally reported a similar number of incidents as 'safer' wards, despite responding that they were significantly less likely to formally report their mistakes. It was hypothesised that high safety risk wards because of poorer safety cultures feel less able to disclose their mistakes and so formally report a comparatively smaller percentage of the errors which *actually occur* compared to lower risk wards. Although there was no difference between wards on the number of *formally* reported incidents, high safety risk wards reported more MAEs than safer wards using the DRBQ which might suggest this measure is a more sensitive measure of errors than those reported through formal channels. Alternatively, this may have been due to the anonymity offered by the DRBQ compared to the non-anonymous system of formally reporting incidents. Future research should focus on the difference between errors and incidents in terms of reporting likelihood.

8.3. Limitations

One particular limitation of the final study was the inclusion of a newly developed and as such unvalidated outcome measure (DRBQ) against which to measure the predictive validity of an equally unvalidated measure of organisational safety (OSQ). Findings revealed that although scores on the OSQ were predictive of scores on the DRBQ, they were not *significantly* predictive of the number of formally reported incidents; the most commonly employed outcome measure in this field of research. Ideally, the predictive validity of the OSQ would have been measured using a validated and reliable measure of errors. However, it is frequently acknowledged within the patient safety literature that there are multiple biases associated within non-anonymous formally reported incident

data (e.g. underreporting/selective reporting etc. see chapter 6, section 6.2 for more detailed discussion) which make its use as an outcome measure questionable. Indeed, qualitative review of a large number of formally reported 'medication incidents' in this study (see study 1, chapter 4) highlighted the difficulties of using this data as an appropriate measure of medication administration errors. For the majority of reported incidents it was not possible to determine whether an error had occurred and it was unclear whether the number of incidents reported represented an accurate reflection of the frequency of these events. Unfortunately, patient safety researchers also acknowledge that there is a distinct lack of an appropriate and reliable alternative measure of error available for this type of applied research (Flin *et al.*, 2000). With this in mind, the DRBQ was developed as an independent measure of errors. While significant attempts were made to ensure this measure had good face and content validity, the final study revealed especially low variability in responding which did not support evidence from previous small-scale pilot studies. It is possible that disseminating this measure alone to a larger pilot study (e.g 100+) may have refined this measure and improved its reliability and ultimately response variability *before* the final study. However, time and recruitment constraints meant that such a pilot study was not possible. Correlational analysis in study 4 revealed little or no relationship between scores on the DRBQ and formally reported incidents. Furthermore, there remains to date no available alternative measure of medication administration or any other type of medical error against which to test the validity of the DRBQ as an error outcome measure. It would be useful for future research to further investigate the validity and reliability of the DRBQ as an independent outcome measure.

In more general terms, this research was conducted in only one NHS Trust which may to some extent compromise the generaliseability of findings. Attempts to recruit two other hospitals were not successful due to ethical and time constraints. It was hoped that by conducting this research in a generic field of medicine (general medicine), findings would be common to multiple areas of care and also to secondary care environments in general. A plethora of supporting evidence for each of

the ten latent failures identified in study 1 (see chapter 4 for supporting evidence) and the relevance of the final tool to thirty-four clinical areas in study 4 suggest findings are generalisable beyond general medicine and indeed beyond this particular NHS Trust. In terms of cultural generaliseability, it was not possible to interview any foreign nurses during study 1 and so all interviews were conducted with white British nurses which may represent a biased sample. However, the aim of conducting multiple qualitative methods such as interviews, observations and reviewing incident reports was to gain a 'holistic' insight into error causation. Observations in particular went some way towards gaining an insight into the issues surrounding foreign nursing (e.g. language barriers, cultural differences); issues which may not have arisen naturally from white British nurses but which could then be further explored during interviews. It was then possible to develop OSQ items to reflect these issues. In terms of questionnaire completion, respondent's ethnicity was not measured and so it is possible that foreign nurses although sampled, did not respond. However, developed items were objectively framed in a clear and concise way and were designed specifically so they would not be open to interpretation and as such culturally sensitive.

Specific problems with poor response variability on the final tool have already been addressed elsewhere in this thesis (see chapter 7). However, it is important to explore the potential reasons why this effect occurred in the final study but was not as evident in previous pilot studies. Firstly, during interviews conducted as part of study 1, nurses frequently suggested that they felt a certain level of mistrust of the organisation based on a variety of factors such as previous negative experiences of reporting mistakes and a general lack of feedback overall. It was therefore essential to promote a high level of trust between the interviewer and nurses to achieve a candid and accurate perspective of error causation, particularly given the sensitive nature of the issues being discussed (e.g. making mistakes and organisational failings). This was achieved in part by placing a conceptual distance between the interviewer and the organisation as is commonly recommended in qualitative research of sensitive issues (Elwood & Martin, 2004). It was frequently emphasised

during interviews and observations that while the organisation was supporting the research and were keen to improve current error management approaches and systems, the researcher would be the only person with access to the information discussed. The information disclosed during interviews and observations was largely supported during subsequent pilot studies where similar efforts were made to emphasise the independent position of the researcher. Since it would not be possible to use face-to-face strategies of gaining participants' trust in the final study due the sample size, several recommended strategies for improving the candour of responses for 'sensitive' questionnaires were implemented (Edwards *et al.*, 2002; Jones & Linda, 1978). These efforts included adopting a 'neutral' stance by emphasising that the research was being conducted by researchers at the University of Leeds (Dillman, 2000; Sudman & Bradburn, 1974, 1982), assuring participants their performance was not being audited by their employer or judged in any way and emphasising their responses would be anonymous and confidential. However, as with any type of postal questionnaire, there are no guarantees that participants read the information provided before completing the instrument. It is also possible that these efforts were undermined to some extent by overt organisational endorsements of the research which included placement of hospital logos on all documentation and several emails sent by a senior NHS manager encouraging participants to complete the questionnaire. With such potentially sensitive items which involved disclosure of errors and reporting of latent failures including supervision and other management activities, it is possible that these overt organisational endorsements introduced some element of responding bias and reduced participant confidence that the questionnaire was not sent as part of a management audit. It is also possible that the response bias observed in the final study was not biased but a truthful and accurate reflection of the presence of latent failures and prevalence of errors which in turn might suggest that information obtained during interviews was not accurate.

It is unlikely that the response bias evident in study 4, which to some extent may have been reflective of a lack of organisational trust, is limited to the health care organisation in which the

research was conducted. Previous research has suggested that health care organisations comprise particularly 'closed' reporting cultures and that 'most errors are not reported' compared with other high-risk industries (Cullen *et al.,* 1995). Patient safety researchers have identified organisational trust as a key determinant of this reporting culture which in turn dictates subsequent organisational learning throughout the NHS (Firth-Cozens, 2003; Reason, 2000). Firth-Cozens (2004) suggests that when conducting this type of research, sharp-end staff need to feel they can trust that they will not be punished by their organisation for their honesty, and reassured that managers will treat disclosed information sensitively and fairly to benefit patient safety. She further argues that in most health care organisations, this level of trust is currently 'low' and proposes that improving this level of organisational trust should be 'high on the agenda of all health care systems' (pp. 56). It is proposed that in a reporting and learning 'standoff', lack of organisational trust throughout health care may have lead to a culture of reporting which is predominantly 'closed' and which has in turn lead to an inaccurate perception of error prevalence. Additionally, errors which are reported are less detailed and contain little information useful for determining causation. As supported by study 1 of this research, workers at the sharp-end of the workforce are not asked for the organisational failures which precede their reported mistakes, which may ultimately mean that reported errors are not being effectively managed, either respectively or proactively. In a cyclical manner, this could result in a further reduction in organisational trust. Indeed, nurses interviewed in study 1 of this research suggested that lack of feedback from previous incident reports played a key role in whether they would report their future mistakes.

It is proposed that this research has provided evidence which is vital for breaking this cycle. Health care organisations need to know more about latent failures for subsequent error management strategies to be effective. This research has made a significant contribution towards achieving this goal. Using evidence such as this, organisations can instigate more appropriately targeted interventions which can be observed by the workforce as making a difference to patient safety.

This could go some way towards improving organisational trust which might foster a more 'open' culture of reporting and should in turn lead to a significant improvement in error learning. Furthermore, the role of organisationally independent research conducted by outside agencies and universities should not be underestimated as a means of improving the level of trust between health care staff and their managers.

8.4. Implications & future research

Using rigorous qualitative methods, this research has made a significant contribution towards identifying latent failures in secondary care and examining their role in error causation. Although there was some empirical support for many of the variables identified, qualitative methods allowed for an in-depth examination of variables not currently highlighted within patient safety literature as important precursors of error. For instance, the latent factor *bed management* was commonly identified as an organisational factor likely to trigger multiple working conditions such as workload, planning and team communication yet is rarely acknowledged as a potential cause of medical errors within the literature. The qualitative findings of this research represent a sound platform on which to base future work which could examine the ways in which these variables interact with one another to lead to accidents and errors. These findings could be used to develop educational material with which to improve organisational learning both at the sharp-end and the blunt end of health care organisations which may in turn improve the overall culture of disclosing such sensitive information. Future research could also use the information to develop appropriate interventions for each of the latent failures similar to error management 'toolkits' available in other industries (HSE, 1995; Shell, 2004)

In line with Department of Health objectives for patient safety research (DoH, 2001), this research has used techniques employed in other high-risk industries to develop a theoretically-based

proactive measure of latent failures. The Organisational Safety Questionnaire differs from other currently available safety culture tools in that it was driven by health care professionals at both ends of the organisational spectrum. The fact that OSQ items were developed to represent tangible indicators of safety as opposed to workers perceptions, attitudes or beliefs also differentiates this tool from other available safety measures. Unlike other safety measures, the OSQ had good levels of concurrent and predictive validity. The 28-item OSQ is a novel and generalisable *proactive* measure of organisational safety not currently available in health care.

Poor responding variability which lead to the removal of multiple items prevented the construction of separate subscales representing each of the ten proposed latent error causes. This would have allowed for the development of safety profiles at a ward level (see chapter 2) to identify the 'worst' variables and which would be useful for the design of appropriate interventions. Nevertheless, OSQ items reflect a broad range of latent conditions which would make it a particularly effective outcome measure of patient safety intervention success and other organisational changes. The tool would also be an appropriate measure of monitoring continuous improvement.

There are several recommendations for future research. Firstly, conducting focus group studies such as those employed to develop the proactive Tripod-DELTA tool (Groeneweg, 1992; see chapter 2) would be an effective means of 'fine-tuning' indicators which represent each latent failure. This would address problems identified in this study with corresponding indicator time frames which may have been inaccurate. By disseminating the revised questionnaire to a larger sample than recruited in the final study of this research, it would be possible to examine the underlying structure of the OSQ as a measure of ten latent factors. This would then allow for a more detailed understanding of the relationship between each of the variables and the way in which this complex interaction might affect error likelihood. It would also be interesting to conduct focus groups with other health care professionals such as doctors or unqualified staff (e.g. nursing

assistants, ancillary staff etc.) to explore whether the latent conditions identified by nurses are also proposed by other disciplines. If these variables were generalisable to other health professions, it would then be possible to develop 'job-specific indicators' as currently recommended within other high-risk industries (Reason, 1997; Groeneweg, 2002). Finally, future research could also test the ability of the OSQ to predict other types of error (e.g. medical, diagnostic etc.) or other patient safety and quality of care measures such patient satisfaction, nosocomial infections rates or length of hospital stay and rate of recovery.

In conclusion, this research has successfully achieved its main objectives; to gain a clearer understanding of the organisational and workplace causes of medication administration errors and to develop a theoretically-based valid and reliable tool with which to measure them. This research has progressed patient safety research forwards to a more detailed level of understanding the causes of error and has developed a novel tool which will be useful in this field of research. It is anticipated that development of this novel tool will enable health care organisations to develop educational material to improve organisational learning and the culture of reporting and more successfully monitor the effects of organisational change on the workforce.

REFERENCES

Adams, A., Bond, S., & Arber, S. (1995). Development and validation of scales to measure organisational features of acute hospital wards. *International Journal of Nursing Studies,* 32 (6), 612-627.

Adams, S., & Vincent, C. (2004). Systems analysis of clinical incidents: the London protocol. Online: Available at: http://www.csru.org.uk.

Agency for Healthcare Research and Quality (2001). *Making Health Care Safer: A Critical Analysis of Patient Safety Practices.* Rockville, MD: AHRQ.

Aiken, L.H., Clarke, S.P., Sloane, D.M., *et al.* (2002). Hospital nurse staffing and patient mortality, nurse burnout and job dissatisfaction. *Journal of the American Medical Association,* 288, 1987-1993.

Allard, J., Carthey, J., Cope, J., *et al.* (2002). Medication errors: causes, prevention and reduction. *British Journal of Haematology,* 116, 255-265.

Altheide, D.L. (1996). *Qualitative media analysis.* Thousand Oaks: Sage.

American Academy of Pediatrics (AAP) Committee on Drugs and Committee on Hospital Care (1998). Prevention of medication errors in the pediatric inpatient setting. *Pediatrics,* 102, 428-430.

American Society of Consultant Pharmacists. (2003). *Guidelines on preventing medication errors in pharmacies through reporting and evaluation.* Accessed at http://ascp.com/resources/policy/upload/Gui97-Med%20Errors.pdf on 1 Sept 2008.

Anderson, D., & Webster, C.S. (2001). A systems approach to the reduction of medication error on the hospital ward. *Journal of advanced nursing,* 35 (1), 34-41.

Anderson, F., Maloney, J., Knight, C., & Jennings, B. (1996). Utilization of supplemental agency nurses in an Army medical centre. *Military Medicine,* 161 (1), 48-53.

Anderson, L. (1997). The introduction of generic workers into the ward team: An exploratory study. *Journal of Nursing Management,* 5 (2), 69-75.

Aneja, S., Bajaj, G., & Mehandiratta, S.K. (1992). Errors in medication in a pediatric ward. *Indian Journal of Pediatrics*, 29, 727-30.

Appleton, J. V., & King, L. (1997). Constructivism: A Naturalistic Methodology for Nursing Inquiry. *Advances in Nursing Science*, 20 (2), 13-22.

Armitage, G. (2004). Using ethnography (or qualitative methods) to investigate drug errors: A critique of a published study. *Nursing Times Research,* 9 (5), 379-387.

Australian Council for Safety and Quality in Health Care (2002). *Second National Report on Patient Safety—Improving Medication Safety.* Australian Council for Safety and Quality in Health: Canberra, Australia.

Awad, S., Fagan, C., Bellows, D., *et al.* (2005). Bridging the communication gap in the operating room with medical team training. *The American Journal of Surgery*, 190 (5), 770 - 774 .

Bain, L. (1996). Preceptorship: A review of the literature. *Journal of Advanced Nursing*, 24 (1), 104-107.

Bainbridge, D.J., Wise, J.A., & Hopkin, D. V. (1999). *Handbook of aviation factors* (Eds.). Mawah, New Jersey: Lawrence Erlbaum Associates.

Baker, R. G., Norton, P.G., Flintoft, V., *et al.* (2004). The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *Canadian Medical Association Journal,* 170 (11), 1678-1686.

Banister, G., Butt, L., & Hackel, R. (1996). How nurses perceive medication errors. *Nursing Management,* 27 (1), 31-34.

Barber, N. (2002). Should we consider non-compliance a medical error? *Quality and Safety in Health Care,* 11 (1), 81-84.

Barker, K.N., Pearson, R.E., Hepler, C.D., *et al.* (1984). Effect of an automated bedside dispensing machine on medication errors. *American Journal of Hospital Pharmacy,* 41, 1352-1358.

Barrows, H. S. (2000). *Problem-Based Learning Applied to Medical Education.* Springfield: Southern Illinois University School of Medicine.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge, UK: Cambridge University Press.

Bates, D.W., Cullen, D.J., Laird, N., *et al.* (1995). Incidence of adverse drug events and potential drug events. *Journal of the American Medical Associaion*, 274, 29-34.

Battles, J.B. (2006). Improving patient safety by instructional systems design. *Quality and Safety in Health Care,* 15, i25-i29.

Battles, J.B., Kaplan, H.S., Van der Schaaf, T.W., & Schea, C. (1998). The attributes of medical event-reporting systems: experience with a prototype medical event-reporting system for transfusion medicine. *Archives of Pathological Laboratory Medicine,* 122 (3), 231-8.

Beaudoin, L.E., & Edgar, L. (2003). Hassles: Their importance to nurses' quality of work life. *Nursing economics,* 21 (3), 106-113.

Bedell, S.E., Deitz, D.C., Leeman, D., et al. (1991). Incidence and characteristics of preventable iatrogenic cardiac arrests. *Journal of the American Medical Association,* 277, 307-311.

Berelson, B. (1952). *Content analysis in communication research.* Glencoe, Illinois: Free Press.

Berg, B.L. (2001). *Qualitative Research Methods for the Social Sciences*. Needham Heights, Mass.: Allyn and Bacon.

Berkson, L. (1993). Problem-based learning: have the expectations been met? *Academic Medicine*, 68 (10), S79-S88.

Blegen, M.A., Goode, C.J., & Reed, L. (1998). Nurse staffing and patient outcomes. *Nursing research,* 47, 43-50.

Blendon, R.J., DesRoches, C.M., Brodie, M., *et al.* (2002). Views of Practicing Physicians and the Public on Medical Errors. *The New England Journal of Medicine,* 347, 1933-1940.

Blum, K.V., Abel, S.R., Urbanski, C.J., & Pierce, J.M. (1988). Medication error prevention by pharmacists. *American Journal of Hospital Pharmacy*, 45, 1902-3.

Bogdewic, S.P. (1992). Participant observation. In B.F. Crabtree & W.L. Miller (Eds.), *Doing qualitative research: Research methods for primary care.* Newbury Park, CA: Sage.

Bognor, M.S. (Ed.) (1994). *Human error in medicine.* Hove, UK: LEA Publishers.

Bola, T.V., Driggers, K., Dunlap, C., & Ebersole, M. (2003). Foreign educated nurses: strangers in a strange land. *Nursing Management,* 34 (7), 39-42.

Bordun, L.A., & Butt, W. (1992). Drug errors in intensive care. *Journal of Paediatric Child Health,* 28, 309-11.

Bottomley, A. (1997). To randomise or not to randomise: Methodological pitfalls of the RCT design in psychosocial intervention studies. *European Journal of Cancer Care,* 6 (3), 222-230.

Bourque, L.B., & Fielder, E.P. (1995). How to conduct self administered and mail surveys. Thousand Oaks, CA: Sage.

Bowler, S., & Mallik, M. (1998). Role extension or expansion: a qualitative investigation of the perceptions of senior medical and nursing staff in an adult intensive care unit. *Intensive and critical care nursing,* 14, 11-20.

Boxwala, A., Dierks, M., Keenan, M., *et al.* (2004). Organization and Representation of Patient Safety Data: Current Status and Issues around Generalizability and Scalability. *Journal of American Medical Informatics Association,* 11, 468-478.

Boyatzis, R.E. (1998). *Transforming qualitative information: Thematic analysis and code development.* London: Sage.

Brace, N., Kemp, R., & Smelgar, R. (2003). *SPSS for psychologists.* New York: Palgrave.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3, 77-101.

Brennan, T.A., Leape, L.L., Laird, N.M., *et al.* (1991). Incidence of adverse events and negligence in hospitalized patients; results from the Harvard Medical Practice Study I. *New England Journal of Medicine,* 324 (6), 370-376.

Broadbent, D.E. (1958). *Perception and Communication.* London, UK: Pergammon.

Brown, C., Hofer, T., Johal, A., *et al.* (2008). An epistemology of patient safety research: a framework for study design and interpretation. Part 3. End points and measurement. *Quality and Safety in Health Care,* 17, 170-177.

Bryan, L.A. (1989). The human factor: Implications for engineers and managers. *Professional Safety,* 34, 15-18.

Bryman, A. (2001). *Social research methods.* New York: Oxford University Press.

Buchan, J. (2002). *International recruitment of nurses: UK case study.* London, UK: World Health Organization.

Burgess, R. (1984). *In the field: an introduction to field research.* London, UK: George Allen & Unwin.

Busse, D.K., & Wright, D.J. (2000). Classification and analysis of incidents in complex medical environments. *Topics in Health Information Management,*20 (4), 1-11.

Butler, R. (1998). Determinants of help seeking: Relations between perceived reasons for classroom help-avoidance and help-seeking behaviors in an experimental context. *Journal of Educational Psychology,* 90, 630–644.

Butler, R., & Neuman, O. (1995). Effects of task and ego achievement goals on help-seeking behaviors and attitudes. *Journal of Educational Psychology,* 87, 261–267.

Calbrese, A.D., Erstad, B.L., Brandl, K., *et al.* (2001). Medication administration errors in adult patients in the ICU. *Intensive Care Medicine,* 27, 1592-1598.

Carothers, N.B. (1999). An analysis of medication errors: Look alike and sound alike names. *International Journal of Trauma Nursing,* 5 (1), 22-24.

Carthey, J., de Leval, M.R., & Reason, J.T. (2001a). Institutional resilience in healthcare systems. *Quality and Safety in Health Care,* 10, 29-32.

Carthey, J., de Leval, M.R., & Reason, J.T. (2001b). The human factor in cardiac surgery: Errors and near misses in a high technology medical domain. *Annals of Thoracic Surgery,* 72, 300-305.

Cattell, R.B. (1978). *The scientific use of factor analysis.* New York: Plenum.

Centre for Reviews and Dissemination. (2001). Undertaking *systematic reviews of research on effectiveness: CRD's guidelines for those carrying out or commissioning reviews.* University of York: CRD.

Chang, A., Schyve, P., Croteau, R.J., *et al.* (2005). The JCAHO patient safety event taxonomy: A standardized terminology and classification schema for near misses and adverse events. *International Journal for Quality in Health Care,* 17 (2), 95-105.

Clark-Canter, D. (2004). *Quantitative psychological research: A student's handbook.* Hove: Psychology Press.

Clark, L.A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment,* 7 (3), 309-319.

Classen, D.C. & Metzger, J. (2003). Improving medication safety: the measurement conundrum and where to start. *International Journal for Quality in Health Care,* 15, 41-47.

Classen, D.C., Pestonik, S.L., Evans, R.S., & Burke, J.P. (1991). Computerized surveillance of adverse drug events in hospital patients. *Journal of the American Medical Association, 266* (20), 2847-2851.

Cohen, M. M., Kimmel, N. L., Benage, M. K., *et al.* (2005). Medication safety program reduces adverse drug events in a community hospital. *Quality and Safety in Health Care,* 14 (3), 169-174.

Cohen, M.R. (2000). Why error reporting systems should be voluntary. *British Medical Journal,* 320 (7237), 728-729.

Cohen, M.R., Senders, J., & Davis, N.M. (1994). Failure mode and effects analysis: A model approach to avoiding dangerous medication errors and accidents. *Hospital Pharmacy,* 29, 319-328.

Colla, J.B., Bracken, A.C., Kinney, L.M., & Weeks, W.B. (2005). Measuring patient safety climate: A review of surveys. *Quality and Safety in Health Care,* 14, 364-366.

Combs, A., & Taylor, C. (1952). The effect of the perception of mild degrees of threat on performance. *Journal of Abnormal Social Psychology,* 47, 420-4.

Conklin, D., MacFarland, V., Kinnie-Steeves, A., & Chenger, P. (1990). Medication errors by nurses: contributing factors. *AARN Newsletter*, 46 (1), 8-9.

Cook, R., Render, M., & Woods, D. (2000). Gaps in the continuity of care and progress on patient safety. *British Medical Journal,* 320 (7237), 791-794.

Cooper, M.C. (1995). Can a zero defects philosophy be applied to drug errors? *Journal of Advanced Nursing,* 21 (3), 487-491.

Cornish, P.L., Knowles, S.R., Marchesano, R., *et al.* (2005). Unintended Medication Discrepancies at the Time of Hospital Admission. *Archives of Internal Medicine,* 165, 424-429.

Costello, A.B., & Osbourne, J.W. (2005). Best practices in Exploratory Factor Analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation,* 10 (7).

Costello, I., Wong, I., Long, P.F., *et al.* (2007). *Paediatric drug handling.* London: Pharmaceutical Press.

Cox, S., & Flin, R. (1998). Safety culture. Philosophers stone or man of straw. *Work and stress,* 12, 189-201.

Coxon, J.P., Pattison, S.H., Parks, J.W. *et al.* (2003). Reducing human error in urology: Lessons from aviation. *British Journal of Urology,* 91, 1-8.

Cresswell, J.W. (1994). *Research design.* Thousand Oaks, CA: Sage.

Cresswell, J.W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches*. London, UK: Sage.

Croskerry, P. (2002). Achieving quality in clinical decision making: Cognitive strategies and detection of bias. *Academic Emergency Medicine,* 9, 1184-1204.

Cullen, D.J., Bates, D., & Leape, L.L. (2000). Prevention of adverse drug events: a decade of progress in patient safety. *Journal of Clinical Anesthesia,* 12, 600-614.

Cullen, D.J., Bates, D., Small, S., *et al.* (1995). The incident reporting system does not detect adverse drug events. *Joint Commission Journal for Quality Improvement,* 21, 541-548.

Cullen, Rt. Hon Lord. (2000a). *The Ladbroke Grove Rail Enquiry: Volume 1.* London, UK: Health and Safety Commission.

Cullen, Rt. Hon Lord. (2000b). *The Ladbroke Grove Rail Enquiry: Volume 2.* London, UK: Health and Safety Commission.

Davis, N.M. (1994). Concentrating on interruptions. *American Journal of Nursing,* 94 (3), 1414.

Davis, P. (2004). Health care as a risk factor. *Canadian Medical Association Journal,* 170, 1688 – 1689.

Davis, P., Lay-Yee, R., Briant, R., *et al.* (2002). Adverse events in New Zealand public hospitals I: occurrence and impact. *New Zealand Medical Journal,* 115 (1167), U271.

Davis, P., Lay-Yee, R., Briant, R., *et al.* (2003). Adverse events in New Zealand public hospitals II: preventability and clinical context. *New Zealand Medical Journal,* 116 (1183), U624.

Dean, B., Barber, N., & Schacter, M. (2000). What is a prescribing error? *Quality and Safety in Health Care,* 9, 232-237.

Dean, B., Schachter, M., Vincent, C., *et al.* (2002). Causes of prescribing errors in hospital inpatients: a prospective study. *The Lancet,* 359, 1373-1378.

Dean, G.E., Scott, L.D., & Rogers, A. (2006). Infants at risk: when nurse fatigue jeopardizes quality care. *Advances in Neonatal Care,* 6 (3), 120-126.

Delamont, S. (2002). *Fieldwork in Educational Settings: Methods, pitfalls and perspectives.* London: The Falmer Press.

Denzin, N.K. (1970). *The research act: A theoretical introduction to sociological methods.* Chicago: Aldine.

Department of Health (2000). *An organisation with a memory: Report of an expert group on learning from adverse events in the NHS.* London, UK: The Stationery Office.

Department of Health (2001). *Building a safer NHS for patients: Implementing an organisation with a memory.* London, UK: The Stationery Office.

Dillman, D.A. (2000). Mail and internet surveys: the tailored design method. 2nd ed. New York: Wiley.

Dixon-Woods, M. (2003). What can ethnography do for quality and safety in health care? *Quality and Safety in Health Care,* 12, 326-327.

Dodd, D.K., & Markwiese, B.J. (1987). Survey response rate as a function of personalized signature on cover letter. *Journal of Social Psychology,* 127, 97–8.

Donchin, Y., Gopher, D., Olin, M., *et al.* (2003). A look into the nature and causes of human errors in the intensive care unit. *Quality and Safety in Health Care,* 12, 143-148.

Easton, K.L., Parson, B.J., Starr, M., & Brien, J.E. (1998). The incidence of drug-related problems as a cause of hospital admissions in children. *Medical Journal of Australia,* 169, 356–359.

Eccles, M., Grimshaw, J., Campbell, M., & Ramsay, C. (2003). Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care,* 12, 47-52.

Eccles, M., McColl, E., Steen, N., *et al.* (2002). Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *British Medical Journal,* 325, 941-948.

Edmondson, A.C. (1996). Learning from mistakes is easier said than done: group and organisational influences on the detection and correction of human error. *Journal of Applied Behavioural Science,* 32 (1), 5-28.

Edwards, P., Roberts, I., Clarke, M., *et al.* (2002). Increasing response rates to postal questionnaires: systematic review. *British Medical Journal,* 324, 1183.

Elwood, S.A., & Martin, D.G. (2004). "Placing" interviews: Location and scales of power in qualitative research. *The Professional Geographer,* 52 (4), 649-657.

Evans, R.S., Pestonik, S.L., Classen, D.C., *et al.* (1992). Prevention of adverse drug events through computerized surveillance. *Proceedings of the Annual Symposium of Computer Applications in Medical Care,* 437-441.

Evans, R.S., Pestotnik, S.L., Classen, D.C. *et al.* (1998). A computer-assisted management program for antibiotics and other anti-infective agents. *New England Journal of Medicine*, 338, 232-8.

Evans, S. M., Berry, J. G., Smith, B. J., *et al.* (2006). Attitudes and barriers to incident reporting: a collaborative hospital study. *Quality and Safety in Health Care,* 15 (1), 39-43.

Fallowfield, R. (1995). Psychosocial interventions in cancer. *British Medical Journal,* 311, 1316-1317.

Feder, G., Eccles, M., Grol, R., *et al.* (1999). Using clinical guidelines. *British Medical Journal,* 318, 728-730.

Field, A. (2005*). Discovering statistics using SPSS.* London: Sage Publications.

Fielding, N.G., & Fielding, J.L. (1986). *Linking Data: The Articulation of Qualitative and Quantitative Methods in Social Research.* Beverly Hills, CA: Sage Publishing.

Firth-Cozens, J. (2003). Learning from error. In: Harrison J, Innes R, Van Zwanenberg T, eds. *Regaining trust in health care*. Abingdon: Radcliffe Medical Press.

Firth-Cozens, J. (2004). Organisational Trust: the keystone to patient safety. *Quality and Safety in Health Care,* 13, 56-61.

Firth-Cozens, J., & Mowbray, D. (2001). Leadership and quality of care. *Quality in Health Care,* 10, 3-7.

Fiske, S.T., & Taylor, S.E. (1978). *Social Cognition.* Reading, MA: Addison-Wesley.

Flin, R., Mearns, K., O'Connor, P., & Bryden, P. (2000). Measuring safety climate: Identifying the common features. *Safety Science,* 34, 177 – 192.

Flin, R., & Yule, S. (2004). Leadership for safety: industrial experience. *Quality and Safety in Health Care,* 13, 45-51.

Flynn, E.A., Barker, K.N., Pepper, G.A., *et al.* (2002). Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. *American Journal of Health-System Pharmacy*, 59 (5), 436-446 .

Folli, H.L., Poole, R.L., Benitz, W.E., & Russo, J.C. (1987). Medication error prevention by clinical pharmacists in two children's hospitals. *Pediatrics*, 79, 718-22.

Forrey, R.A., Pederson, C.A., & Schnieder, P.A. (2007). Inter-rater agreement with a standard scheme for classifying medication errors. *American Journal of Health-System Pharmacy,*64, 175-81.

Fortescue, E.B., Kaushal, R., Landrigan, C.P., *et al.* (2003). Prioritizing strategies for preventing medication errors and adverse drug events in paediatric inpatients. *Pediatrics,* 111, 722-9.

Foster, R.L. (1997). Addressing epistemologic and practical issues in multi-method research: a procedure for conceptual triangulation. *Advances in Nursing Science,* 20 (2), 1-12.

Funk, K. (1991). Cockpit Task Management: Preliminary definitions, normative theory, error taxonomy and design recommendations. *International Journal of Aviation Psychology,* 1 (4), 271-285.

Gastmans, C. (1999). Care as a moral attitude in nursing. *Nursing Ethics,* 6, 214-223.

Gibson, J.M. (1998). Using the Delphi technique to identify the content and context of nurses' continuing professional development needs. *Journal of Clinical Nursing,* 7(5), 451-459.

Gibson, W.H., Megaw, E.D., Young, M.S., & Lowe, E. (2006). A taxonomy of human communication errors and application to railway track maintenance. *Cognition, Technology and Work,* 8 (1), 57-66.

Giles, W.F., & Field, H.S. (2006). Effects of amount, format and location of demographic information on questionnaire return rate and response bias of sensitive and non-sensitive items. *Personnel Psychology,* 31 (3), 549-559.

Gillespie, R. (1991). *Manufacturing knowledge: a history of the Hawthorne experiments*. Cambridge: Cambridge University Press.

Gladstone, J. (1995). Drug administration errors: A study into the factors underlying the occurrence and reporting of drug errors in a district general hospital. *Journal of Advanced Nursing,* 22 (4), 628-637.

Glaser, B. (1992). *Basics of grounded theory analysis.* CA, USA: Sociology Press.

Gold, D.R., Rogacz, S., Bock, N., *et al.* (1992). Rotating shift work, sleep, and accidents related to sleepiness in hospital nurses. *American Journal of Public Health,* 82, 1011-1014.

Gordon, R., Mearns, K., Flin, R., *et al.* (2000). *Factoring the Human into Safety: Translating Research into Practice.* University of Aberdeen: HSE OTO report (Vol 2).

Gorman, R.L., Bates, B.A., Bereitz, W.E., *et al.* (2003). Prevention of medication errors in the pediatric inpatient setting. *Pediatrics,* 112, 431-40.

Gott, M., Seymour, J., Bellamy, G *et al.* (2004). Older people's views about home as a place of care at the end of life. *Palliative Medicine,* 18, 460-467.

Gould, D. (1996). Using vignettes to collect data for nursing research studies: how valid are the findings? *Journal of Clinical Nursing,* 5 (4):207-12.

Gray, J.H., & Densten, I.L. (1998). Integrating Quantitative and Qualitative Analysis Using Latent and Manifest Variables. *Quality & Quantity,* 32, 419–431.

Groeneweg, J. (1992). *Controlling the controllable: The management of safety (1st Ed.).* Leiden, The Netherlands: DSWO Press.

Groeneweg, J. (2000). *Preventing human error using audits.* London, UK: IIR Conference 'Effective Safety Auditing`.

Groeneweg, J. (2002). *Controlling the controllable: The management of safety (5th Ed.).* Leiden, The Netherlands: DSWO Press.

Groeneweg, J., Hudson, P.T.W., & Wagenaar, W.A. (1989). *Towards a complete list of general failure types.* Interim report prepared for Shell Internationale Petroleum Maatschappij, Exploration and Production.

Groeneweg, J., & Roggeveen, V. (1998). *Tripod: Safety performance redefined.* Proceedings of the 1998 SPE Health and Safety Conference, Caracas.

Gully, S.M., Devine, D.J., & Whitney, D.J. (1995). A meta-analysis of cohesion and performance: Effects of levels of analysis and task interdependence. *Small Group Research,* 26, 497-520.

Guzzo, R.A. & Dickson, M.W. (1996). Teams in organisations: Recent research on performance and effectiveness. *Annual Review of Psychology,* 47, 307-338.

Hackman, J.R. (1983). *A normative model of work team effectiveness* (Technical report NO. 2). Yale School of Organization and Management, Research Program on Group Effectiveness,

Harris, C.B., Krauss, M.J., Coopersmith, C.M., *et al.* (2007). Patient safety event reporting in critical care: a study of three intensive care units. *Critical Care Medicine*, 35 (4), 1068-76.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *The American Psychologist,* 39 (12), 1372-1388.

Hautala, K.T., Saylor, C.R., & O'Leary-Kelley, C. (2007). Nurses' perceptions of stress and support in the preceptor role. *Journal for Nurses in Staff Development,* 23 (2), 64-70.

Hawkins, F.H. (1987). *Human factors in flight* (2nd Ed) Aldershot, UK: Ashgate Publishing Company.

Hawkins, M.R., Vichick, D.A., Silsby, H.D., *et al.* (1985). Sleep and nutritional deprivation and performance of house officers. *Journal of Medical Education,* 60, 530-535.

Healey, A.N., Undre, S., & Vincent, C.A. (2004). Developing observational measures of performance in surgical teams. *Quality and Safety in Health Care,* 13, i33-i40.

Health and Safety Executive. (1995). *Improving compliance with safety procedures: Reducing industrial installations.* Sudbury, UK: HSE books.

Healthcare Information and Management Systems Society (2002). *Hot Topic Survey.* Accessed at http://www.himss.org/content/files/surveyresults/hottopicssurvey.pdf on 08 August 2008.

Helmreich, R.L. (2000). On error management: lessons from aviation. *British Medical Journal,* 320, 781-85.

Helmreich, R.L., & Schaefer, J.M. Team performance in the OR. In: Bognor, M.S. (Eds.) (1994). *Human error in medicine.* Hillside, NJ: Erlbaum.

Helmreich, R.L., & Merritt, A. C. (1998). *Culture at work in aviation and medicine.* Aldershot: Ashgate.

Hendrich, A., & Lee, N. (2003). *Evidence-based design on nursing workspace in hospitals.* Report commissioned by the IOM committee on the work environment for nurses and patient safety.

Henriksen, K., & Kaplan, H. (2003). Hindsight bias, outcome knowledge and adaptive learning. *Quality and Safety in Health Care,* 12, 46 – 50.

Herout, P.M., & Erstad, B.L. (2004). Medication errors involving continuously infused medications in a surgical intensive care unit. *Critical Care Medicine*, 32, 428-32.

Hicks, L. (1997). Academic motivation and peer relationships-how do they mix in an adolescent's world? *Middle School Journal,* 28, 18–22.

Hoff, T., Jameson, L., Hannan, E., *et al.* (2004). A review of the literature examining linkages between organisational factors, medical errors and patient safety. *Medical Care and Research Review,* 61 (1), 3-37.

Howanitz, P.J., Renner, S.W., & Walsh, M.K. (2002). Continuous Wristband Monitoring Over 2 Years Decreases Identification Errors. *Archives of Pathology and Laboratory Medicine*, 126, 809–815.

Howell, M. (1996). Prescription for disaster. *Nursing Times,* 92 (34), 30-31.

HSE. (2002). *The Rail Public Inquiries HSC Report on Overall Progress as of July 2002 on all Recommendations from the Rail Public Inquiries.* Norwich, UK: HSE books.

Hsieh, H.F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15* (9), 1277-1288.

Hudson, P. (2005). *Personal email communication dated 02/03/05.*

Hudson, P., Reason, J., Wagenaar, P., *et al.,* (1994). Tripod-Delta: proactive approach to enhanced safety. *Journal of Petroleum Technology,* 40, 58-62.

Hughes, R. (1998). Considering the vignette technique and its application to drug injecting, HIV risk and safer behaviour. *Sociology of Health and Illness*, 20, 381-87.

Hurst, N.W., & Radcliffe, K. (1994). *Development and application of a structured audit technique for the assessment of safety management systems (STATAS). Hazards XII. European advances in process safety.* Rugby: Institute of Chemical Engineers.

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalised linear models.* Thousand Oaks, CA: Sage Publications.

Hutchinson, A., Cooper, K.L., Dean, J.E., *et al.* (2006). Use of a safety climate questionnaire in UK health care: factor structure, reliability and usability. *Quality and Safety in Health Care,* 15, 347-353.

Institute of Medicine. (1999). To *err is human: Building a safer health system.* Washington, DC: National Academy Press.

Institute of Medicine. (2004). *Keeping patients safe: transforming the work environment of nurses.* Washington DC, USA: The National Academies Press.

Ioannidis, J.P.A., & Lau, J. (2001). Evidence on interventions to reduce medical errors: An overview and recommendations for future research. *Journal of General Internal Medicine,* 16, 325-334.

Jastrow, J. (1905). The lapses of consciousness. *The Popular Science Monthly,* 67, 481-502.

Jeffe, D.B., Dunagan, W.C., Garbutt, J., *et al.* (2004). Using focus groups to understand physicians' and nurses' perspectives on error reporting in hospitals. *Joint Commission Journal of Quality and Safety,* 30 (9), 471-479.

Jick, T.D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly,* 24, 602-611.

Jinks, A.M., & Hope, P. (2000). What do nurses do? An observational survey of the activities of nurses on acute surgical and rehabilitation wards. *Journal of Nursing Management,* 8, 273-279.

Johns, G. (1994). How often were you absent—A review of the use of self-reported absence data. *Journal of Applied Psychology,* 79, 574–591.

Jones, B. (2002). Nurses and the code of silence. In *Medical Error*, (Rosenthal & Sutcliffe eds.), San Fransisco: Jossey-Bass.

Jones, S.G., & Tesmer, B. A new tool for investigating and tracking human factors issues in incidents. University of Texas Crew Research Project Website (www.psy.utexas.edu/psy/helmreich/nassau.htm).

Jones, W.H., & Linda, G. (1978). Multiple criteria effects in a mail survey experiment. *Journal of Marketing Research,* 15, 280–4.

Jonsdottir, M.K., Adolfsdottir, S., Dogg, R., *et al.* (2007). A diary study of action slips in healthy individuals. *The Clinical Neuropsychologist,* 21, 875-883.

Kahneman, D. (1973) *Attention and Effort.* Englewood Cliffs, N.J.: Prentice-Hall.

Kanjanarat, P., Winterstein, A.G., Johns, T.E. *et al.* (2003). Nature of preventable adverse drug events in hospitals: a literature review. *American Journal of Health-System Pharmacy*, 60, 1750-1759.

Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other achievement behavior in college students. *Journal of Educational Psychology,* 83, 221–230.

Kass, R.A., & Tinsley, H.E.A. (1979). Factor analysis. *Journal of Leisure Research,* 11, 120-138.

Kaushal, R., Bates, D.W., Landrigan, C., *et al.* (2001). Medication Errors and Adverse Drug Events in Pediatric Inpatients. *Journal of the American Medical Association,* 285, 2114-20.

Kennedy, W. (2004). Beneficence and autonomy in nursing. *British Journal of Perioperative Nursing,* 14 (11), 500-6.

Kim, J., & Mueller, C.W. (1985). *Factor analysis: Statistical methods and practical issues*. Newbury Park, CA: Sage.

Kimchi, J., Polivka, B., & Stevenson, J. S. (1991). Triangulation: Operational definitions. *Nursing Research,* 40 (6), 364-366.

Kline, P. (2002). *An easy guide to factor analysis*. London: Routledge.

Kline, T. J. B., Sulsky, L. M., & Rever-Moriyama, S. D. (2000). Common method variance and specification errors: A practical approach to detection. *The Journal of Psychology,* 134, 401–421.

Knafl, K.A., & Breitmayer, B.J. (1989). Triangulation in qualitative research: issues of conceptual clarity and purpose. In *Qualitative Nursing Research* (Morse J.M. ed.), Aspen, Rockville, Maryland, pp. 209-220.

Knaus, W.A.E., Draper, A., Wagner, D.P., *et al.* (1986). An evaluation of outcome from intensive care in major medical centres. *Annals of Internal Medicine,* 104, 410-418.

Knowles, W.B. (1963). Operator loading tasks. *Human Factors,* 5, 151-161.

Kohn, L.T., Corrigan, J.M., & Donaldson, M.S. (Eds.) (2000). *To err is human: building a safer health system*. Washington, DC: National Academy Press.

Koppel, R., Metlay, J.P., Cohen, A., *et al.,* (2005). Role of computerized physician order entry systems in facilitating medication errors. *Journal of the American Medical Association,* 293 (10), 1197-1203.

Kozer, E., Scolnik, D., Macpherson, A., *et al.* (2002). Variables associated with medication errors in pediatric emergency medicine. *Pediatrics*, 110, 737-42.

Kracquer, S. (1953). The challenge to qualitative content analysis. *Public Opinion Quarterly*, 16 631-642.

Kremsdorf, R. (2005). CPOE: not the first step towards patient safety. *Health Management Technology,* 26 (1), 65-66.

Kuperman, G.J & Gibson, R.F. (2003). Computer Physician Order Entry: Benefits, Costs, and Issues. *Improving Patient Care*, 139 (1), 31-39.

Landsberger, H.A. (1958). *Hawthorne revisited*. New York: Cornell University Press.

Langley, G., Nolan, K.M. & Nolan, T.W. (1992). *The Foundation of Improvement.* Silver Spring, MD: API Publishing.

Lankshear, A.J., Sheldon, T.A., & Maynard, A. (2005). Nurse staffing and health care outcomes: A systematic review of the literature. *Advances in Nursing Science,* 28 (2), 163-174.

Larmour, I., Dolphin, R.G., Baxter, H., *et al.* (1991). A prospective study of hospital admissions due to drug reactions. *Australian Journal of Hospital Pharmacy,* 21, 90–95.

Laschinger, H.K.S., Finegan, J., Sharmian, J., *et al.* (2000). Organisational trust and empowerment in restructured healthcare settings: Effects on staff nurse commitment. *Journal of Nursing Administration,* 30 (9), 413-425.

Latane, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology,* 37 (6), 822-832.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28 (4), 563-575.

Lawton, R. (1998). Not working to rule: Understanding procedural violations at work. *Safety Science,* 28 (2), 77-95.

Lawton, R., & Parker, D. (2002). Barriers to incident reporting in a health care system. *Quality and Safety in Health Care,* 11, 15-18.

Lawton, R., & Parker, D. (2002). Judgments of the rule-related behaviour of health care professionals: An experimental study. *British Journal of Health Psychology,* 7 (3), 253-265.

Lawton, R., & Ward, N.J. (2004). A systems analysis of the Ladbroke grove rail crash. *Accident Analysis and Prevention,* 37, 235-244.

Leape, L.L. (1997). A systems analysis approach to medical error. *Journal of Evaluation in Clinical Practice,* 3 (3), 213-222.

Leape, L.L. (2000). Institute of medicine medical error figures are not exaggerated. *Journal of the American Medical Association,* 284, 95-97.

Leape, L.L., Bates, D.W., & Cullen, D.J. (1995). Systems analysis of adverse drug events. *Journal of the American Medical Association*, 274, 35-43.

Leape, L.L., Brennan, T.A., Laird, N., *et al.* (1991). The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. *New England Journal of Medicine*, 324 (6), 377-384.

Leape, L.L., Kabcenell, A.I., Gandhi, T.K., *et al.* (2000). Reducing adverse events: Lessons from a breakthrough series collaborative. *Joint Commission Journal on Quality and Patient Safety,* 26 (6), 321-331.

Lilford, R.J., Mohammed, M.A., Braunholtz, D., & Hofer, T.P. (2003). The measurement of active errors: methodological issues. *Quality and Safety in Health Care,* 12, 8.

Lingard, L., Espin, S., Rubin, B., *et al.* (2005). Getting teams to talk: development and pilot implementation of a checklist to promote interprofessional communication in the OR. *Quality and Safety in Health Care,* 14, 340-346.

Lingard, L., Espin, S., Whyte, S., *et al.* (2004). Communication failures in the operating room: an observational classification of recurrent types and effects. *Quality and Safety in Health Care,* 13, 330-334.

Lisby, M., Nielsen, L.P., & Mainz, J. (2005). Errors in the medication process: frequency, type and potential clinical consequence. *International Journal for Quality in Health Care,* 17 (1), 15-22.

Lowden, A., Kecklund, G., Axelsson, J., *et al.* (1998). Change from an 8-hour shift to a 12-hour shift, attitudes, sleep, sleepiness and performance. *Scandinavian Journal of Work, Environment and Health,* 24 (3), 69-75.

Lundgren, S., & Segesten, K. (2001). Nurses' use of time in a medical-surgical ward with all RN-staffing. *Journal of Nursing Management,* 9 (1), 13-20.

Mackay, L. (1993). *Conflicts in Care Medicine and Nursing,* London: Chapman & Hall.

Malpass, A., Helps, S.C., & Runciman, W.B. (1999). An analysis of Australian adverse drug events. *Journal of Quality in Clinical Practice,* 19, 27-30.

Malyn-Haw, C., Dickens, G., & Stubbs, J. (2005). A review of medication administration errors reported in a large psychiatric hospital in the United Kingdom. *Psychiatric Services,* 56, 1610-1613.

Mandler, G. (1982). Stress and thought processes. In *Handbook of stress,* Goldberger & Breznitz (Eds) , pp. 88-164, New York: Free Press.

Marck, P.B., Kwan, J.A., Preville, B., *et al.* (2006). Building safer systems by ecological design: using restoration science to develop a medication safety intervention. *Quality and Safety in Health Care,* 15, 92-97.

Marina, B.L., Reinhardt, K., Eichelberger, W.J., *et al.* (2000). Prevalence of errors in a pediatric hospital medication system: implications for error proofing. *Outcomes Management for Nursing Practice,* 4, 129-35.

Marshman, J.A., David, K.U., Lam, R.W.K. & Hyland, S. (2006). Medication error events in Ontario acute care hospitals. *Canadian Journal of Hospital Pharmacy,* 59(5), 243-250.

Mayo, E. (1949). *Hawthorne and the western electric company, the social problems of an industrial civilization.* Routledge.

Mayring, P. (2000). Qualitative content analysis. *Qualitative social research,* 1 (2), Retrieved July 9, 2008, from http://www.qualitative-research.net.

McFadden, K.L., Towell, E.R., & Stock, G.N. (2004). Critical success factors for controlling and managing hospital errors. *Quality Management Journal,* 11 (1), 61–74.

Mearns, K., Flin, R., O'Connor, P. (2001). Sharing 'worlds of risk': improving communication with crew resource management. *Journal of Risk Research,*4 (4), 377-392.

Meringer, R., & Mayer, C. (1895). *Versprechen und Verlesen: Eine psychologische-linguisticische Studie.* Stuttgart: G.J. Goschen'sche Verlagshandlung.

Meurier, C.E. (2000). Understanding the nature of errors in nursing: Using a model to analyse critical incident reports of errors which had resulted in an adverse or potentially adverse event. *Journal of Advanced Nursing,* 32 (1), 202-207.

Miles, M. & Huberman, A. (1994). *Qualitative data analysis (2nd ed.).* Thousand Oaks: Sage.

Miller, D.P., & Swain, A.D. (1987). Human error and human reliability. In G. Salvendy (Ed.), *Handbook of human factors* (p. 219-250). New-York: John Wiley & Sons.

Miller, L.A. (2005). Patient safety ad teamwork in perinatal care: resources for clinicians. *Journal of Perinatal and Neonatal Nursing,* 19 (1), 46-51.

Mishra, J., & Morrisey, M.A. (1990). Trust in employee/employer relationships: a survey of West Michigan managers. *Public and Personnel Management,* 19, 443–61.

Mitchell, E.S. (1986). Multiple triangulation: A methodology for nursing science. *Advances in Nursing Science,* 8 (3), 18-26.

Mitchell, P.H., Armstrong, S., Simpson, T.F., *et al.* (1989). American association of critical care nurses demonstration project: Profile of excellence in critical care nursing. *Heart and Lung,* 18, 219-237.

Molloy, G.J., & O` Boyle, C.A. (2005). The SHEL Model: A useful tool for analyzing and teaching the contribution of human factors to medical error. *Academic Medicine,* 80 (2), 152-155.

Morey, J.C., Simon, R., Jay, G.D., *et al.* (2002). Error Reduction and Performance Improvement in the Emergency Department through Formal Teamwork Training: Evaluation Results of the MedTeams Project. *Health Services Research,* 37 (6), 1553-1581.

Morrison, P. (1991). The caring attitude in nursing practice: a repertory grid study of trained nurses' perceptions. *Nurse Education Today,* 11 (1), 3-12.

Morrison, W., Haas, E., Shaffner, D., & Garrett, E. (2003). Noise, stress, and annoyance in a pediatric intensive care unit. *Critical Care Medicine,* 31 (1), 113-119.

Murphy, M.F., Casbard, A.C., Ballard, S., *et al.* (2007). Prevention of bedside errors in transfusion medicine (PROBE-TM) study: a cluster-randomized, matched-paired clinical areas trial of a simple intervention to reduce errors in the pre-transfusion bedside check. *Transfusion,* 47 (5), 771-780.

Murphy, S.A. (1989). Multiple triangulation: applications in a program of nursing research. *Nursing Research*, 38, 294-298.

Narumi, J., Miyazawa, S., Miyata, H., *et al.* (1999). Analysis of human error in nursing care. *Accident Analysis and Prevention,* 31, 625-629.

National Coordinating Council for Medication Error Reporting and Prevention (1995). National coordinating council established for medication error reporting and prevention. Press release. www.nccmerp.org/press/press1995-07-28.

National Patient Safety Agency. (2005). *Building a memory: preventing harm, reducing risks and improving patient safety*. London: National Patient Safety Agency.

Nebeker, J.R., Barach, P., & Samore, M.H. (2004). Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. *Annals of Internal Medicine,* 140 (10), 795–801.

Needleman, J., Buerhaus, P., Mattke, S., *et al.* (2002). Nurse-staffing levels and the quality of care in hospitals. *New England Journal of Medicine,* 346, 1715-1722.

Newell, A., & Simon, H.A. (1972). *Human Problem Solving.* Englewood Cliffs, N.J.: Prentice-Hall.

Newman, M.A. (2003). *Pilot Systematic Review and Meta-Analysis on the Effectiveness of Problem Based Learning. On behalf of the Campbell Collaboration Systematic Review Group on the effectiveness of problem based learning.* Newcastle, UK: University of Newcastle, Learning and Teaching Support Network.

Newton, R.R., Prensky, D., & Schuessler, K. (1982). Form effect in the measurement of feeling states. *Social Science Research,* 11, 301–17.

Nixon, P., & Dhillon, S. (1996). Medication errors in paediatrics. *Progress in Practice: UKCPA Autumn Symposium*, 18-9.

Norman, D.A. (1968). Towards a theory of memory and attention. *Psychological Review,* 75, 522-536.

Norman, D.A. (1981). Categorisation of action slips. *Psychological Review,* 88, 1-15.

Norman, D.A. (1988). *The Psychology of everyday things.* New York: Basic Books.

Norman, D.A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behaviour. In R.J. Davidson, G.E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation.* New York: Plenum.

Norris, A.E., & Aroian, K.J. (2004). To transform or not transform skewed data for psychometric analysis: that is the question. *Nursing Research,* 53 (1), 67-71.

Nunnally, J.C., & Bernstein, I. (1994). *Psychometric theory.* New York: McGraw-Hill.

O'Hara, D.A. & Carson, N.J. (1995). Reporting of adverse events in hospitals in Victoria. *Medical Journal of Australia*, 166, 460-463.

O'Neil, A., Petersen, L., Cook, E., *et al.* (1993). Physician reporting compared with medical-record review to identify adverse medical events. *Annals of Internal Medicine,* 119, 370-376.

O'Shea, E. (2001). Factors contributing to medication errors: a literature review. *Journal of Clinical Nursing,* 5 (5), 496-504.

Osbourne, J., Blais, K., & Hayes, J. (1999). Nurses' perceptions: When is it a medication error? *Journal of Nursing Administration*, 29 (4), 33-38.

Pani, J.R., & Chariker, J.H. (2004). The psychology of error in relation to medical practice. *Journal of Surgical Oncology,* 88, 130-142.

Parker, D., Reason, J.T., Manstead, A.S.R., *et al.,* (1995). Driving errors, driving violations and accident involvement. *Ergonomics,* 38 (5), 1036-1048.

Parsons, H.M. (1974). What happened at Hawthorne? *Science,* 183 (4128), 922-932.

Paton, J., & Wallace, J. (1997). Medication errors. *Lancet*, 349, 959-60.

Patton, M. Q. (1990). *Qualitative Evaluation and Research Methods*. NewBury Park, CA: SAGE Publications.

Paul, H. (1880). *Prinzipien der Sprachgeschichte.* Halle: Niermeyer.

Perneger, T.V. (2005). The Swiss cheese model of safety incidents: are there holes in the metaphor? *Health Services Research,* 5 (1), 70-78.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies.* New York: Basic books.

Phillips, J., Beam, S., Brinker, A., *et al.* (2001). Retrospective analysis of mortalities associated with medication errors. *American Journal of Health-System Pharmacy*, 58 (19), 1835-1841.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P. (2003). Common method biases in behavioural research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology,* 83 (5), 879-903.

Polit, D.F. (1996). Factor analysis. In *Data analysis and statistics for nursing research* (pp. 345–379). Stamford, CT: Appleton & Lange.

Potier, D., Monteiro, P., Mooktiar, M., & Shaw, A. (2005). Pilot study to show the loss of important data in nursing handover. *British Journal of Nursing,* 14 (20), 1090-1093.

Potter, P., Boxerman, S, Wolf, L., *et al.* (2004). Mapping the nursing process: A new approach for understanding the work of nursing. *Journal of Nursing Administration,* 34 (2), 101-109.

Priest, H., Roberts, P., & Woods, L. (2002). An overview of three different approaches to the interpretation of qualitative data. Part I: Theoretical issues. *Nurse Researcher,* 10 (1), 30-42.

Pringle, C., & Churchhill, B. (1995). The randomised controlled trials in general practice: gold standard or fool's gold. *British Medical Journal*, 311, 1381-1382.

Prot, S., Fontan, J.E., Alberti C., *et al.* (2005). Drug administration errors and their determinants in pediatric in-patients. *International Journal for Quality in Health Care*, 17 (5), 381-389.

Rahman, N. (1996). Caregivers' sensitivity to conflict: the use of vignette methodology. *Journal of Elder Abuse and Neglect*, 8, 35-47.

Raju, T.N., Kecskes, S., Thornton, J.P., *et al.* (1989). Medication errors in neonatal and paediatric intensive care units. The *Lancet,* 2, 374-6.

Rasmussen, J. (1983). Skills, rules, knowledge: signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions: Systems, Man & Cybernetics,* 13, 257-267.

Rasmussen, J., & Jensen, A. (1974). Mental procedures in real-life tasks: A case study of electronic trouble-shooting. *Ergonomics,* 17, 293-307.

Raz, T., & Kraus, G.P. (1989). Prediction of medication error probability based on patient characteristics. *Quality Assurance in Health Care*, 1 (4), 207-16.

Reason, J. (1979). Actions not as planned. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness* (pp. 67–89). London: Academic Press.

Reason, J. (1987). Generic Error Modelling System (GEMS): a cognitive framework for locating common human error forms. In *New Technology and Human Error,* edited by Rasmussen, J., Duncan, K., & Leplat, J. London, UK: Wiley.

Reason, J. (1990). *Human Error.* New York: Cambridge University Press.

Reason, J. (1993). *Review. I. Overview. II. Theory.* Derby: British Railways Board.

Reason, J. (1993). The human factor in medical accidents. In Vincent, C.A. (ed.) *Medical Accidents.* Oxford: Oxford Medical Publications.

Reason, J. (1995). *Comprehensive Error Management in Aircraft Engineering: A Manger's Guide.* London Heathrow: British Airways Engineering.

Reason, J. (1997). *Managing the risks of organisational accidents.* Aldershot: Ashgate Publishing Ltd.

Reason, J. (1998). Achieving a safe culture: Theory and practice. *Work and Stress,* 12 (3), 293-306.

Reason, J. (2000). Human error: models and management. *British Medical Journal,* 320 (7237), 768-770.

Reason, J., & Hobbs, A. (2003). *Managing maintenance error: A practical guide.* Hampshire, UK: Ashgate.

Reason, J., Manstead, A., Stradling, S., *et al.,* (1990). Errors and violations on the road: a real distinction? *Ergonomics,* 33 (10/11), 1315-1332.

Reason, J., Parker, D., & Lawton, R. (1998). Organisational controls and safety: The varieties of rule-related behaviour. *Journal of Occupational and Organisational Psychology,* 71, 289-304.

Reason, J., Shotton, R., Wagenaar, W.A.., *et al.* (1989). *TRIPOD—A Principled Approach to Safety.* Leiden University: Report for Shell International.

Risser, D.T., Rice, M.M,, Salisbury, M.L., *et al.* (1999). The potential for improved teamwork to reduce medical errors in the emergency department. *Annals of Emergency Medicine,* 34 (3), 373-383.

Rogers, A., Hwang, W., & Scott, L.D. (2004). The effects of work breaks on staff nurse performance. *Journal of Nursing Administration,* 34 (11), 512-519.

Roseman, C., & Booker, J.M. (1995). Workload and environmental factors in hospital medication errors. *Nursing Research,* 44 (4), 226-230.

Ross, L.M., Wallace, J., & Paton, J.Y. (2000). Medication errors in a paediatric teaching hospital in the UK: five years operational experience. *Archives of Diseases in Childhood,* 83, 492-7.

Rousseau, D.M., & Libuser, C. (1997). Contingent workers in high risk environments. *California Management Review,* 39 (2), 103-123.

Rudolph, J.W., & Repenning, N.P. (2002). Disaster dynamics: Understanding the role of stress and interruptions in organisational collapse. *Administrative Science Quarterly,* 47, 1-30.

Runciman, W. B. (1993). Qualitative versus quantitative research — balancing cost, yield and feasibility. *Quality and Safety in Health Care,* 11, 146-147.

Runciman, W.B. (2002). Lessons from the Australian Patient Safety Foundation: Setting up a national patient safety surveillance system – is this the right model? *Quality and Safety in Health Care,* 11, 246-251.

Runciman, W.B., Edmonds, M.J., & Pradhan, M. (2002). Setting priorities for patient safety. *Quality and Safety in Health Care,* 11, 224–229.

Runciman, W.B., Roughead, E.E., Semple, S.J., & Adams, R.J. (2003). Adverse drug events and medication errors in Australia. *International Journal for Quality in Health Care,* 15, i49-i59.

Runciman, W.B., Williamson, J.A.H., Deakin, A., *et al.* (2006). An integrated framework for safety, quality and risk management: An information and incident management system based on a universal patient safety classification. *Quality and Safety in Health Care,* 15, i82-i90.

Ryan, A. M., Hicks, L., & Midgley, C. (1997). Social goals, academic goals, and avoiding seeking help in the classroom. *Journal of Early Adolescence,* 17, 152–171.

Ryan, A. M., & Pintrich, P. R. (1997). Should I ask for help?: The role of motivation and attitudes in adolescents' help seeking in math class. *Journal of Educational Psychology,* 89, 329–341.

Ryan, A.M., Pintrich, P.R., & Midgley, C. (2001). Avoiding Seeking Help in the Classroom: Who and Why? *Educational Psychology Review* ,13 (2), 93-114.

Sandal, G.M. (1999). The effects of personality and interpersonal relations on crew performance during space simulation studies. *International Journal of Life Support Biosphere Science,* 5, 226-37.

Sasou, K. & Reason, J. (1999). Team errors: definition and taxonomy. *Reliability Engineering and System Safety,* 65, 1-9.

Savage, J. (2000). Ethnography and health care. *British Medical Journal,* 321, 1400-1402.

Savin-Baden, M. (2000). *Problem-based learning in Higher Education: Untold Stories.* Buckingham: Society for Research in Higher Education and Open University Press.

Schiff, G.D., Aggarwal, H.C., Kumar, S. & McNutt, R.A. (2000). Prescribing potassium despite hyperkalemia: medication errors uncovered by linking laboratory and pharmacy information systems. *The American Journal of Medicine*, 109 (6), 494-497.

Schilling, J. (2006). On the pragmatics of qualitative assessment: Designing the process for content analysis. *European Journal of Psychological Assessment, 22* (1), 28-37.

Schmitt, N. (1994). Method bias: The importance of theory and measurement. *Journal of Organizational Behavior,* 15, 393–398.

Schmitt, N., & Stults, D.M. (1985). Factors defined by negatively keyed items: the result of careless respondents? *Applied Psychological Measurement,* 9, 367-373.

Schneider, M.P., Cotting, J., & Pannatier, A. (1998). Evaluation of nurses' errors associated in the preparation and administration of medication in a pediatric intensive care unit. *Pharmacy World Science*, 20, 178-82.

Scholz, D.A. (1990). Establishing and monitoring an endemic medication error rate. *Journal of Nursing Quality Assurance,* 4 (2), 71-85.

Schwappach, D. L. B., & Koeck, C. M. (2004). What makes an error unacceptable? A factorial survey on the disclosure of medical errors. *International Journal for Quality in Health Care*, 16 (4), 317-326.

Seago, J.A., Williamson, A., & Atwood, C. (2006). Longitudinal analyses of nurse staffing and patient outcomes: More about failure to rescue. *Journal of Nursing Administration,* 36 (1), 13-21.

Segatore, M., Miller, M., & Webber, K. (1994). Medication out of control. *The Canadian Nurse*, 90 (8), 35-39.

Selbst, S.M., Fein, J.A., Osterhoudt, K., & Ho, W. (1999). Medication errors in a pediatric emergency department. *Pediatric Emergency Care*, 15, 1- 4.

Sexton, J.B., Thomas, E.J., & Helmreich, R.L. (2000). Error, stress and teamwork in medicine and aviation: cross sectional surveys. *British Medical Journal,* 320, 745-749.

Sexton, T., & Helmreich, R. L. (2004).Translating teamwork behaviours from aviation to healthcare: development of behavioural markers for neonatal resuscitation. *Quality and Safety in Health Care*, (13), i57-i64.

Shappell, S.A., & Wiegmann, D.A. (1997). A human error approach to accident investigation: The taxonomy of unsafe operations. *International Journal of Aviation,* 7 (4), 269-291.

Shell International Exploration and Production B.V. (2004). 'Hearts *and minds' toolkit:* London: Energy Institute.

Sheppard, J.A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin,* 113, 67-81.

Shih, F.J. (1998). Triangulation in nursing research: Issues of conceptual clarity and purpose. *Journal of Advanced Nursing,* 28 (3), 631-341.

Sim, J., & Wright, C. (2000). *Research in health care: Concepts, design and methods.* Cheltenham, UK: Nelson Thornes.

Skibinski, K.A., White, B.A., Lin, L., *et al.* (2007). Effects of technological interventions on the safety of a medication-use system. *American Journal of Health-System Pharmacy*, 64 (1), 90-96.

Smidts, C., Song-Hua, S., & Mosleh, A. (1995). A taxonomy and root-cause analysis of human cognitive behaviour based on a cognitive model. *Reliability and Maintainability Symposium Proceedings,* 16-19, 303-314.

Smith, J.A., & Osborn, M. (2003). Interpretative phenomenological analysis. In smith, J.A., (Ed.) *Qualitative psychology: a practical guide to methods*. London, UK: Sage.

Smith, S. (1996). Ethnographic inquiry in physiotherapy research: Illuminating the working culture of the physiotherapy assistant. *Physiotherapy,* 82, 342-349.

Sorra, J.S., & Nieva, V.F. (2004). *Hospital Survey on Patient Safety Culture*. Rockville, MD: Agency for Healthcare Research and Quality.

Spath, E. (2000). *Error reduction in health care.* San Francisco: Josey Bass.

Spencer, F.C. (2000). Human Error in Hospitals and Industrial Accidents: Current Concepts. *Journal of the American College of Surgeons,* 191 (4), 410-418.

Spigelman, A., & Swan, J. (2005). Review of the Australian Incident Monitoring System. *ANZ Journal of Surgery,* 75, 657-661.

Stein, L. (1967). The doctor-nurse game. *Archives of General Psychiatry,* 16 (6), 699-703.

Stein, L., Watts, D.T., & Howell, T. (1990). The doctor-nurse game revisited. *New England Journal of Medicine,* 322 (8), 546-549.

Streiner, D.L. (1994). Figuring out factors: The use and misuse of factor analysis. *Canadian Journal of Psychiatry*, 39, 135–140.

Sudman, S., & Bradburn, N. (1974). Response effects in surveys: a review and synthesis. Chicago, IL: Aldine.

Sudman, S., & Bradburn, N. (1982). Asking questions: a practical guide to questionnaire design. San Francisco, CA: Jossey-Bass.

Sully, J. (1881). *Illusions: A Psychological Study.* London, UK: C. Kegan Paul & Co.

Sutcliffe, K.M., Lewton, E., & Rosenthal, M.M. (2004). Communication failures: An insidious contributor to medical mishaps. *Academic Medicine,* 79 (2), 186-194.

Sweet, S.J., & Norman, I.J. (1995). The nurse-doctor relationship: A selective review. *Journal of Advanced Nursing,* 22, 165-170.

Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics* (4th edition). Boston: Allyn & Bacon.

Tamuz, M. & Harrison, M.I. (2006). Improving patient safety in hospitals: Contributions of high-reliability theory and normal accident theory. *Health Services Research*, 41 (4 II), 1654-1676.

Taxis, K., & Barber, N. (2000). Causes of intravenous medication errors: An ethnographic study. *Quality and Safety in Health Care,* 12, 343-347.

Temple, W.J. (2004). Medical errors: A challenge to all physicians. *Journal of Surgical Oncology,* 88, 111.

Tesch, R. (1990). *Qualitative Research: Analysis Types & Software Tools.* Bristol, PA: The Falmer Press.

Thorne, S. (2000). Data analysis in qualitative research. *Evidence-Based Nursing,* 3, 68-70.

Thurmond, V.A. (2001). The point of triangulation. *Journal of Nursing Scholarship,* 33 (3), 253-258.

Tierney, W.M., Miller, M.E., Overhage, J.M. & McDonald, C.J. (1993). Physician inpatient order writing on microcomputer workstations. Effects on resource utilization. *Journal of the American Medical Association,* 269, 379-83.

Topf, M. (2000). Hospital noise pollution: An environmental stress model to guide research and clinical interventions. *Journal of Advance Nursing,* 31 (3), 520-528.

Tourangeau, A.E., Cranley, L.A., & Jeffs, L. (2006). Impact of nursing on hospital patient mortality: A focused review and related policy implications. *Quality and Safety in Health Care,* 15, 4-8.

Tourangeau, R, & Smith, T.W. (1996). Asking sensitive questions: The impact of data collection mode, question format and question context. *Public Opinion Quarterly,* 60, 275-304.

Tucker , A.L. (2004). The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management,* 22 (2), 151-169.

Tucker, A.L., & Spear, S.J. (2006). Operational failures and interruptions in hospital nursing. *Health Services Research,* 41 (3), 643-662.

Upenieks, V. (2002). What constitutes successful nurse leadership? A qualitative approach utilizing Kanter's theory of organizational behavior. *Journal of Nursing Administration*, 32 (12), 622-632.

Uribe, C.L., Schweikhart, S.B., Pathak, D.S., *et al.* (2002). Perceived barriers to medical-error reporting: An exploratory investigation. *Journal of Healthcare Management,* 47 (4), 263-280.

Van der Schaaf, T. (2005). *PRISMA-Medical: A brief description*. Netherlands: Eindhoven University of Technology.

Vaughn, D. (1996). *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago, Illinois: University of Chicago Press.

Vernon, D.T.A., & Blake, R.L. (1993). Does Problem-based Learning Work? A Meta-analysis of Evaluative Research. *Academic Medicine*, 69 (7), 550-563.

Vincent, C. (2001). *Clinical risk management: enhancing patient safety (*2nd eds). London, UK: BMJ publications.

Vincent, C. (2004). Analysis of clinical incidents: a window on the system not a search for root causes. *Quality and Safety in Health Care,* 13, 242-243.

Vincent, C. (2006). *Patient Safety.* London, UK: Elsevier.

Vincent, C., Neale, G., & Woloshynowych, M. (2001). Adverse events in British hospitals: preliminary retrospective record review. *British Medical Journal*, 322, 517 – 519.

Vincent, C., Taylor-Adams, S., Chapman, E.J., *et al,* (2000). How to investigate and analyse clinical incidents: Clinical Risk Unit and Association of Litigation and Risk Management protocol. *British Medical Journal*, 320 (7237): 777 - 781.

Vincent, C., Taylor-Adams, S. & Stanhope, N. (1998). Framework for analysing risk and safety in clinical medicine. *British Medical Journal,* 316, 1154-1157.

Vincer, M.J., Murray, J.M., Yuill, A., *et al.* (1989). Drug errors and incidents in a neonatal intensive care unit. *American Journal of Diseases of Children*, 143, 737-40.

Vogus, T.J., & Sutcliffe, K.M. (2007). The Safety Organizing Scale: Development and validation of a behavioral measure of safety culture in hospital nursing units. *Medical Care,* 45 (1), 46-54.

Wagenaar, W.A., Groeneweg, J., Hudson, P., & Reason, J. (1994). Promoting safety in the oil industry. *Ergonomics,* 37 (12), 1999-2013.

Wakefield, D.S., Wakefield, B.J., Uden-Holman, T., *et al.* (1999). Understanding why medication administration errors may not be reported. *American Journal of Medical Quality,* 14, 81-88.

Wakefield, D.S., Ward, M.M., Groath, D., *et al.* (2008). Complexity of medication-related verbal orders. *American Journal of Medical Quality,* 23 (1), 7-17.

Walston, S., & Kimberly, J. (1997). Reengineering hospitals: Evidence from the field. *Hospital and Health Services Administration,* 42 (2), 143-163.

Walters, J.A. (1992). Nurses' perceptions of reportable medication errors and factors that contribute to their occurrence. *Applied Nursing Research,* 5 (2), 86-88.

Walton, H.J., & Matthews, M.B. (1989). Essentials of Problem Based Learning. *Medical Education,* 23, 542- 558.

Walton, M. (2004). Creating a "no blame" culture: have we got the balance right? *Quality and Safety in Health Care,* 13, 163-164.

Waring, J. (2005). Beyond blame: cultural barriers to medical incident reporting. *Social Science and Medicine,* 60 (9), 1927-1935.

Webb, E.J., Campbell, D.T., Schwartz, R.D., *et al.* (1981). *Non-reactive Measures in the Social Sciences.* Boston: Houghton Mifflin.

Weber, R.P. (1990). *Basic Content Analysis.* Beverly Hills, Calif.: Sage Publications.

Wensing, M., & Grol, R. (1994). Single and combined strategies for implementing changes in primary care: A literature review. *International Journal for Quality in Health Care,* 6 (2), 115-132.

Werner, O., & Schoepfle, G.M. (1987). *Systematic fieldwork.* CA: Sage Publications.

West, E. (2000). Organisational sources of safety and danger: sociological contributions to the study of adverse events. *Quality in Health Care,* 9, 120-126.

Wieman, T.J., & Wieman, E.A. (2004). A systems approach to error prevention in medicine. *Journal of Surgical Oncology,* 88, 115-121.

Williams, A. (1996). How to avoid mistakes in medicine administration. *Nursing Times,* 92 (13), 40-41.

Williams, P.L., & Webb, C. (1994). The Delphi Technique: A methodological discussion. *Journal of Advanced Nursing,* 19 (1),180-6.

Wilson, D.G., McArtney, R.G., Newcombe, R.G., *et al.* (1998). Medication errors in paediatric practice: insights from a continuous quality improvement approach. *European Journal of Pediatrics*, 157, 769-74.

Wilson, H.S., & Hutchinson, S.A. (1991). Triangulation of qualitative methods: Heideggerian hermeneutics and grounded theory. *Qualitative Health Research,* 1 (2), 263-276.

Wilson, R.M., Runciman, W.B., Gibberd, R.W *et al.* (1995). The Quality in Australian Health Care Study. *Medical Journal of Australia,* 163 (4), 458-471.

Wolf, L.D., Potter, P., Sledge, J.A., *et al.* (2006). Describing nurses' work: Combining quantitative and qualitative analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 48 (1), 5-14.

Wong, I.C.K., Ghaleb, M.A.G., Franklin, B.D., *et al.* (2004). Incidence and nature of dosing errors in paediatric medications. *Drug Safety,* 27, 661-70.

Woolf, S.H., Grol, R., Hutchinson, A., *et al.* (1999). Potential benefits, limitations and harms of clinical guidelines. *British Journal of Medicine,* 318, 527-530.

Yi, M., & Jezewski, M. (2000). Korean nurses adjustment to hospitals in the United States of America. *Journal of Advanced Nursing,* 32 (3), 721-729.

Yu, K.H., Nation, R.L., & Dooley, M.J. (2005). Multiplicity of medication safety terms, definitions and functional meanings: when is enough enough? *Quality and Safety in Health Care,* 14, 358-363.

Zhang, J., Patel, V.L., Johnson, T.R., & Shortliffe, E.H. (2004). A cognitive taxonomy of medical errors. *Journal of Biomedical Informatics,* 37, 193-204.

Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on micro accidents in manufacturing jobs. *Journal of Applied Psychology,* 85, 587-596.

Appendix I: Articles included in systems interventions review

ARTICLE	AUTHOR(S)	YEAR	TITLE	SOURCE
1	Bates, D.W., *et al.*	1998	Effect of computerized physician order entry and a team intervention prevention of serious medication errors	Journal of the American Medical Association, 280 (15), 1311-1316.
2	Bates, D.W., *et al.*	1999	The impact of computerized physician order entry on medication error prevention	Journal of the American Medical Informatics Association, 6 (4), 313-321.
3	Walsh-Sukys, M., *et al.*	2001	Reducing light and sound in the neonatal intensive care unit: An evaluation of patient safety, staff satisfaction and costs	Journal of Perinatology, 21, 230-235.
4	Fasting, S., & Gisvold, S.E.	2000	Adverse drug errors in anesthesia, and the impact of coloured syringe labels	Canadian Journal of Anesthesia, 47 (11), 1060-1067.
5	Landrigan, C.P., *et al.*	2004	Effect of reducing interns work hours on serious medical errors in intensive care units	New England Journal of Medicine, 351, 1838-48.
6	Bizovi, K.E., *et al.*	2002	The effect of computer-assisted prescription writing on emergency department prescription errors	Academic Emergency Medicine, 9, 1168-1175.
7	Greengold, N.L., *et al.*	2003	The impact of dedicated medication nurses on the medication administration error rate	Archives of Internal Medicine, 163, 2359-2367.
8	Shaw, J., *et al.*	2003	Error reduction: Academic detailing as a method to reduce incorrect prescriptions	European Journal of Clinical Pharmacology, 59, 697-699.
9	Dean-Franklin, B., *et al.*	2007	The impact of a closed loop electronic prescribing and administration system on prescribing errors, administration errors and staff time: a before and after study.	Quality and Safety in Health Care, 16, 279-284.

Appendix I (continued): Articles included in systems interventions review

ARTICLE	AUTHOR(S)	YEAR	TITLE	SOURCE
10	Spencer, D.C., *et al.*	2005	Effect of a computerised prescriber-order-entry system on reported medication errors	American Journal of Health-System Pharmacy, 62, 416-419.
11	Davey, A.L., *et al.*	2008	Decreasing paediatric prescribing errors in a district general hospital	Quality and Safety in Health Care, 17, 146-149.
12	Donihi, A.C., *et al.*	2006	Use of a standardised protocol to decrease medication errors and adverse events related to sliding scale insulin	Quality and Safety in Health Care, 15, 89-91.
13	King, W.J., *et al.*	2003	The effect of computerised physician order entry on medication errors and adverse drug events in pediatric patients	Pediatrics, 112, 506-509.
14	Silver, M.P., & Antonow, J.A.	2000	Reducing medication errors: A peer review organisation collaboration	Joint Commission Journal on Quality Improvement, 26 (6), 332-340.
15	Upperman, J.S., *et al.*	2005	The impact of hospital-wide computerised physician order entry on medical errors in a pediatric hospital	Journal of Pediatric Surgery, 40, 57-59.
16	Larsen, G.Y., *et al.*	2005	Standard drug concentrations and smart pump technology reduce continuous medication infusion errors in pediatric patients	Pediatrics, 116 (1), 21-25.
17	Cordero, L., *et al.*	2004	Impact of computerised physician order entry on clinical practice in a newborn intensive care unit	Journal of Perinatology, 24, 88-93.
18	Potts, A.L., *et al.*	2004	Computerised physician order entry and medication errors in a pediatric critical care unit	Pediatrics, 113 (1), 59-63.
19	Rothschild, J.M., *et al.*	2005	A controlled trial of smart infusion pumps to improve medication safety in critically ill patients	Critical Care Medicine, 33 (3), 533-540.

Appendix II: Summary of key findings from systems interventions review

ARTICLE #	Systems factor(s) intervention							Dependent variable(s)						Significant decrease in errors			
	Standard operating procedures	Decision support	Education / training	Design (environment/ equipment)	Human resources	Local working conditions	Team communication	Medication errors (all)	Prescribing errors	Medication administration errors	Adverse Drug Events	'Serious' medication errors	Potential Adverse Drug Events	Yes	No	Some components of intervention successful	Not reported
1	✓	✓					✓					✓				✓	
2	✓	✓						✓			✓		✓			✓	
3				✓				✓							✓		
4			✓	✓						✓					✓		
5						✓						✓		✓			
6	✓								✓								✓
7					✓					✓					✓		
8		✓	✓						✓					✓			
9	✓								✓	✓				✓			

Appendix II (continued): Summary of key findings from systems interventions review

ARTICLE #	Systems factor(s) intervention							Dependent variable(s)						Significant decrease in errors			
	Standard operating procedures	Decision support	Education / training	Design (environment/ equipment)	Human resources	Local working conditions	Team communication	Medication errors (all)	Prescribing errors	Medication administration errors	Adverse Drug Events	'Serious' medication errors	Potential Adverse Drug Events	Yes	No	Some components of intervention successful	Not reported
10	✓							✓							✓		
11		✓	✓						✓							✓	
12	✓								✓					✓			
13	✓							✓						✓			
14	✓		✓	✓				✓						✓			
15	✓							✓						✓			
16	✓	✓		✓						✓				✓			
17	✓	✓							✓								✓
18	✓	✓							✓					✓			
19		✓						✓			✓		✓		✓		
Total	**12**	**8**	**4**	**4**	**1**	**1**	**1**	**7**	**7**	**4**	**2**	**2**	**2**	**9**	**5**	**3**	**2**

Appendix III: Reviewed articles - interventions in detail

ARTICLE	INTERVENTIONS		Statistically significant decrease in medication errors	
	Type	Further details	YES	NO
1	CPOE	System not specified	✓	
	Decision Support	Detection of drug-allergy interactions		
		Detection of drug-drug contraindications		
		Consequent orders		
	Team functioning	Increase of Pharmacist availability on wards		✓
		Recommended dilutions charts		
		Computerised drip-rate infusion calculation program		
		Standardised labelling of IV bags, tubes and pumps		
		Pharmacy-nurse communication log book		
2	CPOE	In-house system	✓	
	Decision support	Detection of drug-allergy interactions		✓
		Detection of duplicate orders		
		Detection of drug-drug contraindications		
		Maximum dosage limits		
3	Sound modifications	Weather stripping on doors and drawers		✓
		Replacement of metal containers with rubber		
		Carpeting		
		Sound absorbing acoustic material		
	Light modifications	Reduction of fluorescent lighting		
		Addition of halogen spotlighting		

Appendix III (continued): Reviewed articles - interventions in detail

ARTICLE	INTERVENTIONS		Statistically significant decrease in medication errors	
	Type	Further details	YES	NO
4	IV label design	Colour-coded to prevent improve drug selection		✓
	Education program	Mechanisms behind drug error		
5	Shift pattern changes	Change in no. of working hours per week	✓	
		Change in no. of working hours per shift		
6	Computer-assisted prescription writing	EmStat, (Cyberplus Corporation, Texas)	Not reported	
7	Change of nursing job role	Specialist 'medication' nurses		✓
	Educational program	Education of nurses in basic pharmacology & principles of safe medication administration.		
8	Academic detailing	One-to-one prescribing troubleshooting interview	✓	
		Bookmark prompt on safe prescribing practice		
9	Closed-loop prescribing & administration system	Ward-based & hand-held prescribing terminals	✓	
		Ward-based automated dispensing units		
		Electronic drug trolleys		
		Bar-coded patient ID wristbands		
		Electronic medication administration records		
10	CPOE	Siemens Medical Solutions system		✓
11	Educational program	Prescribing skills tutorial	✓	
	Decision Support	Introduction of bedside prescribing guidelines		✓
12	Implementation of insulin prescribing protocol	Provision of new prescribing guidelines	✓	
		Standardised order form		

Appendix III (continued): Reviewed articles - interventions in detail

ARTICLE	INTERVENTIONS		Statistically significant decrease in medication errors	
	Type	Further details	YES	NO
13	CPOE	Sunrise Clinical manager	✓	
14	Improving information access	Redesign of medication administration records	✓	
		Increased availability of allergy information on wards		
		Automated pharmacy checking allergy information		
		Improved access to drug information		
		Training in antidote information		
	Standardising and simplifying medication procedures	Standardisation of lethal medication ordering procedures		
		Reducing variation in use of materials		
		Reducing demands on memory		
		Improving communication		
	Restricting access to lethal drugs	Such as chemotherapy drugs, insulin & Potassium Chloride		
	Education program	Assess medication knowledge deficiencies		
		Increase awareness for error potential		
		Pharmacist led discharge patient education		
	Workplace design	Separation of look- and sound-alike drugs		
		Changing vendors for medications with dangerous packaging		
15	CPOE	Customised system (unspecified)	✓	

Appendix III (continued): Reviewed articles - interventions in detail

ARTICLE	INTERVENTIONS		Statistically significant decrease in medication errors	
	Type	Further details	YES	NO
16	Decision support	Development of standardised concentrations of 32 common IV infusion medications	✓	
	Smart pump technology	Medex system (California)		
		Maximum dosage alerts		
		Drug library		
	IV Label design	Pharmacy generated medication labels based on human-engineering principles		
17	CPOE	Invision 24 (Siemens Medical solutions)	Reduction but statistical significance not reported	
	Decision support	Detection of drug-allergy interactions		
		Detection of duplicate orders		
		Detection of drug-drug contraindications		
		Cross-sensitivity checking		
		Maximum dosage limits		
		Corollary order reminders		
		Weight-based dosing		
		Drug route restrictions		
18	CPOE	Horizon Expert Order System (McKesson, Atlanta)	✓	
	Decision support	Detection of drug-allergy interactions		
		Dose checking		
		Detection of drug-drug contraindications		
19	Smart pump technology	Medley pump (Alaris Medical Systems, California)		✓
		Drug library providing standardised concentrations		
		Automatic weight-based volume and rate calculations		
		Dosage and rate limits		

Appendix IV: Senior manager's invitation letter

Bradford Teaching Hospitals NHS
An NHS Foundation Trust

Exploring the role of organisational factors in medication error

Dear Sir/Madam,

I am a researcher based at the Institute of Psychological Sciences at the University of Leeds investigating the role of organisational factors in medication errors in secondary care. The research project will be carried out in collaboration with Bradford NHS Trust under the supervision of Dr John Wright (Director of Research) and is funded by the Economic and Social Research Council. For the initial stage of the project I am keen to recruit NHS senior managers and both clinical and non-clinical ward-based staff for a series of short interviews. I would be extremely grateful if you could spare a little time to read the enclosed project information sheet and consider whether you would like to be involved in this valuable and innovative research project.

Many Thanks

Samantha Beaumont
Institute of Psychological Sciences
University of Leeds
Leeds
LS2 9JT
Email: psc2sjb@leeds.ac.uk

Appendix V: Interviews - participant information sheet

The role of organisational factors in medication error

My name is Sam Beaumont and I am a researcher based at the Institute of Psychological Sciences at the University of Leeds and am working in collaboration with Bradford Teaching Hospitals Trust on my PhD research project. I am currently recruiting a variety of staff members working for Bradford NHS Trust, who might like to be involved in a new and exciting research project and would very much like you to be involved. **I am interested in interviewing a wide variety of clinical staff working on general medical wards and also NHS management staff working in risk, clinical governance, research, nursing, patient services and pharmacy.** Here are a few important details and questions you might like answering before making a decision to participate.

Who am I & who do I work for?

The research project is funded by the Economic and Social Research Council and Bradford Teaching Hospitals Trust as part of a PhD-CASE collaborative studentship. My academic background is in the field of Psychology (BSc Psychology - University of York; MSc Health Psychology - University of Leeds) and I also have an employment background in psychiatric health care services. I will be based partly in the Institute of Psychological Sciences at the University of Leeds and also at Bradford Royal Infirmary for the duration of the project (3 years in total).

What is the research project about?

The overall aim of the research project is to investigate the organisational or systems factors which may give rise to medication errors made in secondary care. From initial analyses of these organisational factors, the project aims to develop a *proactive* measurement tool in the form of a safety checklist, similar to those already developed and employed in industry. This tool will actively predict the factors most likely to give rise to errors *before* they have occurred. The tool will ultimately aim to provide feedback on ways such areas might be improved and thus reduce the likelihood of future errors occurring.

Why study this subject?

Industrial evidence has emphasised the importance of investigating the role of organisational or workplace factors which may have made individual errors possible. The current research project will apply this 'organisational theory' to the field of health care and will attempt to uncover which organisational factors are most likely to predict the occurrence of *medication* errors.

How will you be involved?

You will be asked to take part in an informal interview, lasting approximately 30-40 minutes, scheduled at your convenience over the next few months. All interviews will be conducted by me and will be tape-recorded to ensure I can pay full attention to the issues being discussed. The information discussed in these interviews will be strictly confidential and any details which would make you identifiable will not be reported on any documentation.

What will happen during the interviews?

During the first part of the interview you will be presented with a printed hypothetical error scenario (a short story describing one or more medical errors in the delivery of patient care). These scenarios have been developed from the medical literature and are based on actual events - although not at Bradford Hospitals. Once you have read this scenario we will go on to discuss any ideas you have about the causes behind the errors, in terms of the individuals involved, contributing environmental factors and finally the role of the organisation.

We will move on to discuss a 'real-life' example of medication error which you have either been involved in (in clinical terms), or an error you simply have knowledge of in enough detail to be able to describe it to me. This error does not necessarily need to have ended in patient harm - it could simply be a near miss or a common mistake which you think has the potential to result in patient harm. You should come prepared to the interview prepared to discuss a real error or 'near miss' you have had experience of. We will not discuss specific details of any individuals involved in the error or near miss you describe and you will not be asked to divulge any personal information which could lead to identification of the individuals concerned. Similarly you will not be asked to talk about details which might cause you any discomfort and you are free to withdraw from the interview at any point. Since there are no right or wrong answers the interviews will be exploratory and aim to uncover a wide range of viewpoints.

If you would like to take part - what should you do now?

If you are interested in participating in this research project or would like more information before making a decision to participate please get in touch with me on any of the contact details listed below and I will be more than happy to discuss the project or arrange a time for us to meet.

Contact details:
Ms Sam Beaumont
Institute of Psychological Sciences
University of Leeds
Leeds
LS2 9JT
Email. psc2sjb@leeds.ac.uk

Appendix VI: Letter of invitation to nurses

Hi,

My name is Sam Beaumont and I am a 2nd year PhD student in the Institute of Psychological Sciences at the University of Leeds. I am conducting research into the possible systems causes of medication errors. This research has been developed because it is my belief that there are many reasons errors occur in medicine – not just because people make mistakes. These causes could be things like the way the ward is set up or the design of tasks or equipment for example (although I am more interested in what you think – rather than what I think!)

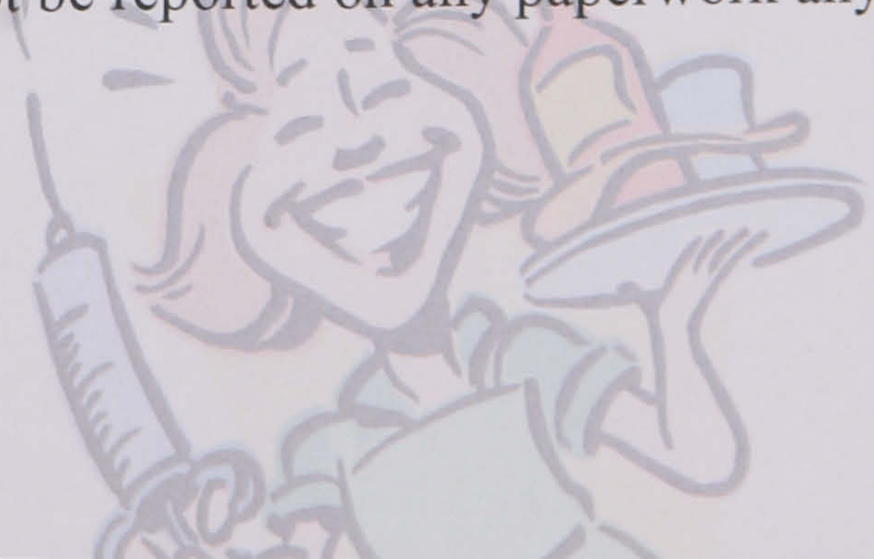
Part of the project involves interviewing nurses working in Bradford hospitals to find out what they believe are important considerations when thinking about medical errors. Your name has been suggested by ..... as someone who has lots of nursing experience and who might have a good insight into the nature of errors.

If you're still not convinced it's a good idea to be involved – here are just a few reasons that might persuade you;

1. The interviews are really informal and just like a chat between me & you really – not like going for a job interview! You are not being tested, as there are no right or wrong answers – just your opinion.

2. The research is important for nurses because finding out what else might have caused error might change the way errors are dealt with in the future.

3. The aim of the project is to get a broad range of views from several nurses on more than one medical ward – not to single individuals out. Your name, position and ward number will not be reported on any paperwork anywhere.



4. As part of being involved I will supply lunch…….from Marks & Spencers!!

5. Last (but by no means least!) ……

You will be entered into a prize draw to win £50!!!!!!!

Names will be drawn out of a hat once the last person has been interviewed and considering only 20 people will be entered you stand an excellent chance of being the lucky winner! There will also be two runners up prizes of wine and chocolates.

What now?

Our interview has been set up for …… at …….. and will last no more than 1 hour (this time has already been agreed by your matron/sister). I will telephone you before the interview (usually on the same day) to find out what you would like for lunch (no caviar or champagne requests please!). The interviews should be fun and should definitely not be a cause for nerves. If you really do not want to take part in the study - don't worry. Just ring me on my mobile (listed below) and let me know before the day. If you want to know a bit more before agreeing to take part you can either ring me or email me & I will be more than happy to give you the details.

Thank you for taking the time to read this information
See you soon!

Sam Beaumont
Email. Psc2sjb@leeds.ac.uk

Appendix VII: Hypothetical error scenarios

1. *Mr Smith was transferred from Ward A to Ward B at **1pm**. Although Mr Smith's medical notes were with him when he arrived, his drug chart was missing. John, the Senior Nurse on Ward B rang Ward A to ask them where the missing drug chart was. The qualified nurse who answered the phone said she was almost certain they had sent Mr Smith's drug chart down with his medical notes. John asked her if Mr Smith had received his **12pm** medication before being transferred to Ward B to which she replied she "wouldn't have thought so", since he was due to be transferred. The nurse told John what medication Mr Smith should have received before being transferred and said they would have another look for the drug chart and send someone down to Ward B if it turned up. John gave Mr Smith the medication he had supposedly not received at lunchtime. At **5pm** when the tea-time medication round was started, Mr Smith's drug chart was returned to Ward B by the Pharmacy department. It was recorded that although the medication John had given to the patient was correct, it had already been administered at lunchtime **(12pm)**. Therefore Mr Smith had received his lunchtime dose twice.*

2. *An operating surgeon prescribed the intravenous administration of Heparin (to prevent blood clots) after a patient's hip surgery and asked the medical technician to fetch a bag of IV Heparin. The medical technician mistakenly took a bag from the place Heparin is usually kept and does not notice that the bag is actually Lidocaine-Hydrochloride not Heparin. Heparin and Lidocaine bags look almost identical and are stored side by side. The anaesthesiologist who accompanied the patient back to the ward also did not notice the wrong bag hanging on the IV drip. Neither the afternoon nor evening nursing staff noticed the error. The following day the Staff Nurse arriving on the early shift noticed the IV bag by the patients' bedside was Lidocaine not Heparin. Whilst administering Lidocaine to a patient who is not prescribed it is not serious, the omission of Heparin is potentially fatal. Later incident analysis revealed that the nursing staff, medical technician and anaesthesiologist who did not notice the error said the IV bag looked the way they expected it to look had it been Heparin.*

3. *A 67-year-old man was admitted to hospital for surgery. He had an allergy to antibiotic drugs, which was noted on his medical record. The surgeon did not notice the information about the allergy and ordered an antibiotic to be given at the end of the surgery. A hospital nurse gave the patient the antibiotic*

4. *When conducting the 9pm medication round on an elderly care ward, the Staff Nurse noticed that a particular patient was not wearing his identification wristband. On this ward, it was correct practice to ensure the I.D. number on the patient's wristband and the number on the drug chart were the same before giving them any medication. On this occasion, there were no other staff members around for the nurse to liaise with and she was reluctant to leave the drug trolley. However, the name on the white board above the patients' bed matched the patient name on the drug chart at the foot of the patient bed, so believing the patient in front of her to be the right patient; the Staff Nurse gave him the medication prescribed on the drug chart. Thirty minutes later, she discovered that this patient was not the patient named on the drug chart. The patient had been confused when returning from the bathroom and had got in the wrong bed by mistake.*

5. *At **9am**, Senior Nurse Julie was trying to find the duty doctor to write a discharge prescription for patient Mrs Brown who was due to be discharged later that day. An emergency admission meant Julie was unable to follow this task up. When Julie had finished with the new admission she had forgotten about the discharge prescription and only remembered later in the day during afternoon handover. Senior Nurse Claire on the afternoon shift arranged for a doctor to attend the ward to write the prescription which she personally took to the pharmacy department at **2.30pm**. When the ambulance arrived to take Mrs Brown home at **4pm** her medication was not ready for her to take since Pharmacy had not had sufficient time to process the request. Claire rang them and asked if they could dispense the medication as a priority. They said they could have the medication ready by **7pm** if the patient could wait a little longer or else she could have someone collect it for her. Mrs Brown would not wait and went home by ambulance. Mrs Brown's relative was not able to collect the medication until **8pm** the following when they had finished work.*

6. *During the morning medication round at **9am**, a patient complained to the nurse on duty that he was in pain and would like some pain killers. The patient had been transferred from A&E at **11pm** in the evening the previous day. Aspirin was listed on the patients drug chart as a possible PRN (prescription required as needed) medication since the patient complained of severe headaches so the nurse gave him two 75mg tablets. Three hours later, admission notes from A&E arrived on the ward and the nurse discovered this patient had been admitted due to suspected severe liver disease. In such cases, aspirin should be avoided due to it increasing the risk of gastro-intestinal bleeding and impaired kidney function. Deterioration of kidney function in these patients can lead to rapid and life-threatening deterioration of their liver disease.*

7. *Mr. Jameson was a patient in an acute ward. One of his prescription charts was rewritten by a doctor who did not usually treat him. The prescription chart was faxed to a pharmacist, Mr. Cryer, who was busy because it was the Friday before a bank holiday weekend. Mr. Cryer dispensed hydralazine (for high blood pressure) instead of hydroxyzine (for itching). The medicine was sent to the ward in a bottle labelled with the patient's name. Mr. Jameson was given the wrong drug three times a day for five days before the error was recognised by another pharmacist, who was checking prescription charts in the ward. All the staff treating Mr Jameson had given him the drug over the five days but no one had noticed that it was the wrong drug.* (Foresight training scenarios, NPSA)

8. *Miss Doyle was being given pethidine via a patient-controlled analgesia (PCA) infusion device. She was on pethidine because she was allergic to morphine. Two junior nurses, Nurse Khan and Nurse James, were changing Miss Doyle's infusion. Neither nurse was competent to carry out this procedure without being supervised by another nurse. There was a record about Miss Doyle's allergy in her notes and she was also wearing an allergy wristband. However, Nurse Khan and Nurse James set up a morphine syringe for Miss Doyle with neither noticing the wristband. In addition, the pethidine had originally been set up using a syringe labelled morphine. The word morphine had been crossed out and pethidine had been written over it.* (Foresight training scenarios, NPSA)

Appendix VIII: interview schedule

Short introduction to the structure of the interview (3 main areas of interest; person, environment, organisation)….give RAIL PROMPTS

A. Person-related questions (active failures)

1. In terms of the *people* described in this scenario – what *actions* do you think could have led to this incident?

- ***Rail example: If a train driver drives through a stop signal and subsequently collides with another train he;***
  - ***might not have noticed the signal*** **[slip/lapse]**
  - ***could have seen the signal but misread it*** **[mistake]**
  - ***may have simply ignored it*** **[violation]**

2. Do you think their actions could *realistically* have been performed any differently? If not, what factors do you think could have affected this?

- ***Rail example: It might not always be possible to simply pay more attention or 'know' something other than what you believe to be true. Tasks might not always be appropriate or achievable according to the 'rules'.***

B. Environment-related questions (local working conditions)

1. Do you think there were any problems relating to the *immediate working conditions* which could have made this error more likely to occur / made this task more difficult?

- ***Rail example: The driver might not have been able to see the stop light properly because it was obscured by overhanging trees/sunlight shining made it difficult to see properly. Driver might have ignored stop signal because they believed it was giving them incorrect information and they usually just drove through them anyway.***

2. Do you think there were any problems with the *specific task* which might have made this error more likely to occur?

- ***Rail example: The driver was supposed to be sent on two more training courses before driving alone but had had to start earlier and not receive the training due to staff shortages.***

C. Organisational-related questions

1. In terms of the workplace factor "x" you mentioned earlier – what do you think the organisational or management factors are which could have contributed to this problem?

- ***Rail example: More closely monitored safety training to make sure all systems were adequate for the job and all staff knew what all signals meant.***

Staff problems with signalling systems were not identified and actioned early enough.

2. Do you think there are any changes which could be made at an organisational level which could prevent this type of error occurring?

3. To what extent do you consider this to be a *global problem* relating to medical errors in general?

4. Can you think of any other examples whereby problems with this organisational factor (name) might result in medical error?

<u>Follow-up questions</u>

- How big a problem do you think this type of error is?
- What type of error do you think this is if you had to classify it?
- Tell me how common you think this problem is in your experience?
- Do you think medication error is a frequent problem?
- Can you give me an example of a medication error you have personal experience of? It does not have to involve you, it could be an error you have observed and simply heard of on the grapevine – you do not have to be involved or disclose anyone's name?

Appendix IX: Interview consent form

Researcher: Samantha Beaumont

Please initial box

1. I confirm that I have read and understand the information sheet dated 18 April 05 (version 3) for the above study and have had the opportunity to ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw from the study at any time, without giving any reason for doing so. ☐

3. I consent to being tape-recorded whilst taking part in the research. I understand that this recording will be confidential, will only be used for the purposes of the research being undertaken and will be destroyed when the research is completed. ☐

4. I agree to take part in the above study. ☐

5. I will not break the confidentiality of any participant concerning information discussed in the interview ☐

Name of Participant	Date	Signature
Researcher	Date	Signature
Name of Participant	Date	Signature

Appendix X: Everyday driving error examples

1. Slips/lapses

- Attempting to drive away from traffic lights in third gear
- Forgetting where you left your car in the multi-storey car park
- Switching on headlights when you meant to turn the wipers on
- Locking yourself out of your car with the keys still inside
- Forgetting to put headlights on when setting off

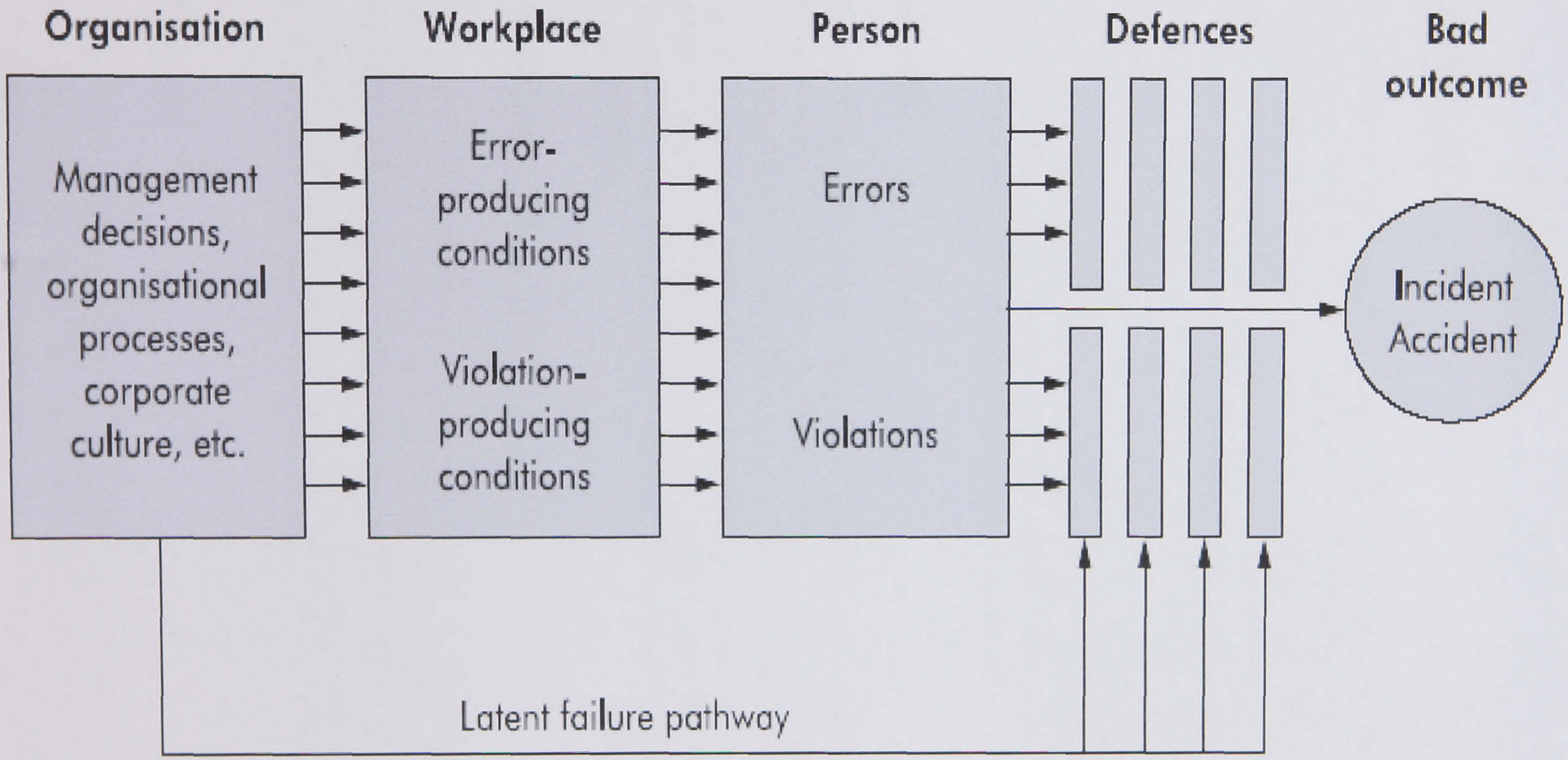
2. Mistakes

- Misjudge your gap in a car park and nearly (or actually) hit another parked car
- Overtake a line of slow moving vehicles and discover they were slowing due to the lane ending ahead
- Getting into the wrong lane at a roundabout
- Steer the wrong way into a skid

3. Violations

- Become impatient with a slow driver in outer lane and undertake (overtake them on inside lane) them to get past
- Driving too close to the car in front or flashing them to move out of the way
- Speeding
- Speeding up when approaching an amber light when you could have stopped in time

Appendix XI: Reason's model of organizational accidents



Appendix XII: Final version of the OSQ

Organisational Safety Questionnaire
Information

What is this study looking at?

This questionnaire aims to assess the way your environment functions. In the same way you might monitor changes in a patient's medical condition for example, this tool will assess the 'health' of your environment so that any future improvements can be most effectively targetted. Organisational or other workplace issues such as human resources, training or team communication which are not running as smoothly as they could be are often early indicators that quality of care or patient safety might be compromised in the future. This questionnaire is one way of monitoring your environment to *prevent* future errors and avoid the potentially serious consequences for patient care and for the health care professionals involved. The questionnaire itself has been developed by researchers at the University of Leeds in conjunction with several nursing teams across Bradford NHS Trust who felt that these workplace issues played a vital role in the likelihood of medication errors.

What will you have to do?

The only thing you need to do is read each statement carefully and circle the appropriate frequency with which each event happens in your environment. It should take you no more than 15-20 minutes to complete. Some statements might be more relevant to your role or unit than others so feel free to circle 'not applicable' if you think you cannot respond. While your participation in this study is entirely voluntary, it would be a good opportunity for you to tell us about your environment so that resources can be most effectively targeted at the most needed improvements in the future.

What will happen to the information you provide?

The information you provide is **completely anonymous** and will be combined and analysed with the rest of your ward or unit to obtain a 'Safety Profile'. **YOU WILL NOT BE PERSONALLY IDENTIFIABLE FROM THIS QUESTIONNAIRE**. We do not ask for your name, only very brief details regarding your unit for the purposes of analysis. Combined ward/unit data will be fed back to Bradford Teaching Hospitals NHS Trust in order to aid future improvements.

What should you do now?

Simply turn over this page and begin completing the questionnaire. When you have finished simply place it into the freepost envelope provided and post it back to me - Sam Beaumont at the University of Leeds ***before Friday 30th November***. If you have any questions about the study please do not hesitate to contact me either by email (psc2sjb@leeds.ac.uk) or telephone (0113 3436680).

Organisational Safety Questionnaire

Please read each statement below carefully and respond by circling the number which best represents the frequency with which each event happens on your ward or department (if you do not work on a ward). Please bear in mind the time scales reported for each question (e.g. in the last 2 weeks, 4 weeks, 2 months etc). If you cannot answer the question because you do not know whether this happens in your department or it is not relvant to you then circle 'not applicable' (N/A).

Not at all	Occasionally	Quite Often	Frequently	Nearly all the time	Not Applicable
1	2	3	4	5	N/A

In the last **4 weeks**, a patient's clinical investigation paperwork (e.g. blood test, ecg report etc.) has been sent to this ward/unit/dept and subsequently lost	1	2	3	4	5	N/A
In the last **2 months**, I have encountered details on a patient's ID wristband which did not correspond with their drug chart	1	2	3	4	5	N/A
In the last **2 months**, I have been unable to coordinate the ward/unit/dept because of my patient allocation	1	2	3	4	5	N/A
In the last **4 weeks**, I have worked a shift when an agency nurse was requested but did not arrive, leaving the ward/unit/dept understaffed	1	2	3	4	5	N/A
In the last **3 months**, I have worked 3 or more 'long shifts' (7am – 9pm) one after another	1	2	3	4	5	N/A
Senior nurses on this ward/unit/dept prefer *all* patient queries to come through them in the first instance – even if there are other nurses on shift who would know the answer	1	2	3	4	5	N/A
When a mistake is made that *could* harm the patient, but does not, how often is this reported on your ward/unit/dept?	1	2	3	4	5	N/A
In the last **4 weeks,** I have given or *almost* given a patient medication I was not sure had been prescribed correctly	1	2	3	4	5	N/A
In the last **3 months**, a patient *from this ward* has been 'slept out' elsewhere but had to return the next day because they were unwell	1	2	3	4	5	N/A
All senior nurses on this ward/unit/dept demonstrate a safe example of patient care for staff to follow	1	2	3	4	5	N/A
In the last **6 months**, I have performed a clinical procedure or used a piece of equipment for which I was not formally trained	1	2	3	4	5	N/A
In the last **4 weeks**, I have discovered a drug chart belonging to one patient in a different patient's nursing/medical notes/bedside locker/bed	1	2	3	4	5	N/A
This ward regularly *uses* a standardised list (**written document**) of tasks which should be carried out for each new admission	1	2	3	4	5	N/A
In the last **4 weeks**, I have sped up during a drug round to avoid a backlog of jobs later in the shift	1	2	3	4	5	N/A

Not at all	Occasionally	Quite Often	Frequently	Nearly all the time	Not Applicable
1	2	3	4	5	N/A

In the last **4 weeks,** I have given or *almost* given a patient the wrong dose of the correct medication	1	2	3	4	5	N/A
In the last **2 months**, as a qualified nurse I have had to leave the ward to accompany a patient to another department leaving the ward short of qualified nurses	1	2	3	4	5	N/A
In the last **2 months**, I have worked 6 or more hours of my shift without anything to eat or drink	1	2	3	4	5	N/A
In the last **2 months**, a senior nurse has complained that jobs from the previous shift were left unfinished	1	2	3	4	5	N/A
I have attended a ward/unit/dept meeting in the last **3 months**	1	2	3	4	5	N/A
This ward *uses* a standardized list (**written document**) of information which *must* be passed on to colleagues during patient handover	1	2	3	4	5	N/A
In the last **2 months**, a doctor has refused to carry out something I needed them to do because the patient concerned was not theirs	1	2	3	4	5	N/A
In the last **2 months**, I have encountered boxed medication which did not correspond with the outer package labelling	1	2	3	4	5	N/A
In the last **2 months**, I have refrained from bleeping a particular doctor because I wasn't confident they had the necessary skills to help me	1	2	3	4	5	N/A
In the last **4 weeks,** I have given or *almost* given a patient a dose of a drug which was 10 times more *or* less than it should have been (e.g. 25mgs instead of 2.5mgs)	1	2	3	4	5	N/A
In the last **12 months**, there has been discussion of a policy or guideline that I was not aware existed until *after* an incident had occurred	1	2	3	4	5	N/A
In the last **3 months**, I have been pressured into arranging to discharge a patient quicker than had been originally planned	1	2	3	4	5	N/A
All of the senior nurses on this ward/unit/dept are approachable if a nurse needs help or advice with a work issue	1	2	3	4	5	N/A
In the last **4 weeks**, I have given or *almost* given a patient medication which was intended for a different patient	1	2	3	4	5	N/A
In the last **6 months**, I have been taught a clinical procedure or how to use a piece of equipment by a nurse who had not received the relevant 'formal' training	1	2	3	4	5	N/A
In the last **12 months**, I have mentored a newly qualified nurse *before* I have attended mentorship and teaching and assessment programs	1	2	3	4	5	N/A
In the last **3 months**, this ward has transferred a patient to another ward during the night or early hours of the morning because their bed was needed for another patient	1	2	3	4	5	N/A
In the last **12 months**, I have signed to say I have read and understood a new policy, procedure or guideline when I haven't either read it or understood it	1	2	3	4	5	N/A

Not at all	Occasionally	Quite Often	Frequently	Nearly all the time	Not Applicable
1	2	3	4	5	N/A

In the last **2 months**, I have been asked by a consultant to carry out a task which was really the job of a junior doctor rather than a nurse	1	2	3	4	5	N/A
In the last **2 months**, I have rushed through outstanding jobs towards the end of my shift so the next shift did not have to do them	1	2	3	4	5	N/A
In the last **2 months**, I have worked with an agency nurse who did not know how to carry out a skill or procedure required of their nursing grade	1	2	3	4	5	N/A
In the last **2 months**, I have encountered a patient on this ward whose allergy wristband was on the opposite wrist to their ID wristband	1	2	3	4	5	N/A
In the last **4 weeks**, I have encountered a patient's allergy status written on their main drug chart but not on their PRN chart	1	2	3	4	5	N/A
In the last **4 weeks**, I have sped up during a drug round to avoid making the next drug round late	1	2	3	4	5	N/A
In the last **2 months**, I have been unable to get hold of IV medication equipment because I could not find it in the hospital or it was all in use on other wards	1	2	3	4	5	N/A
In the last **4 weeks**, I have had to break off from a drug round because I was interrupted by another health professional (e.g. doctor, nurse, physio)	1	2	3	4	5	N/A
In the last **4 weeks**, a patient's drug chart has gone missing from this ward/unit/dept and had to be rewritten	1	2	3	4	5	N/A
In the last **12 months**, I have received an updated version of a written policy, protocol or guideline without knowing which part of the original version had changed	1	2	3	4	5	N/A
In the last **4 weeks,** I have given or *almost* given a patient the same medication twice in the same medication round	1	2	3	4	5	N/A
In the last **12 months**, *all* newly qualified nurses starting on this ward/unit/dept have been given and have worked through a preceptorship pack	1	2	3	4	5	N/A
In the last **2 months**, a senior nurse has advised me to cut corners in order to complete all jobs planned before the end of the shift	1	2	3	4	5	N/A
In the last **12 months**, I have been asked to read a written policy, protocol or guideline but only skimmed it to get the gist because it was too long	1	2	3	4	5	N/A
In the last **4 weeks**, I have given or *almost* given a patient a drug which I later discovered should have been withheld or discontinued	1	2	3	4	5	N/A
In the last **2 months**, I have refrained from bleeping a particular doctor again because I knew they would be annoyed I had bleeped them several times that shift	1	2	3	4	5	N/A
In the last **2 months**, a patient on this ward has missed 3 or more doses of their prescribed medication because it had been ordered but not sent from pharmacy	1	2	3	4	5	N/A

Not at all	Occasionally	Quite Often	Frequently	Nearly all the time	Not Applicable
1	2	3	4	5	N/A

In the last **2 months**, an agency nurse has been sent to this ward/unit/dept who was less qualified than the grade of nurse we requested	1	2	3	4	5	N/A
In the last **3 months**, this ward has had an 'inappropriate admission' because A&E did not want to breach the 4-hour waiting rule	1	2	3	4	5	N/A
In the last **4 weeks**, I have given or *almost* given a patient medication which had expired	1	2	3	4	5	N/A
In the last **4 weeks**, I have encountered a drug chart where the details of a particular drug (e.g. dosage, frequency/route of administration) had been crossed out and altered 2 or more times	1	2	3	4	5	N/A
In the last **3 months**, I have made or witnessed an error which I did not 'formally' report because the patient wasn't harmed by it in any way	1	2	3	4	5	N/A
In the last **3 months**, this ward has had to rush an existing patient's discharge to admit a new patient from MAU or A&E to avoid a patient breaching the 4 hour rule in A&E	1	2	3	4	5	N/A
In the last **2 months**, a patient from this ward/unit/dept has been discharged without taking home their medication because it had been ordered but not been sent by pharmacy	1	2	3	4	5	N/A
In the last **4 weeks**, the number of doctors arriving for morning ward round has slowed my drug round progress	1	2	3	4	5	N/A
In the last **4 weeks**, I have sought assistance from a colleague to interpret the handwriting on a patient's drug chart	1	2	3	4	5	N/A
In the last **4 weeks**, I have infused or *almost* infused IV fluid over an incorrect time period (e.g. over 8 hours instead of 16)	1	2	3	4	5	N/A
In last **6 months**, I have formally reported a safety concern on this ward/unit/dept without receiving any feedback	1	2	3	4	5	N/A
In the last **2 months**, I have found it difficult to prioritize duties for the day because all tasks seemed equally important	1	2	3	4	5	N/A
When a mistake is made, but is *caught and corrected before affecting the patient*, how often is this reported on your ward/unit/dept?	1	2	3	4	5	N/A
In the last **2 months**, I have been confused as to what jobs were supposed to be my responsibility during the course of a shift	1	2	3	4	5	N/A
In the last **2 months**, 2 or more new admissions have arrived within minutes of each other on this ward	1	2	3	4	5	N/A
In the last **2 months**, I have worked an entire shift without a rest break	1	2	3	4	5	N/A
In the last **4 weeks**, a doctor or other health professional has written instructions in a patients notes for the nursing team to carry out without verbally advising them they had done so	1	2	3	4	5	N/A
In the last **4 weeks**, I have given or *almost* given a patient a drug that they were allergic to (e.g. Heliclear with penicillin allergy)	1	2	3	4	5	N/A

Not at all	Occasionally	Quite Often	Frequently	Nearly all the time	Not Applicable
1	2	3	4	5	N/A

In the last **3 months**, I have encountered medication (e.g. boxed or IV) in the drug trolley or clinic room which had expired by more than a month	1	2	3	4	5	N/A
In the last **3 months**, a patient has been discharged from this ward before all the necessary agencies needed to care for them at home (e.g. community nurse, social worker) had been arranged	1	2	3	4	5	N/A
In the last **2 months**, senior nurses have been rushing around the ward/unit/dept completing most tasks while other nurses have less to do	1	2	3	4	5	N/A
In the last **6 months**, I have booked a training course or study day which was cancelled due to budgetary constraints	1	2	3	4	5	N/A
In the last **4 weeks**, I have withheld or *almost* withheld a drug which should have been administered	1	2	3	4	5	N/A
In the last **4 weeks**, I have encountered drugs packaging for two different types of medication which was virtually identical	1	2	3	4	5	N/A
In the last **6 months**, there has been a deterioration in the level of experience within the team on this ward/unit/dept	1	2	3	4	5	N/A
In the last **2 months**, I have had to telephone the ward after my shift had ended because I remembered information I had previously forgotten to pass on during handover	1	2	3	4	5	N/A
In the last **4 weeks**, I have bleeped a doctor 2 times or more without receiving a response	1	2	3	4	5	N/A
In the last **3 months**, I have followed instructions against my better judgement because a doctor or more senior nurse told me it was what I should do	1	2	3	4	5	N/A
In the last **12 months**, I have encountered a written policy or procedure that I thought would be impossible to follow in the course of my job	1	2	3	4	5	N/A
In the last **2 months**, there has been an important task outstanding at the end of the shift which I didn't do because I thought another nurse had done it	1	2	3	4	5	N/A
In the last **4 weeks**, I have given or *almost* given a patient someone else's medication to take home upon discharge	1	2	3	4	5	N/A
In the last **2 months**, a doctor on this ward has refused to wait until I was finished with a particular task before conducting their ward round without me	1	2	3	4	5	N/A
In the last **6 months**, there has been an increase in senior nursing turnover on this ward/unit/dept	1	2	3	4	5	N/A
In the last **2 months**, I have experienced difficulty writing all the information passed on to me during patient handover because I didn't hear or understand everything or there wasn't time to write everything	1	2	3	4	5	N/A
In the last **2 months**, I have encountered a patient who had been on the ward for more than 24 hours who did not have *this ward's* ID wristband on	1	2	3	4	5	N/A

Not at all	Occasionally	Quite Often	Frequently	Nearly all the time	Not Applicable
1	2	3	4	5	N/A

There has been an incident debriefing or team meeting to discuss the *last* 'significant' clinical incident to happen on this ward/unit/dept within one month of its occurrence	1	2	3	4	5	N/A
In the last **4 weeks**, I have infused or *almost* infused an *incorrect volume* of IV fluid to a patient (e.g. 75mls instead of 50mls)	1	2	3	4	5	N/A
In the last **4 weeks**, I have experienced difficulty getting an on-call doctor to come to the ward out of hours	1	2	3	4	5	N/A
In the last **3 months**, a patient has been 'slept out' *to this ward* who needed specialised care I thought this ward could not effectively deliver	1	2	3	4	5	N/A
In the last **4 weeks**, I have encountered a patient's drug chart which had been changed by a doctor since the previous drug round without my knowledge	1	2	3	4	5	N/A
In the last **2 months**, I have encountered a patient on this ward wearing an allergy wristband whose allergy was not documented on their drug chart or vice versa	1	2	3	4	5	N/A
In the last **3 months**, I have experienced difficulty challenging a doctor about a prescription I thought they had written incorrectly	1	2	3	4	5	N/A
In the last **6 months**, I have booked a training course or study day which was cancelled due to a clash with work requirements or staff shortages	1	2	3	4	5	N/A
In the last **2 months**, I have come across details written on the board above a patient's bed for the patient who had *previously* occupied that bed	1	2	3	4	5	N/A
In the last **4 weeks**, I have given or *almost* given medication to a patient too soon after their last dose (e.g. only 4 hours after previous dose instead of 8 hours)	1	2	3	4	5	N/A
When a mistake is made, but has *no potential to harm the patient*, how often is this reported on your ward/unit/dept?	1	2	3	4	5	N/A
In the last **12 months**, all the reasons have been made clear why a new procedure or policy has been introduced	1	2	3	4	5	N/A
In the last **12 months**, *all* newly qualified nurses starting on this ward/unit/dept have received the standard induction training programme required	1	2	3	4	5	N/A
In the last **4 weeks**, an *unexpected* new admission on this ward has meant that other important tasks were incomplete by the end of the shift	1	2	3	4	5	N/A

Not at all	To a very limited extent	To a limited extent	To a moderate extent	To a considerable extent	To a great extent	To a very great extent
1	2	3	4	5	6	7

We have a good "map" of each other's talents and skills	1	2	3	4	5	6	7
We talk about mistakes and ways to learn from them	1	2	3	4	5	6	7
We discuss our unique skills with each other so we know who on the unit has relevant specialized skills and knowledge	1	2	3	4	5	6	7
We discuss alternatives as to how to go about our normal work activities	1	2	3	4	5	6	7
When giving report to an oncoming nurse, we usually discuss what to look out for	1	2	3	4	5	6	7
When attempting to resolve a problem, we take advantage of the unique skills of our colleagues	1	2	3	4	5	6	7
We spend time identifying activities we do not want to go wrong	1	2	3	4	5	6	7
When errors happen, we discuss how we could have prevented them	1	2	3	4	5	6	7
When a patient crisis occurs, we rapidly pool our collective expertise to attempt to resolve it	1	2	3	4	5	6	7

Background information

- **Please give your work area/unit in this hospital an overall grade on patient safety. Mark ONE answer only.**

❑ Failing ❑ Poor ❑ Acceptable ❑ Very good ❑ Excellent

- **<u>In the last 12 months</u>, how many incident reports have <u>you personally</u> filled out and submitted?**

❑ None ❑ 1 to 2 ❑ 3 to 5 ❑ 6 to 10 ❑ 11 to 20 ❑ 21 or more

- **How long have you worked in <u>your current hospital work area/unit</u>?**

❑ Less than 1 year ❑ 1 to 5 years ❑ 6 to 10 years ❑ 11 to 15 years ❑ 16 to 20 years ❑ 21 years or more

- **Typically, how many hours per week do you work in this hospital?**

❑ Less than 10 hours per week ❑ 10 to 19 hours per week ❑ 20 to 39 hours per week
❑ 40 to 59 hours per week ❑ 60 to 79 hours per week ❑ 80 hours or more

- **What is your staff position in this hospital?**

❑ Matron ❑ Sister ❑ Senior Staff Nurse ❑ Staff Nurse ❑ Student Nurse

❑ HCA / NA ❑ Midwife ❑ Other, please specify:.....

- **Do you administer medication to patients as part of your nursing role?** YES / NO

Please feel free to write any comments about patient safety, error, or incident reporting in your hospital (attach extra sheet if space allowed is insufficient).

We would like to thank you for completing this questionnaire.

Appendix XIII – DRBQ participant information sheet

Drug Round Behaviour Questionnaire
Information

What is this study looking at?

While it is important to evaluate significant clinical incidents which directly affect patient safety, it is also important to keep track of smaller 'events' which might occur more frequently but generally go unnoticed. Such events may often seem insignificant at the time, rarely resulting in a bad outcome for the patient. However they are often symptomatic of future large scale incidents. This short questionnaire has been developed as a way of monitoring these smaller events which might occur whilst you undertake your routine drug round. These events could be in the form of a memory lapse, concentration or attention difficulties. They might also represent a situation where you acted in a way you thought was right at the time but which you later discovered was incorrect. Finally, an event could be a situation where you performed an action in a way you knew was not exactly how it should be done believing the 'rules' to be unsuitable or irrelevant to the circumstances. While these events may not have resulted in a bad outcome for the patient (or any outcome at all in some cases) it is important that such events are monitored in order to prevent more potentially serious (for the patient and for the nurses involved) incidents in the future.

What will you have to do?

The only thing you need to do is read each question carefully and circle the frequency with which you think it happens to you. It should take you approx 5-10 minutes to complete.

What will happen to the information you provide?

You do not need to give us your name as the information you give us is ***completely anonymous***. We only ask that you provide us with a few simple demographic details at the end of the questionnaire to help us to analyze the data accurately. Individual responses will be viewed only by researchers at the University of Leeds' Institute of Psychological Sciences. These responses will then be combined to produce an average 'Medication Round Behaviour Profile' for your ward. You will not be personally identifiable from this questionnaire. Combined ward data will be fed back to Bradford Teaching hospitals NHS Trust.

What should you do now?

Simply turn over this page and begin completing the questionnaire. If you have any questions about the study please do not hesitate to contact Sam Beaumont at the University of Leeds either by email (psc2sjb@leeds.ac.uk).

Drug Round Behaviour Questionnaire
TASK 1 - ANSWER SHEET

For each of the 27 questions on the Drug Round Behaviour Questionnaire please consider the following 3 issues;

1. Comprehension – is the question clear and easy to understand? If no, please provide details.

2. Frequency – would you be able to rate the number of times this behaviour happens to you? (answer yes or no)

3. Examples – can you think of examples of when this might have occurred either to you *or to someone else*? (you do not need to provide examples here, simply answer yes or no)

	1.Comprehension (if no, provide details)	2. Frequency	3. Examples
Q1	Y / N	Y / N	Y / N
Q2	Y / N	Y / N	Y / N
Q3	Y / N	Y / N	Y / N
Q4	Y / N	Y / N	Y / N
Q5	Y / N	Y / N	Y / N
Q6	Y / N	Y / N	Y / N
Q7	Y / N	Y / N	Y / N
Q8	Y / N	Y / N	Y / N
Q9	Y / N	Y / N	Y / N
Q10	Y / N	Y / N	Y / N

	1.Comprehension (if no, provide details)	2. Frequency	3. Examples
Q11	Y / N	Y / N	Y / N
Q12	Y / N	Y / N	Y / N
Q13	Y / N	Y / N	Y / N
Q14	Y / N	Y / N	Y / N
Q15	Y / N	Y / N	Y / N
Q16	Y / N	Y / N	Y / N
Q17	Y / N	Y / N	Y / N
Q18	Y / N	Y / N	Y / N
Q19	Y / N	Y / N	Y / N
Q20	Y / N	Y / N	Y / N
Q21	Y / N	Y / N	Y / N
Q22	Y / N	Y / N	Y / N
Q23	Y / N	Y / N	Y / N
Q24	Y / N	Y / N	Y / N
Q25	Y / N	Y / N	Y / N
Q26	Y / N	Y / N	Y / N
Q27	Y / N	Y / N	Y / N

Please write here any comments you would like to make about the Drug Round Behaviour Questionnaire in general, i.e. structure, clarity, wording, formatting or more general aspects of its use. All your comments will be **completely anonymous** so feel free to be as honest as you like!

We would like to thank you for being involved with this research.

Appendix XV: Personalised OSQ cover letter for Study 4

Sam Beaumont MSc
Institute of Psychological Sciences
University of Leeds
LS2 9JT
Email: psc2sjb@leeds.ac.uk
Tel no: 0113 3436680

Dear Anne,

Please find enclosed the **Organisational Safety Questionnaire** and FREEPOST return envelope. This questionnaire aims to measure particular aspects of your working environment (e.g. workload, human resources, policies etc.) which are most likely to make everyday tasks more prone to error. **You will not be personally identifiable from the information you provide on this questionnaire.**

You will have already received this questionnaire a few weeks ago. Unfortunately, several staff members felt that since they did not strictly work on a ward it was not applicable to them. At this stage I need to find out *exactly which questions are relevant for each ward / unit or department.* **If you do not work on a ward**, there will be other less specific questions on the questionnaire which may be relevant to your area of work, such as training issues or your relationship with other departments and health professionals for example.

Although involvement in this research is voluntary, please be aware that the information you provide on this questionnaire is **completely anonymous**. This means I do not know who has or hasn't already returned their completed questionnaires. If you have already done so, I would like to thank you for your time and effort. I realise how difficult it can be fitting extra activities into your schedule. If you have yet to fill in a questionnaire and you wish to be involved please return completed questionnaires to me, Sam Beaumont at the University of Leeds in the freepost envelope provided by **Friday 30th November**.

Kind regards

Sam Beaumont

Appendix XVI: Item loadings, mean scores & inter-item correlations on 28-item Organisational Safety Questionnaire

Item	Item code	Item loading	Mean (Std)	Corrected item-total correlation
In the last **3 months**, I have been pressured into arranging to discharge a patient quicker than had been originally planned	BED2	.49	1.90 (1.09)	.48
In the last **3 months**, this ward has transferred a patient to another ward during the night or early hours of the morning because their bed was needed for another patient	BED3	.44	2.34 (1.42)	.43
In the last **2 months**, a senior nurse has complained that jobs from the previous shift were left unfinished	CLT2	.58	2.26 (1.07)	.51
In the last **2 months**, a doctor on this ward has refused to wait until I was finished with a particular task before conducting their ward round without me	CLT11	.47	1.68 (.93)	.44
In the last **2 months**, I have rushed through outstanding jobs towards the end of my shift so the next shift did not have to do them	CLT5	.56	2.62 (1.27)	.51
In the last **6 months**, I have formally reported a safety concern on this ward/unit/dept without receiving any feedback	CLT9	.47	1.48 (.88)	.45
In the last **4 weeks**, I have worked a shift when an agency nurse was requested but did not arrive, leaving the ward/unit/dept understaffed	HR1	.55	2.05 (1.02)	.57
In the last **2 months**, as a qualified nurse I have had to leave the ward to accompany a patient to another department leaving the ward short of qualified nurses	HR2	.42	2.55 (1.32)	.44
In the last **2 months**, I have worked with an agency nurse who did not know how to carry out a skill or procedure required of their nursing grade	HR4	.56	2.33 (1.31)	.54
In the last **6 months**, there has been a deterioration in the level of experience within the team on this ward/unit/dept	HR6	.57	2.18 (1.24)	.57

Appendix XVI: continued

Item	Item code	Item loading	Mean (Std)	Corrected item-total correlation
In the last **2 months**, I have been unable to get hold of IV medication equipment because I could not find it in the hospital or it was all in use on other wards	LWC4	.50	1.60 (.90)	.47
In the last **2 months**, a patient from this ward/unit/dept has been discharged without taking home their medication because it had been ordered but not been sent by pharmacy	LWC6	.52	2.52 (1.27)	.50
In the last **2 months**, I have worked an entire shift without a rest break	LWC7	.45	2.11 (1.25)	.34
In the last **4 weeks**, I have encountered drugs packaging for two different types of medication which was virtually identical	LWC9	.50	1.92 (1.08)	.42
In the last **2 months**, I have had to telephone the ward after my shift had ended because I remembered information I had previously forgotten to pass on during handover	RP5	.51	1.93 (.90)	.47
In the last **2 months**, I have encountered a patient who had been on the ward for more than 24 hours who did not have *this ward's* ID wristband on	RP7	.54	1.91 (1.01)	.55
In the last **2 months**, senior nurses have been rushing around the ward/unit/dept completing most tasks while other nurses have less to do	SP5	.44	1.96 (1.12)	.46
In the last **2 months**, there has been an important task outstanding at the end of the shift which I didn't do because I thought another nurse had done it	SP6	.43	1.33 (.55)	.37
In the last **4 weeks**, I have encountered a patient's drug chart which had been changed by a doctor since the previous drug round without my knowledge	TC10	.70	2.13 (1.01)	.68
In the last **4 weeks**, I have discovered a drug chart belonging to one patient in a different patient's nursing/ medical notes/bedside locker/bed	TC2	.41	1.61 (.76)	.33
In the last **4 weeks**, I have encountered a patient's allergy status written on their main drug chart but not on their PRN chart	TC4	.46	2.47 (1.24)	.50

Appendix XVI: continued

Item	Item code	Item loading	Mean (Std)	Corrected item-total correlation
In the last **4 weeks**, a patient's drug chart has gone missing from this ward/unit/dept and had to be rewritten	TC5	.52	1.84 (.89)	.44
In the last **4 weeks**, I have encountered a drug chart where the details of a particular drug (e.g. dosage, frequency/route of administration) had been crossed out and altered 2 or more times	TC6	.46	1.56 (.66)	.45
In the last **4 weeks**, a doctor or other health professional has written instructions in a patients notes for the nursing team to carry out without verbally advising them they had done so	TC8	.62	2.62 (1.16)	.55
In the last **2 months**, I have found it difficult to prioritize duties for the day because all tasks seemed equally important	WL5	.57	2.10 (1.07)	.53
In the last **4 weeks**, an *unexpected* new admission on this ward has meant that other important tasks were incomplete by the end of the shift	WL7	.60	2.30 (1.05)	.54
In the last **12 months**, there has been discussion of a policy or guideline that I was not aware existed until *after* an incident had occurred	WPP1	.46	1.32 (.57)	.42
In the last **12 months**, I have encountered a written policy or procedure that I thought would be impossible to follow in the course of my job	WPP5	.42	1.60 (.86)	.38