PREDICTING HUMAN BEHAVIOUR IN EMERGENCIES

GLYN LAWSON, BSc(Hons) MSc MIEHF AHEA

**Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy**

JULY 2011

Abstract

The outcome of an emergency is largely determined by the behaviour of the people involved. To improve the safety of buildings and to increase the effectiveness of response procedures and training programmes it is often necessary to predict human behaviour in emergency situations. There are several approaches which can be used to make these predictions, but not all had previously been systematically analysed and therefore their appropriateness for any given application was unknown.

This thesis describes an analysis of approaches for predicting human behaviour in emergencies. The research focussed on approaches which could be used by human factors professionals to extend the contribution this systems-oriented and user-focussed discipline can make to managing risks and reducing danger. The investigated approaches were evaluated against criteria for judging their quality, including validity, reliability, resources, sensitivity and ethics.

In research conducted to test the approaches, fire drills, virtual environments (VEs) and a new talk-through approach, in which participants describe the hypothetical actions they would take in an emergency scenario, demonstrated potential for predicting behaviour in emergency situations. These approaches were subsequently evaluated in a standardised comparison, in which each one was applied to analyse the behaviour demonstrated during an evacuation from a university building. The observed frequencies of behaviour produced by each approach were significantly correlated, as were the sequences of behaviour. All of the approaches demonstrated replicability. The resources required to apply each approach were relatively low, especially for the talk-through approach.

Based on the findings from this research, and drawing upon previous work from the scientific literature, guidance was provided for selecting approaches and methods for behavioural prediction in emergency situations. The talk-through approach is suitable for use during the concept phase of a design as it is quick to implement and requires low resources. VEs and simulation tools are more appropriate for design activities when detailed CAD models become available. Fire drills can provide useful measures of human behaviour in evacuation scenarios, but require a physical representation of the building or environment under investigation. Fire drills, VEs and simulation tools can be used to inform emergency response procedures. Predictions from all of the aforementioned approaches can support the development of training programmes. This guidance was previously unavailable to human factors professionals and now serves both to inform design work and support the evaluation of existing evacuation procedures and protocols.

Publications

The following publications have been written within the timescale of this PhD research:

Book chapters

Lawson, G., and D'Cruz, M. (in press). Ergonomics methods and the digital factory. Accepted for publication in: L. Canetta, C. Redaelli, and M. Flores (Eds.), *Intelligent Manufacturing Systems DiFac*, Heidelberg: Springer.

Conference proceedings

Lawson, G., Sharples, S., Clarke, D., and Cobb, S. (2009). Development of a technique for predicting the human response to an emergency situation. In: D Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics, Proceedings of the 8th International Conference, EPCE 2009*, *HCI International 2009.* San Diego 19-24 July 2009. Heidelberg: Springer, pp.22-31.

Lawson, G., Sharples, S., Clarke, D., and Cobb, S. (2009). The use of experts for predicting human behaviour in fire. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.493-500.

Lawson, G., Sharples, S., Cobb, S., and Clarke, D. (2009). Predicting the human response to an emergency. In: P.D. Bust (Ed.), *Contemporary Ergonomics 2009, Proceedings of the Ergonomics Society's Annual Conference*. London 22-23 April 2009. London: Taylor and Francis, pp.525-532.

Karasova, V., and Lawson. G. (2008). Methods for predicting human behaviour in emergencies: an analysis of scientific literature. In: *Proceedings of Geographical Science Research UK 2008, GISRUK*, Manchester 2-4 April 2008. pp.217-221.

Lawson, G., D'Cruz, M., Bourguignon, D., and Pentenrieder, K. (2007). Training in the Digital Factory. Paper presented at the *IFAC Workshop on Manufacturing, Modelling, Management and Control.* Budapest, 14-16 November 2007.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Dr Sarah Sharples, Professor David Clarke, and Dr Sue Cobb, who have been unsurpassed sources of knowledge, support and inspiration throughout my research.

I am also thankful to my examiners, Professor John Wilson and Dr Shamus Smith, for a supportive viva and for their positive feedback on my thesis.

I would like to acknowledge the support from the staff and students of the Human Factors Research Group at the University of Nottingham. While many members of the group contributed to my research in different ways, particular thanks go to Dr Mirabelle D'Cruz for her help in obtaining my PhD. I would also like to give particular thanks to Anne Floyde and Kirstie Dane for their administrative support.

Acknowledgements go to the DiFac (IST-5-035079) EU-funded research project, from which the idea for this thesis arose and which provided the foundations of much of my further research. I owe particular gratitude to my colleague Dr David Bourguignon, who offered encouragement and advice throughout and beyond DiFac.

I would also like to acknowledge Eric Jones of Nottinghamshire Fire and Rescue Services for his contributions to my research.

My gratitude goes to all of my family and friends who have supported and encouraged me during the last four years, with particular thanks to Renze Lowrie, Damyan Velkov, Dave Richards, Claire Michel and Catherine Coles.

Finally, I remain indebted to Dr Rose Saikayasit for her limitless support and encouragement throughout the entirety of my PhD.

Contents

List of figures

List of tables

1. Introduction

1.1 Chapter overview

This chapter provides an introduction to this thesis, including its origins in project work at the Human Factors Research Group at the University of Nottingham. The aim of the research, to investigate approaches for predicting human behaviour in emergency situations, is explained. To achieve this aim, the work involved the development and analysis of a selection of approaches which were applied to the same scenario to compare their suitability for use by human factors professionals. An overview of the specific studies and investigations is presented, as is the structure of the thesis.

1.2 Background

Human factors (HF) is a discipline concerned with researching the capabilities and characteristics of human beings and applying this knowledge to improve the design of products and systems with which they interact (Wilson, 2005). Identifying and resolving human factors issues early in the development phase can reduce the costs associated with change following investment into specific lines of engineering (Laughery, 2005). A proactive approach to human factors can also more effectively optimise human performance by preventing rather than reacting to incidents or injuries (Haag, 1992; Reason, 2000). These aims can be achieved by predicting human performance during interaction with the system, ranging from the physical fit of equipment and workplaces to the person (Pheasant, 1996) to cognitive and behavioural aspects (Kirwan and Ainsworth, 1992). However, publications in the HF literature are often concerned with predicting human performance during typical tasks with the aim of supporting system design (e.g. Kieras and Meyer, 2000) or predicting the likelihood of human error in an attempt to avoid a catastrophic situation (e.g. Kirwan, 1994). The applicability of these approaches to emergency situations is unclear, and may not be appropriate. They tend to rely on knowledge of the tasks people must perform within a system, which is not necessarily understood for an emergency situation. For example, assuming in an emergency that the primary goal is safe evacuation, people have displayed behaviours which do not support this goal, such as shutting down a computer before evacuating (Gershon et al., 2007). Thus, for emergency situations identifying what actions people would take could in fact be one of the main motivations for conducting the analysis.

There is recognition that systems engineering principles, as used in human factors/ergonomics, can make useful contributions to an understanding of human behaviour and safety in fire and emergencies (Sime, 1995; Pauls, 1999; Shields and Proulx, 2000; Zachary Au and Gilroy 2009). Working in this area, ergonomists would use a variety of approaches for predicting behaviour in emergency situations to help guide the design of buildings, aircraft and boats as well as evacuation procedures and training for emergency response (Meguro et al., 1998; Brooks et al., 2001; Purser and Bensilum, 2001; Perry and Lindell, 2003; Kanno et al., 2006; Lawson et al., 2007b; Pentenrieder et al., 2007; Chittaro and Ranon, 2009; Deere et al., 2009; Tubbs and Meacham, 2009). To do this effectively, they would need to understand fully the approaches available (Annett, 2002). Approaches and methods which ergonomists and HF professionals can use to predict human behaviour in emergencies is the main topic of this thesis.

The requirement for further research into approaches for predicting human behaviour in emergencies was identified during two separate projects with which the author became involved while working as a Research Fellow at the University of Nottingham. For both projects it was realised that existing knowledge and approaches were insufficient.

First, the EU-funded DiFac project (Digital Factory for Human-Oriented Production System: IST-5-035079) aimed to develop a desktop simulator for training emergency response teams in the correct procedures during a factory fire. The simulator was to include avatars which displayed realistic human behaviour to increase the effectiveness of the training simulator (Lawson et al., 2007b). However, it became apparent during development that no suitable established method for predicting and modelling human behaviour in emergencies existed within the human factors discipline. Approaches used in other disciplines, such as developing simulation tools based on models of behaviour published in the scientific literature, were also problematic (Cornwall et al., 2002).

Second, a local emergency response team was interested in predicting human behaviour during a large-scale emergency response scenario. They had conducted logistical predictions concerning the suitability of their equipment, but were also interested in finding out in more detail the human behavioural response to the event. They wanted specific, quantifiable predictions about the public's likely actions. The predictions could not be based on previous events as the scenario was rare and incident reports were difficult to obtain. Furthermore, theories published in the scientific literature were insufficiently quantified or detailed to satisfy their requirements (e.g. Fischer, 2002, 2003; Mawson, 2005).

The research work conducted for this thesis developed and analysed approaches for behavioural prediction, based on the practical human factors applications from which it originated. It evolved to investigate a variety of approaches, determining their strengths, weakness and limitations. It was completed under part-time registration while the author continued to work in the Human Factors Research Group at the University of Nottingham.

1.3 Aims and objectives

This research investigated approaches and methods for predicting human behaviour in emergencies. Throughout the research, consideration was given to the purpose of the prediction, for example improving the design of buildings to support safe evacuation and for developing emergency response procedures. The work was based on the following aims:

Aim 1: To analyse approaches for predicting human behaviour in emergencies against established criteria for assessing their quality. The criteria, taken from the Human Factors (HF) literature (Chapter 3), were used to determine the strengths, weaknesses and limitations of the approaches for HF applications, such as the design of buildings and structures, and developing training courses and emergency response procedures. The use of established criteria was to ensure the assessment was thorough, and that the approaches were suitable for HF professionals.

This aim was addressed iteratively throughout the course of the research, as the outcome of each study was reviewed against the criteria. This begun with a review of the approaches used in previous studies by other researchers (Chapters 2 and 3), followed by an evaluation of all studies conducted for this thesis.

Aim 2: To develop new approaches for predicting human behaviour in emergency situations. New approaches for predicting human behaviour were to be developed in response to any identified opportunities or requirements for them. In particular, development work was carried out to address shortfalls in the performance of previous approaches against the established criteria as identified in the evaluation process conducted for Aim 1. The newly developed approaches would themselves be subject to an evaluation against the criteria for judging the quality of an approach, in an iterative process of optimisation.

The studies which were conducted to satisfy this aim are presented in Chapters 5 and 7. Chapter 5 presents a new approach for predicting behaviour in which experts make judgements for the likelihood of behaviour, using the paired comparisons technique. Chapter 7 details three development studies on a new approach in which participants are asked to describe their anticipated actions in response to a hypothetical emergency situation.

Aim 3: To make a systematic comparison of approaches for predicting human behaviour in emergencies. The selection of approaches was to be based on the outcome of Aims 1 & 2. Thus, both existing and new approaches for predicting human behaviour in emergencies were to be evaluated in a systematic comparison. This was to obtain a comparative review of the approaches, by applying them to the same (standardised) scenario. It also enabled a detailed assessment against the criteria for judging their quality (Aim 1).

A standardised comparison of fire drills, virtual environments, the talk-through approach and the use of literature for predicting human behaviour in emergencies is presented in Chapter 8.

Aim 4: To develop recommendations and guidance for HF professionals responsible for behavioural predictions in emergency situations. An important intended outcome of the work was recommendations for the selection and administration of approaches and methods. The recommendations (Chapter 9) were written to provide users with guidance based on the desired data/measures but also factors which may affect their prediction, such as limited time or budget. The recommendations were derived not only from the studies conducted for the thesis, but were augmented with a review of previous research available from the scientific literature.

The relationship between the aims is shown graphically in Figure 1.1 below:

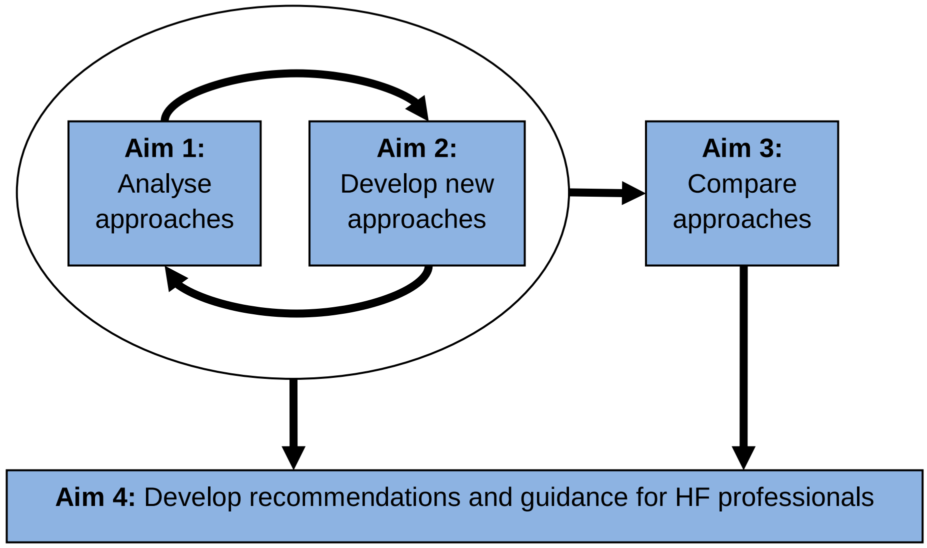

Figure 1.1. Relationship between the research aims

Thus, the research aimed to develop and analyse approaches for predicting human behaviour in emergencies. The primary anticipated users of the knowledge presented in this thesis are HF professionals, due to the contribution this field can make to human safety and emergency planning, as discussed in Section 1.2. Furthermore, the lack of existing guidance on predictive approaches (Section 1.2) and reports that HF professionals do not always consider sufficiently the quality of the methods that they use (Annett, 2002) further justifies this work. However, the output is also likely to prove useful to those from other disciplines such as fire safety and emergency preparedness, given that many approaches used in these areas for behavioural prediction had not previously been fully evaluated against criteria for judging their quality (Chapter 2).

In the course of this work data were collected on the actual behaviours people demonstrated in a range of emergency situations. These data were analysed only to support further understanding of the approaches and methods.

1.4 Novel contribution

Chapter 2 demonstrates that not all approaches for predicting behaviour had previously been systematically analysed. Therefore, the analysis of approaches within this thesis contributes to an understanding of their quality and value, and is one of the main novel contributions to knowledge.

The approaches for predicting behaviour presented in Chapters 5 (expert predictions) and 7 (talk-through approach) were developed during the work conducted for this thesis. These new approaches, while drawing on existing methods, had never previously been applied to generate predictions of human behaviour in emergencies, thus providing further novel contribution.

Chapter 2 demonstrates that while previous researchers have conducted comparisons of approaches and methods, (e.g. Olsson and Regan, 2001; Kuligowski, 2003; Gwynne et al., 2005; Christoffersen and Soderlind, 2009), none had conducted a comparison of fire drills, VEs, the talk-through approach and the use of literature. These approaches were used for the standardised comparison in Chapter 8. Furthermore, no previous research had evaluated these approaches against the criteria used in this thesis, in particular the range of measures used to evaluate the validity of the approaches.

During the project work described in the background section (1.2), the author found little guidance in the Human Factors literature when attempting to predict human behaviour in emergencies. Thus a further novel contribution is the development of guidance for HF practitioners when selecting a predictive approach. This novel contribution, associated with Aim 4, is provided in Chapter 9.

The novel contribution of this work is revisited in the discussion (Chapter 9), to review more fully the achievements and novel contribution against previous research.

1.5 Definition of approaches, methods, and measures

It is pertinent to define the terms approaches, methods, and measures as used within this thesis, particularly as there are differences in their use by different authors. For example, Fahy (2005) describes “laboratory experiments”, “post-incident surveys and interviews” and “videotaped observations” as types of *method*. Wilson (2005) provides a more specific taxonomy of methods, techniques and measures/outcome, for example direct observation is described as a *method*, human recording is a *technique* and event frequency is a *measure* or *outcome*.

Research in human behaviour in emergencies often consists of a *setting* to which a *method* is applied. To apply this definition to the examples given by Fahy (2005) above, post-incident surveys and interviews would involve a *setting* (the incident) to which *methods* (surveys and interviews) are applied. Similarly, videotaped observations would involve a *setting* (the evacuation) to which a *method* (observation) is applied. In this thesis, the application of a *setting* and *method* (or *methods*) to achieve a particular *measure* is defined as an *approach*. Methods themselves are further categorised as *data collection methods* and *data analysis methods* where appropriate.

1.6 Thesis overview

This section presents a general overview of the structure of the thesis, and a description of each contributory chapter. The organisation of the research is illustrated in Figure 1.2. The research was guided initially by a review of approaches which have previously been used for analysing human behaviour in emergencies. Thereafter, it was separated into two main sections:

1. **Phase 1: tests of approaches.** This research aimed to investigate and develop approaches for predicting human behaviour in emergencies.

2. **Phase 2: standardised methods comparison.** This contains a systematic comparison of several approaches for predicting behaviour using the same emergency scenario.

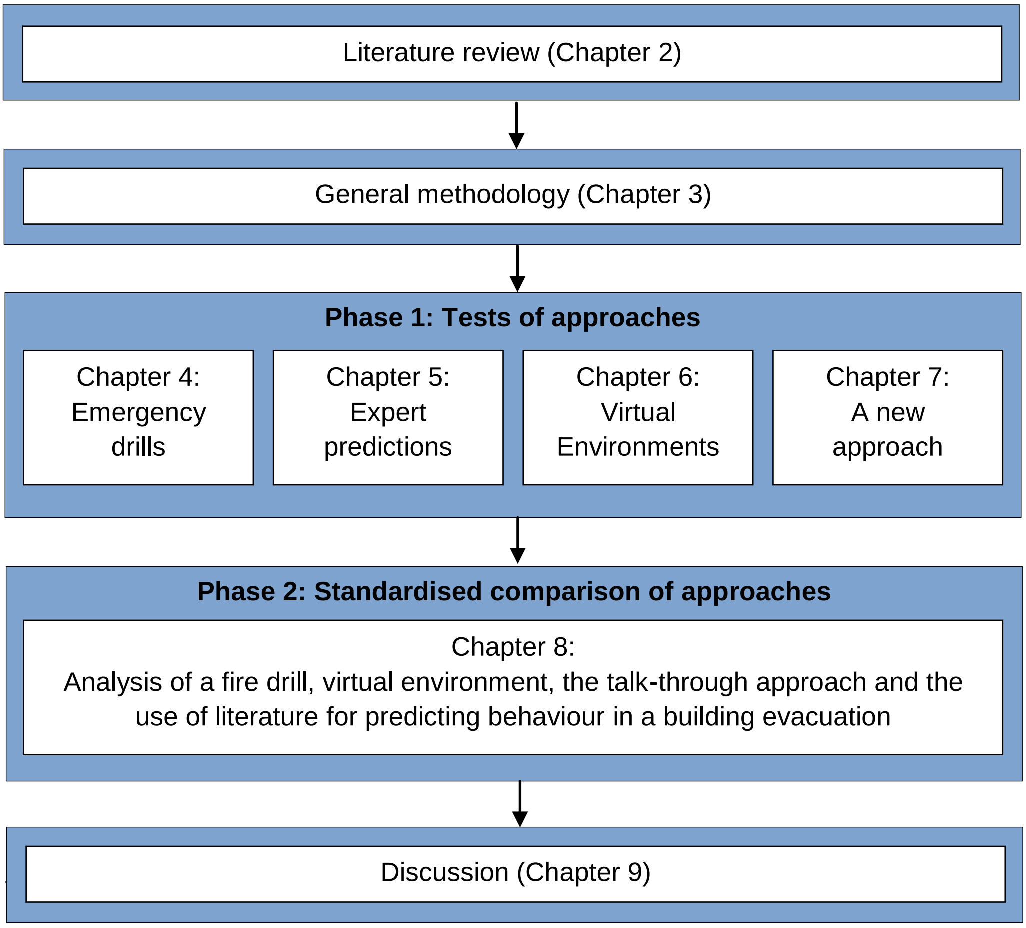

Figure 1.2. Research overview

Chapter 2. Predicting human behaviour in emergencies

Chapter 2 summarises previous literature in the field of predicting human behaviour in emergencies. It provides a review of the predictive methods and approaches used within human factors and other related disciplines. Where appropriate, the findings of the predictive investigations are reported to give an indication of the likely behavioural responses to emergency situations.

Chapter 3. General methodology

This chapter describes the approach taken to analyse the predictive approaches. In particular, it justifies the criteria selected for judging the quality of the approaches. The approaches selected for further analysis are presented, based on a review of previous research against the criteria. This chapter also provides comment on the general methodology taken for evaluating the approaches against the criteria.

Chapter 4. Emergency drills

Chapter 4 presents research into emergency response drills, and their suitability for predicting human behaviour. Several large-scale emergency response drills were observed in which emergency response teams practised their response actions. This chapter also contains Study 1, which analysed the behaviours demonstrated during a fire alarm evacuation of a hotel. Five evacuees were asked to complete a questionnaire in the days after the alarm. Their reported actions and perceptions of danger were compared to data from actual emergency situations.

Chapter 5. The use of experts for predicting human behaviour in fire

Chapter 5 presents research into the use of experts for predicting human behaviour in fire. It presents the outcome of Study 2, in which experts' predictions for the likelihood of different behaviours in a domestic fire were assessed against a reference study of human behaviour in fire (Canter et al., 1980).

Chapter 6. Virtual environment for evacuation studies

This chapter presents research into the use of virtual environments for emergency preparedness, including both training and predictive applications. It includes work on the EU-funded project DiFac (Digital Factory for Human-Oriented Production System: IST-5-035079) and a Masters student dissertation project, supervised by the author, which investigated fire drills in the Second Life VE.

Chapter 7. A new approach for predicting the human response to emergency situations

Chapter 7 describes the development of a new approach for predicting the human response to emergencies, which was based on the talk-through method (Kirwan and Ainsworth, 1992) and sequential analysis (Bakeman and Gottman, 1986). The new approach involved describing a scenario to participants, who then reported the hypothetical actions they predicted they would take. The first investigation of this approach is presented in Study 3a using a domestic fire scenario. The approach was developed and repeated in Study 3b to investigate its reliability. In Study 3c, the scenario was changed to a hotel fire to investigate the generalisability of the new approach.

Chapter 8. Standardised comparison of approaches for predicting human behaviour in emergency situations

Chapter 8 presents the research conducted in Phase 2: a standardised comparison of approaches. The approaches were selected from those which demonstrated greatest potential in Phase 1. Each approach was used to predict behaviour in the same scenario: evacuation of a building at the University of Nottingham. The output from each of the approaches was then compared against criteria for judging their quality (see Chapter 3). The specific studies were:

Study 4a: The behaviours of building occupants during an actual fire drill evacuation of the Psychology building were analysed, to investigate this as an approach for predicting emergency behaviours.

Study 4b: The use of virtual environments was investigated by analysing a virtual evacuation of the Psychology building using Second Life.

Study 4c: Use of the talk-through approach described in studies 3a-c, but applied to a hypothetical evacuation of the Psychology building.

Study 4d: As a further investigation the evacuation behaviour reported in scientific literature was investigated and compared to the outcome of the evacuation study.

Chapter 9. General discussion

The outcome of the research is discussed in Chapter 9, including discussion of the relative merits of the various approaches. The limitations of the research work are also presented.

Chapter 10. Conclusions and future work

The research is concluded and recommendations are made for further work into approaches for predicting behaviour in emergencies.

1.7 Cautionary note

Throughout the thesis, mention will be made of the limitation of the behavioural predictions, as it is important to not over-interpret the power of the analysed techniques. Prediction in this research refers to predicting the outcome of an event based on a set of conditions. Ethical considerations prevent running studies which cause danger or distress to participants (Dane, 1990), and consequently the predictions were generally evaluated against studies of previous emergencies. The predictions should only be considered in light of the conditions of each study, and the data sources against which they were compared. The method of validation and limitations of the predictions will be specified after each study, and in full in the discussion in Chapter 9.

1.8 Chapter summary

This chapter has laid the foundations for the thesis by introducing the background and origins of the research into human factors approaches for predicting human behaviour in emergencies. The aims and objectives of the research have been described. The structure of the various studies and investigations within the thesis has been presented using a research overview diagram (Figure 1.2). This chapter has prepared the reader to follow the evolution of the research as it investigated approaches for predicting human behaviour in emergencies.

2. Predicting human behaviour in emergencies

2.1 Chapter overview

This chapter presents an overview of the literature on predicting human behaviour in emergencies. The reviewed material was mainly from scientific journals and conference papers. Given the nature of the topic, reference was also made to incident reports and other documents from institutions working in emergency planning or fire safety.

The studies reported in the literature cover a variety of approaches and methods for investigating and predicting behaviour. While there were several possible categorisations and interlinking themes, the following groupings emerged and consequently formed the basis of the structure for this chapter:

- reports by survivors of emergency situations
- predictions based on scientific literature
- simulation and modelling
- predicting behaviour using Virtual Environments (VEs)
- fire drills and experimental studies
- predictions made by study participants
- expert predictions

The literature review contains an analysis of each general approach for its utility in behavioural prediction. In addition to describing the approaches which have been used to predict behaviour in emergencies, where appropriate the predicted behaviours themselves have been presented. The purpose of this was twofold: to provide the reader with an indication of the type of predictions arising from the various approaches, and also to summarise the behaviours which have been observed and predicted in emergency situations. A summary table of all the predicted behaviours is presented in Appendix I.

2.2 Reports by survivors of emergency situations

Obtaining data from survivors of emergency situations has been used as an approach to gain valuable insights into their experiences and to help understand their behaviours (Canter et al., 1980; Drury et al., 2006; Galea et al., 2007; Gershon et al., 2007; Averill et al., 2009; Galea et al., 2009; McConnell et al., 2009). Fahy and Proulx (2005) argued that this understanding is necessary to make predictions about behaviour in fire situations. It is also necessary for investigating the impact of different factors on the outcome of an evacuation (Galea et al., 2007). Furthermore, accounts by survivors can be used to obtain a description of an event, in addition to their perception and reaction to the emergency (Proulx and Reid, 2006).

Recent examples of using this approach for understanding behaviour in emergency situations can be seen following the 2001 terrorist attack on the World Trade Center (WTC) (Galea et al., 2007; Gershon et al., 2007; Averill et al., 2009; Galea et al., 2009; Gershon, 2009; McConnell et al., 2009). Data collection methods have included telephone interviews, face to face interviews, focus groups and questionnaires. Participant sample sizes varied in these studies, but generally reached high numbers, exceeding a thousand responses in some studies (Gershon et al., 2007; Averill et al., 2009). The objectives of the investigations and approaches taken have also varied between studies.

The World Trade Center Evacuation Study was conducted to investigate the factors associated with time to start evacuation, time taken to evacuate, and the risk of injury (Gershon et al., 2007; Gershon, 2009). The early phases of this study involved semi-structured interviews and focus groups with 50 volunteering survivors. The interview transcripts were analysed to identify factors which facilitated or hindered evacuation and were categorised as individual, organisational and environment (Gershon et al., 2007). These data were used to form a questionnaire, for which 1441 responses were received from evacuees of the WTC (Gershon, 2009). The research took a participatory approach, involving researchers, consultants and participants in all phases of the investigation. This was reported to have increased the value of the data by enabling survivors of the WTC disaster to help with the interpretation of the results and generation of more relevant and practical guidelines (Gershon, 2009).

Averill et al. (2009) used telephone and face to face interviews and focus groups to obtain data from over 1000 survivors of the WTC disaster. The telephone interview results were analysed to generate causal models for delays to initiate evacuation and stairwell evacuation time, each of which reportedly accounted for between 49 and 56% of the variance in total evacuation time (Averill et al., 2009).

Galea et al. (2009; *see also* Galea et al., 2007) conducted interviews with 271 evacuees, coding their responses in a database for further analysis as part of the High-rise Evacuation Evaluation Database (HEED) project. They attempted to increase the richness of the data by using unconstrained interview techniques, which started with evacuees being asked to provide free-flow narratives of their experiences. This was reported to have helped memory recollection, and captured events which may have been missed in a more structured process. The technique was followed with semi-structured interviews, which expanded on the outcome from the free-flow interviews. They also used computer generated animations to support participants' recollections of crowd densities in the stairwell, and risk perception questionnaires to understand their perceived level of risk (Galea et al., 2007, 2009). A sub-sample of the interviewees' responses was investigated using content analysis and populated the HEED database (Galea et al., 2009). McConnell et al. (2009) used the HEED data to investigate in further detail the recognition and response phases of the WTC evacuations.

The main findings from these studies are summarised below, which focus on the exhibited behaviours and actions taken by the survivors. Despite differences between the studies, there was sufficient overlap to enable a collective review of the findings.

Considering first the initial response, cues such as feeling the impact of the planes, hearing the explosion, swaying of the building and smelling burning fuel facilitated rapid evacuation (Gershon et al., 2007). Previous experience of emergencies was a factor which also contributed to a rapid response (Gershon et al., 2007). Those with a higher perceived risk responded more quickly than those with a lower perceived risk (Galea et al., 2009); those who continued working reported a significantly lower perception of risk than those who did not (McConnell et al., 2009).

Around 80% of evacuees responded within 8 minutes of the impact (Galea et al., 2009; McConnell et al., 2009). However, on average each evacuee completed four activities before evacuating (Galea et al., 2009). McConnell et al. (2009) reported the three most common initial response activities as 'seek information on the event' (24.6% of all initial response activities), 'collect belongings' (17.5%) and 'provide verbal instructions to evacuate' (15%). The latter was more common for those with a managerial role (McConnell et al., 2009). The three most common activities throughout the entire evacuation were 'collect belongings (50%), 'provide verbal instruction to evacuate' (31.7%) and 'seek information on event' (31%) (McConnell et al., 2009).

Other non-evacuation activities included: making phone calls, shutting down computers, securing items, changing footwear, and seeking permission to leave (Gershon et al., 2007; Galea et al., 2009). The number of tasks completed prior to evacuation was significantly correlated with the delay to evacuation (Galea et al., 2009). Seeking additional information was found to be one of the best predictors of evacuation initiation delay (Averill et al., 2009).

Delays were confounded by the lack of managers, which resulted in further investigation by the occupants (Gershon et al., 2007). In fact, lack of leadership was found to be one of the main factors which caused delay to evacuation (Gershon, 2009). Other delays were caused by some evacuees' concerns about their ability to walk down the stairs or unfamiliarity with the building (Gershon et al., 2007; Gershon, 2009). Difficulty locating fire exits and poor signage were primary causes of delays to initiate evacuation. Lack of previous participation in drills was reported to have caused delays to initiate evacuation, and to hinder the progression of the evacuation (Gershon, 2009).

Other factors, including emergency preparedness training and experience, increased evacuation progress; those without experience followed those with (Gershon et al., 2007). Progression was increased by other group social activities, for example praying out loud in the stairwell, and cheering after descending each flight of stairs. Clear direction from a perceived authority figure of greater seniority aided evacuation progress; direction from a more junior authority figure was acted upon less rapidly, as were ambiguous messages. Some natural leaders emerged, based in part on the clear direction they gave and on their authoritative voices (Gershon et al., 2007). Assistance from co-workers and emergency responders supported evacuation (Averill et al., 2009). The communication of emergency plans to employees prior to the event also improved evacuation (Gershon et al., 2007).

Sensory cues enhanced evacuation progress, in spite of announcements that it was safe to return to the offices (Gershon et al., 2007). These announcements were found to be a significant constraint to the evacuation of WTC 2 (Averill et al., 2009; Gershon, 2009). Considering the design of the environment, physical safety features such as lighting, handrails on the stairs, reflective tape, and floor lighting increased evacuation progress (Gershon et al., 2007; Averill et al., 2009).

Some footwear (e.g. high heels, slip-on shoes) were reported to have slowed progression, which resulted in particular problems when crossing broken glass (Gershon et al., 2007). Structural damage, congestion on the stairway, slow-moving occupants, crowdedness, fire-fighter counter flow, debris and glass in the lobby, smoke and water in the stairs, and occasionally locked doors all hindered evacuation (Gershon et al., 2007; Averill et al., 2009; Galea et al., 2009; Gershon, 2009). 85% of all participants stopped at least once; 43% of all stoppages were due to congestion and only 8% were for rest (Galea et al., 2009). Of the people who stopped for rest, 85% were female (Galea et al., 2009). Physical disabilities or impairment significantly increased the evacuation time, as did delaying activities such as stopping to make phone calls (Gershon, 2009). Survivors reported barriers to mobility from pre-existing injuries, medication/medical treatments and occasionally wheelchairs, pregnancy or older age (Averill et al., 2009). However, no correlation was found between travel speed and fitness, or fitness and number of stoppages (Galea et al., 2009). This result was explained by the crowd density causing natural breaks which inevitably enabled evacuees to rest (Galea et al., 2009).

Averill et al. (2009) used the interview data to identify travel speeds of 0.2 m/s for evacuees in the stairwells of WTC1. This figure was considered to be towards the low end of the range reported in the scientific literature, due to the crowding and obstacles encountered in the stairwells. Galea et al. (2009) report slightly higher speeds of 0.29 m/s.

Comments on the information yielded through these approaches are presented towards the end of this section, following a review of similar approaches applied to scenarios other than the WTC evacuation.

Other work which has benefitted from reports by survivors includes a questionnaire survey of the occupants of a high rise building in Chicago, 2003, in which six people died due to smoke inhalation (Proulx and Reid, 2006). The authors used a questionnaire as they recognised that it would have been “very difficult” to identify and meet all survivors in person. They received responses from 89 people who had evacuated from the building fire.

Proulx and Reid (2006) found that attempting to gain more information was one of the most common actions undertaken by occupants (39%), similar to the findings from the WTC investigations (McConnell et al., 2009). The top three motivations reported for evacuating were ‘perceiving a fire cue’ (47.4%), ‘hearing a P.A. message’ (30.8%) and ‘interaction or behaviour of other occupants’ (26.0%)(Proulx and Reid, 2006).

The mean pre-movement time from realising something unusual was happening to starting evacuation was 5 minutes (SD=4.7). However, the range was reported as 0 (immediate) to 30 minutes, with some occupants continuing to work or converse. The majority of participants gathered personal effects (71%), while fewer gathered both personal effects and job-related material (7%) or just job-related material (4%); 4% took emergency equipment and 14% took nothing. However, the security videos showed most evacuees wearing jackets (Proulx and Reid, 2006).

51% of evacuees reported using the lifts, despite emergency procedures advising against this. This was postulated by the authors to be related to time of day (around 5pm) with many occupants preparing to leave work. Lift use was associated with floor level, with more people on higher floors using them (Proulx and Reid, 2006).

Several of the occupants (17%) in the stairwell tried to re-enter the building to escape from smoke, crowding, blockages or to look for other people. However, re-entry to the building was prevented by self-locking doors and led to smoke inhalation, resulting in six fatalities and several injuries. 44% of participants moved through smoke in the stairwell, sometimes up to 27 floors of smoke (Proulx and Reid, 2006).

Interviews and questionnaires have also been used to capture and analyse the experiences of survivors of fires in low-rise buildings of a variety of types (Canter et al., 1980; Edelman et al., 1980; Wood, 1980). Canter et al. (1980) used interviews to obtain detailed and empirical data on human behaviour in fire, specifically to understand sequences of actions and behaviour during the early stages of fire development. Initially, they obtained information from the fire brigade about different types of fire incidents. They then attempted to obtain statements from witnesses involved in the fires that had occurred, supplementing this with information from press reports and police witness statements. The witnesses were asked to describe exactly what they did from the time they noticed that something abnormal was happening, until after they exited the building. These descriptions were transcribed and coded against a common taxonomy of behaviours. The researchers then performed sequence analysis to generate decomposition diagrams (Figure 2.1) demonstrating the “strength of associations” between behaviours in the sequence. These values, shown on the arrows in Figure 2.1, were described by Canter et al. (1980) as indicating the likelihood of the second act occurring, based on the occurrence of the first act. The frequencies of acts and sequences of acts (decomposition diagrams) were reported for domestic, multiple occupancy (including hotels) and hospital fires (Canter et al., 1980).

Based on the decomposition diagrams, Canter et al. (1980) attempted to extract generic models of behaviour which were displayed in all types of residency studied. They used these to create a general model of behaviour in fire situations (Canter et al., 1980). They noted the recurrence of behaviours at several different steps in the general model, emphasising the importance of understanding the sequences of behaviours rather than just their total number of occurrences. They also identified key nodes in the model, the first of which occurs after perceiving the initial cues of a fire, which can lead to either a misinterpretation sequence of events, or investigation. After seeing smoke, the evacuee then engages in one of three preparatory sequences, “instruct”, “explore” or “withdraw”. After these have been executed, the evacuee enters one of four act sequences, namely “evacuate”, “fight”, “warn” or “wait”. Canter et al. (1980) draw attention to the increase in variety of actions as the sequence of behaviour progresses. Further details of these results are presented in Chapters 4 and 7, as they were used to validate studies conducted for this thesis.

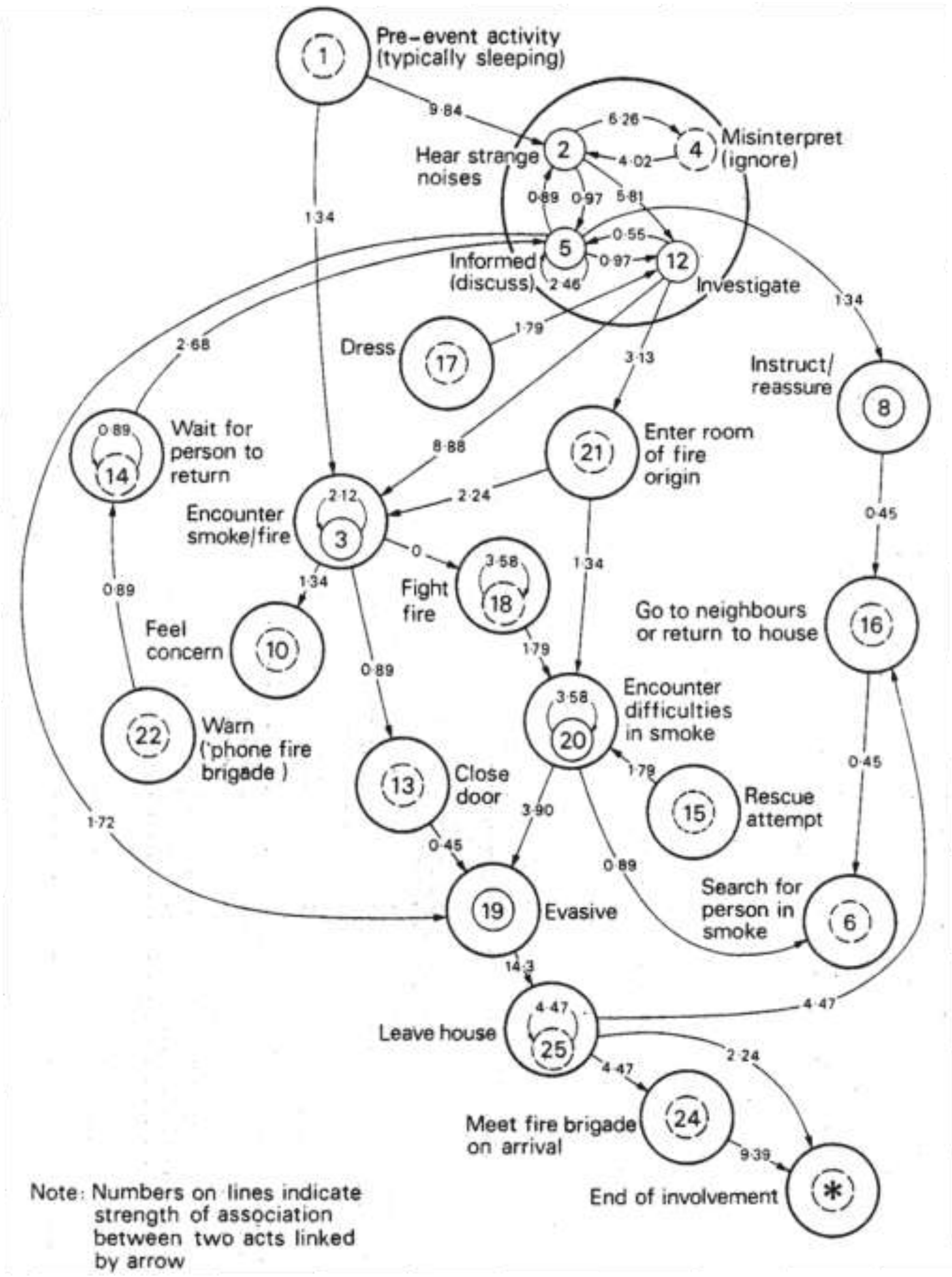

Figure 2.1. Decomposition diagram (Canter et al., 1980)

In a further study of emergency survivors, Wood (1980) obtained information on emergency behaviour using a questionnaire which was administered by fire service personnel. They collected data from 2191 participants who had been involved in 952 fire incidents in a variety of factory, residential and institutional buildings. The main findings from this study are shown in Table 2.1, which shows the behaviours investigated and the variables associated with those behaviours.

Table 2.1. Main variables associated with behaviour in fire (Wood, 1980)

Behaviour	Associated variable
First actions:	
Leave the building (rather than attempt to fight the fire)	Greater consideration that the fire is serious
Attempt to save personal effects	Less than completely familiar with building
Raising alarm or organising evacuation	More frequent training or instruction on what to do in a fire
Fight fire or minimise risk	Previous experience of a fire incident
Minimise risk	Gender: male
Warning others	Gender: female
Leave building immediately	Gender: female
Request assistance	Gender: female
Evacuate family	Gender: female
Fight fire	Gender: male Age: 10-59 years
Decreased likelihood to leave immediately	Previous experience of a fire incident
Evacuation:	
Increased evacuation	Extensive smoke spread
	Home environment (vs. work environment)
	Lack of previous involvement in a fire
	Gender: female
	Age: younger
	Untrained
	Complete familiarity with building
	Any presence of smoke
Increased re-entry to building	Gender: male
	Time of day: daytime
	Any presence of smoke
	Previous involvement in a fire
Movement through smoke	Gender: male
	More extensive smoke spread
	Home environment (rather than work)
	Time of day: daytime
	Complete familiarity with building

In another study of this type, Edelman et al. (1980) used a multi-stage approach for interviewing people who have been involved in fires. The first stage of the interview included questions regarding their location and activities prior to the fire. They then probed the cues the evacuees used to determine the existence of the fire. Thereafter, the respondents' were asked to track their movements on a plan of the building, and questions were asked about their perception of the environment for each major stage of the evacuation. The final stage of the process involved collecting background information about the respondents, and an evaluation of their physical and psychological status by the interviewer.

When applied to an evacuation of a nursing home, Edelman et al. (1980) elicited many findings relevant to human behaviour in fire. They found that the alarm often did not prompt immediate evacuation, with people more often taking no action or returning to their rooms. This was possibly due to the number of false alarms previously experienced in the nursing home. Furthermore, several respondents could not recall hearing the alarm after the event. Hearing screams of “fire” was a greater motivator to action, causing people to find further information. However, staff warnings and witnessing fire and/or smoke contributed the most towards convincing evacuees that there was a fire. They found that in general residents’ level of action (e.g. evacuate, return to room, gain information) reflected their perception of the environment (Edelman et al., 1980). Another interesting finding was that the majority of evacuees used the main stairwell, despite several possible emergency exits, many of which were closer to the residents’ rooms. This was attributed to lack of previous use of the emergency stairwells (these were alarmed and the residents were scolded if they used them), poor instruction by the staff, and the behaviour of other residents. The authors mention the benefit of flexibility in the semi-structured interview approach (Edelman et al., 1980).

Drury et al. (2006) used interviews with 21 survivors of 11 emergency events in a further study to prove that co-operation rather than panic prevails in emergency situations. They conducted qualitative analysis of the interview data to demonstrate that co-operation was more common. This was attributed to the continued influence of pre-existing social roles and everyday norms, in addition to the shared threat creating solidarity within the crowd (Drury et al., 2006).

Other work with survivors has included a study of evacuation from trains in order to understand the risks to passengers (Kecklund et al., 2009). This investigation used a questionnaire survey to identify a series of issues with train evacuations, including unclear communication of information to passengers, the time delay before the decision was taken to evacuate, and lack of training in some staff. The authors investigated only low-threat evacuations, but extrapolated recommendations to high threat situations by identifying which factors they considered to be more influential in the latter (Kecklund et al., 2009).

Jeon and Hong (2009) used questionnaires and interviews to investigate survivor behaviour following the 2003 Daegu subway fire in Korea. Some of the most pertinent findings were that several passengers (24%) acted passively after becoming aware of the fire. Impediments to evacuation included poor visibility and lack of recognition of the guide lights; only 12% of evacuees used the optimum evacuation routes (Jeon and Hong, 2009).

Despite the informative examples provided above, there are limitations to using reports by survivors of emergency situations. The resources required to conduct interviews and focus groups can be extensive; for example, in Gershon et al. (2007)

transcripts ran to 3000 pages, which were each read four times by independent researchers as part of the analysis process (Gershon et al., 2007). The HEED project (Galea et al., 2009) involved interviewing 271 evacuees, and ran to nearly 6000 pages of transcripts. Galea et al. (2007) commented upon the considerable logistics involved with the data collection of accounts from WTC survivors, including participant recruitment, arranging the interviews, and ethics approval. Timings were provided for processing and entering the interview data into the HEED database as follows: transcribing each interview (1-1.5 days); editing the transcription (1-3 days); coding the transcript (1-2.5 days); and entering the data into the HEED database (1-4 hours) (Galea et al., 2007). Wood (1980) also reported resource problems when using fire service personnel to collect data after evacuations, as this took them away from other important duties.

Ethics is also a consideration when interviewing people who may have experienced a traumatic event. This problem is not insurmountable but has required consideration and ethics approval (Gershon et al., 2007; Gershon, 2009). In the World Trade Center Evacuation Study (Gershon, 2009) participants were screened for suitability pre-test and at several time periods to determine any adverse impact of participation. Furthermore, ethics considerations resulted in a total of 10 different consent forms and disclosure statements, giving an indication of the necessary sensitivity when working with disaster survivors (Gershon, 2009).

Concerning the generalisability of this approach, the findings from a study in one scenario may be limited to the circumstances from which they were derived (Edelman et al., 1980). For example, the findings from a study into behaviour in a business-oriented high-rise study, such as those by Gershon et al. (2007), may not necessarily be relevant to residential buildings. This concern was raised by Aguirre et al. (1998) when using questionnaire responses from survivors of the 1993 WTC bombing to prove predictions drawn from Emergent Norm Theory (Section 2.3). Furthermore, Drury et al. (2006) reported unexplained variance in their data: in one emergency they found a lack of evidence of cooperation, yet were unable to explain this from the interview results. The questionnaire results concerning human behaviour in fires in the UK reported by Wood (1980) were compared to a study in the US, and several significant differences were found; however this may be attributed to different dwelling types, and therefore occupant types in the different buildings, rather than unreliability of the approach.

When using interviews and focus groups of past events, memory can influence the findings (Edelman et al., 1980; Aguirre et al., 1998; Gershon et al., 2007). Edelman et al. (1980) specifically mentioned that memory may affect results with participants inaccurately recalling their perception of an incident. Fahy and Proulx (2005) attempted to overcome these concerns by analysing first-person accounts of the WTC evacuation in the media, email exchanges and web-sites, thus enabling analysis of data which were reported much closer to the event than were possible to obtain using other research approaches (Fahy and Proulx, 2005). In recognition of concerns such as lack of control, unknown questioning or treatment of the responses by journalists, and over-representation of dramatic stories, Fahy and Proulx (2005) treated this as a study to identify themes for further research, rather than results which could be generalised to all evacuees of the WTC. However, Fahy and Proulx (2005) identified many similar findings to the interview/questionnaire studies (e.g. Gershon et al., 2007; Galea et al., 2009), for example: evacuees made phone calls before evacuating; crowds, smoke and debris were experienced in the stairwells; occupants with experience of the 1993 event more readily evacuated; and safety features aided evacuation.

Proulx and Reid (2006) reported some uncertainty in their questionnaire data of the high-rise building evacuation, particularly in self-reported quantitative data such as evacuation and pre-movement time. Similarly, interviews with survivors have also relied on estimates from the evacuees, for example asking them to estimate the time they spent waiting in line (Galea et al., 2009). Wood (1980) recognised the importance of the criticism that what people say they did is not necessarily what they actually did. To investigate this, they asked fire brigade officers to compare the evacuees' statements to their own experiences of the events; less than 1% of the responses were rejected through this validation check. Furthermore, they found high comparability between accounts from the same events (Wood, 1980). These results indicate that, in this study, evacuees' recall of events was accurate.

It is also recognised that the results from this approach are incomplete as it does not include the experiences of those who died within the emergency (Gershon et. al., 2007), or those who chose not to respond (Aguirre et al., 1998), thus opening the possibility for sample bias in the results (Wood, 1980).

The approach can give valuable and rich insights to the experiences of those involved in an emergency (Gershon et al., 2007). However, this depends on the interview technique implemented; in the Canter et al. (1980) study, the approach focussed on the performed acts and did not investigate the motivations of the building occupants. Wood (1980) also surveyed the actions taken by occupants and recognised the lack of insight this approach gave into their decision making processes.

In conclusion, reports by survivors have been used to obtain an understanding of human behaviour in fire. However, the resources required to apply the approach can be substantial. Ethics considerations exist as survivors may be asked to remember a potentially distressing event. This approach relies on survivors' memories of an incident, which may be inaccurate. It may also be difficult or impossible for the sample in any study to represent all those involved in the incident.

2.3 Predicting behaviour from scientific literature and technical reports

An alternative approach for predicting behaviour is to base predictions upon published work in journal papers, conference proceedings or incident reports. This approach has included reviewing several papers and attempting to extract behavioural rules which may be applicable in a variety of situations (Mawson, 2005; Pan et al., 2006; Proulx, 2007). It is related to the approach described in the previous section in that published reports containing data collected from survivors have sometimes been used to make behavioural predictions (e.g. Fischer, 2002). This section presents only studies which base predictions or models of behaviour based on secondary data collection and analysis. Studies based upon primary data collection methods are presented in the other sections of this chapter. Within this section, behavioural predictions are summarised where appropriate to provide an indication of the type of behavioural data obtained from scientific literature and technical reports.

It should be emphasised that comments on the use of secondary data for behavioural predictions do not apply to the anticipated purpose of this thesis. The research reported in this thesis provides an analysis of approaches for behavioural prediction, rather than a source of behavioural data for further predictions.

Several studies have used literature and reports to make predictions about human behaviour in fire (Proulx, 2001; Purser and Bensilum, 2001; Proulx et al., 2006; Proulx, 2007; Kuligowski, 2009; Tubbs and Meacham, 2009). The data sources have included technical reports (Proulx, 2007), a combination of studies and data from previous incidents (Purser and Bensilum, 2001) and scientific literature (Proulx, 2001; Proulx et al., 2006; Kuligowski, 2009; Tubbs and Meacham, 2009).

This approach has been used to investigate behaviour during the initial stages of an evacuation, such as the response to fire alarms (Proulx, 2001; Purser and Bensilum, 2001; Proulx, 2007; Kuligowski, 2009). Predictions have been made that fire alarms may not actually warn of a fire, as occupants have been reported to not draw this conclusion without additional cues (Proulx, 2007). Without further confirmation such as instruction or the sight or smell of smoke, alarms rarely trigger evacuation. Proulx (2007) used several studies and reports to provide reasons why occupants fail to respond, which include: a relatively low percentage of people recognise the signal as an alarm; occupants do not know the correct response to an alarm; a lack of confidence due to false-alarms; and inability to hear the alarm (Proulx, 2007). Proulx

(2001) cited studies which support the prediction that voice communications will improve evacuation behaviour, although pre-recorded messages were not recommended as they will not be specific enough to guide people to safety.

Purser and Bensilum (2001) also found that occupants were slow to respond to alarms in a review of incident data. They report that people in buildings have a commitment to their prior activities and need to recognise the importance of an event to stop these activities. This problem is confounded by the often ambiguous early cues of a fire. Furthermore, people misunderstand the speed with which fires can develop (Purser and Bensilum, 2001).

After hearing an alarm, occupants are likely to conduct non-evacuation activities before evacuating. These could include gathering family members, pets or valuables, warning others, fighting the fire, or completing work tasks. The activities may last several minutes (Proulx, 2001; Purser and Bensilum, 2001).

In a later review, Proulx (2007) found further evidence for these findings, and predicted that the behavioural response of an occupant to a fire alarm is dependent upon their role in the building: visitors are more likely to wait or expect instructions, whereas employees (based on their sense of responsibility) are more likely to act quickly (Proulx, 2007).

In an earlier study on a related issue, Proulx et al. (2006) used literature to develop best and worst case estimates for egress times from single family houses. Proulx et al. (2006) reviewed scientific reports relating to the factors which can influence the required safe escape time. Amongst others, these included data on the location and causes of fire and studies of the effects of age, sleep stage and drug and alcohol consumption on occupants' response times. Proulx et al. (2006) used this data to generate best and worst-case estimates for the required safe escape times from a single family house (2 minutes to 16 minutes 10 seconds).

Kuligowski (2009) used literature to investigate the factors which influence the perception of fire, the definition of the situation as fire, and definition of risk. The direction of influence is also shown i.e. whether each factor increases or decreases the likelihood of perception and definition. Kuligowski (2009) recognised the need to extend the model to include the factors influencing decisions about actions, and executing actions. The data is presented in Table 2.2.

Table 2.2. Influential factors (and their direction) on the likelihood of perception of cues and definition of the situation as fire and of the risk to self/others (adapted from Kuligowski, 2009)

Factor	Likelihood of perceiving a cue	Likelihood of defining situation as a fire	Likelihood of defining risk to self/others
Occupant-based pre-event factors			
Experience with fires	+	+	+
Knowledge of fire/training	+	+	+
Habituation with environment	-	*	
Has knowledge of routes			-
Frequent experience with false alarms		-	
Feeling of security in building		-	
Perceptual disability	-		
Older adult	-		+
Gender: female	+		+
Speaks same language as others	+		
Frequent interaction with family	+		
Occupant-based event factors			
Having a higher stress/anxiety level	-		
Perceived time pressure	-	-	+
Presence of others(especially loved ones)	-		+
Proximity to fire/visual access	+		
Sleeping	-		
High number (>1) of behavioural processes		+	
Defines situation as fire		N/A	+
Cue-based factors			
Higher number of cues	Mixed	+	+
Consistent cues		+	+
Unambiguous cues		+	
Social cues consistent with understanding of fire		+	+
Official source	+	+	
Familiar source		+	
Higher dose of toxic gases		-	
Extreme/dense cues	-		+
Visual/audible cues	+		
Risk information		+	

*Blank cells indicate no research found

Literature has also been used for quantitative analysis on pre-movement time (Purser and Bensilum, 2001). For well-managed offices, shop waiting rooms and assembly spaces, this was predicted to be most often less than 2 mins, but could extend to around 4 mins. This range was specifically derived from unannounced evacuation drills, which did not contain sleeping occupants. Furthermore, management were informed of the drills in these studies, which the authors expected to have some impact on the outcome. Purser and Bensilum (2001) concluded from reports of evacuation studies that pre-movement evacuation times take a positive skew, with some people taking much longer to evacuate. They also reported that with good fire safety management systems, pre-movement times were short and had less variation than without such systems. However they excluded multi-enclosure sleeping accommodation from this generalisation due to lack of data (Purser and Bensilum, 2001).

Predictions for behaviour in high rise buildings have also been made based on literature, for example there has been a prediction that occupants may not follow instructions telling them to stay in place, particularly if they have seen other people evacuating (Tubbs and Meacham, 2009). The authors also predicted that many occupants will have difficulty descending a large number of stairs, particularly the elderly and people with disabilities or medical conditions (Tubbs and Meacham, 2009). This latter point was in accordance with the results found from interviews with survivors of the WTC disaster (Gershon et al., 2007; Galea et al., 2009).

Literature has also been used to generate (or enhance) models and theories of behaviour, related to stress (Proulx, 1993; Ozel, 2001), rationality (Pauls and Jones 1980b; Sime, 1995; Proulx, 2001), and social behaviour in emergencies (Pauls and Jones, 1980b; Sime, 1995; Aguirre et al., 1998; Mawson, 2005; Pan et al., 2006).

Proulx (1993) applied research into decision making and information processing to generate a model of stress in a fire. The model demonstrates that information which must be processed in an emergency situation is mainly ambiguous. This contributes to stress, which causes people to experience and progress through the emotional states of: control; uncertainty; fear; worry; and confusion. The stress experienced increases through stages in the model, although it can be decreased if the person feels their decisions are leading to a problem solution. People may remain within one stage of the model for some time before moving to the next stage. Proulx (1993) argued that providing information to evacuees can reduce stress, through supporting decision-making and problem solving. The model was compared to an evacuation experiment and to the results of the King’s Cross fire of 1987 to prove its viability (Proulx, 1993).

Ozel (2001) also used theories in the literature to investigate the impact of stress on the evacuation and route selection process. Ozel (2001) argued that some stress can increase vigilance, although too much can lead to a “hyper-vigilant” state, in which people do not make use of the available information, due to rapid processing, or filtering, of information. Ozel (2001) also argued that decisions tend towards the less risky option when under time pressure, which may contribute to the common selection of familiar exit routes in an emergency. Time pressure was also predicted to result in an emphasis on negative dimensions of a choice. Stress (and the contributory time pressure) was argued to result in a narrowing of the perceptive field, which would mean that fewer cues are utilised in an emergency (Ozel, 2001). Pan et al. (2006) supported these predictions (again through the use of literature), claiming that in emergencies decision making differs due to a higher level of stress. They predicted that sub-optimal stress levels could result in a focus on immediate survival goals in some individuals, rather than altruistic behaviour (Pan et al., 2006).

Publications have been used in secondary research to prove that panic is unlikely in a fire and that most behaviour will be rational, despite the occupant being scared or nervous (Pauls and Jones, 1980b; Sime, 1995; Proulx, 2001). Movement in emergencies and crowds, even when subsequent analysis demonstrates that the behaviour did not contribute to a successful and safe evacuation, is rational from the view point of those in the situation (Pauls and Jones, 1980b; Sime, 1995).

Considering crowds and social behaviours in emergencies, scientific literature has been used to show that crowds are not homogeneous masses, and that individuals retain their rationality (Pan et al., 2006). Social structures, interactions and pre-existing social relationships below a crowd level exert a strong influence over behaviours, and altruistic and group-oriented behaviours will predominate rather than highly individualistic, selfish behaviours (Pauls and Jones, 1980b; Sime, 1995; Shaw, 2001; Drury, 2004; Pan et al., 2006). Often, these factors have manifested as clustering within the crowd, but other implications are that groups may seek to exit together. Group bonds are so strong that a separated member may re-enter a building to reform the group such that members can exit together (Pan et al., 2006).

Pan et al. (2006) made predictions based on the theory of bounded rationality, which explains that individuals are capable of making rational decisions in an emergency, albeit with limited information and cognitive capacities. The authors used this theory to explain that individuals exit through the same way they entered the building rather than evaluate all alternatives; the appearance of a problem takes longer to perceive; and immediate situations receive more focus than future scenarios. They also predicted that this may lead to apparently non-supportive crowd behaviours, for example if a queue stops moving an individual may push the person in front to resolve their immediate situation (Pan et al., 2006).

Considering interactions between individuals, Pan et al. (2006) predicted that decision making is based upon social identity, and that highly-individualistic crowd members are more likely to demonstrate behaviours which are not altruistic. The importance of personal space was recognised, and the authors reported that in a crowded situation individuals are likely to experience higher levels of stress and attempt to regain their personal space. They also recognised the concept of social proof, and predicted that individuals will look to others to guide their own actions in uncertain scenarios (Pan et al., 2006).

Pan et al. (2006) also predicted that higher crowd density, environmental constraints (such as dim lighting, too narrow exits, poor signage), and high emotional arousal can result in behaviours which do not contribute to the safety of an evacuee or other building occupants.

Aguirre et al. (1998) made predictions from the Emergent Norm Theory (ENT) which were tested using a survey of people who were in the World Trade Center at the time of the 1993 bomb. ENT explains that in unusual circumstances collective behaviours occur as people re-define their situation and interact to form a new social structure which guides their behaviour (Aguirre et al., 1998). Aguirre et al. (1998) tested several aspects of this theory against questionnaire data from 420 people who were inside the World Trade Center when a car bomb exploded in an underground parking garage in 1993. They found that occupants in small groups responded quicker if the situation was perceived as dangerous, while the opposite was found for people in large groups. Aguirre et al. (1998) explained this through ENT in that extended interaction was required to define and propose new actions in the larger groups. Moreover, pre-existing relationships were shown to delay response, which the authors argued was due to greater efforts help friends and colleagues, rather than evacuate immediately. The resources available to groups extended the time delay due to the time required to process the information. However, the opposite was found for people in groups who knew each other well; pre-existing social relationships enabled groups to utilise resources more efficiently. Aguirre et al. (1998) found that cooperativeness increased the time it took to join the evacuation, explained through extended time to search for meaning and initiate action. However, cooperativeness decreased the time to start evacuation in groups in which people knew each other well, which showed that social relationships reduce the effects of a threat. Finally, working in areas which contained groups of workers from other firms increased evacuation time, which was attributed to the interaction between groups to determine the appropriate course of action (Aguirre et al., 1998).

Mawson (2005) used literature to generate a social attachment model of human behaviour in fire, which emphasises the importance of attachment figures to human beings. The model predicts that the common response to a threat is not to flee but to move towards familiar people and places, and that the presence of such attachment figures reduces the threat of danger. A typology of the behavioural response to threat and disaster is presented, with four outcomes based on the perceived degree of danger and the presence of attachment figures. With a mild degree of perceived danger, and in the presence of attachment figures (i.e. family members, colleagues) the behavioural response is described as affiliation, which would include contacting familiar people and remaining in familiar places. With a mild degree of perceived danger, yet in the absence of attachment figures, the behavioural response would include orderly evacuation from danger and towards familiar people or places. This could include a move towards family or home for someone experiencing an emergency alone in an unfamiliar location. With a severe threat of danger, and in the presence of attachment figures, people would increase their affiliation behaviours, but would be likely to evacuate in an orderly fashion in a severe disaster. Finally, with a severe threat of danger and in the absence of attachment figures, intense flight towards attachment figures outside the unfamiliar location were predicted. Mawson (2005) cited several studies to support the model.

Other work has focussed on the anticipated response to Weapons of Mass Destruction. Fischer (2002, 2003) presented predictions, drawn from scientific literature, of the behavioural response to disaster with respect to terrorism in two separate papers. The predictions used evidence from the 2001 attack on the World Trade Center to prove the validity of the model. Further evidence was given to the prediction that panic is rare, explaining that, although frightened, occupants behaved rationally, and moved away from danger. Fischer (2002) cited video footage and conversations with survivors to support these claims. The model predicted altruism, and examples are provided of emergency response teams as well as the general public providing generous support and help to the victims (Fischer 2002, 2003).

Fischer (2003) made other interesting predictions about the behavioural response to a biological or chemical terrorist attack. Immediately after impact, several people are predicted to converge on the area, including media, relatives and others who are merely curious (Fischer, 2003). Fischer (2003) claimed that the first people to be affected by an agent will not realise it at first, with symptoms appearing several days after exposure. A similar concern was raised by Kanno et al. (2006) with regards to the behavioural response to a nuclear disaster, as information about the event will not be directly perceivable by the public. With a delay in symptoms, people may take steps to treat themselves, go to a local GP, or to the hospital (Fischer, 2003). However, if the terrorists announce that an attack has been made, hysteria was predicted to hasten overloading of the healthcare system. A large proportion of the population would hesitate to evacuate the area, waiting for all family members to gather or meet before leaving together. Others will refuse to evacuate, fearing for their property or that they will actually increase their exposure to the agent, or stay to care for critically ill family members (Fischer, 2003). Fischer (2003) actually quantified his prediction, stating that one-third to one-half of the population would evacuate. Concerning quarantines, Fischer (2003) declared that while most will cooperate with quarantine orders, many will be outside the city area before symptoms appear; others will evacuate before the announcement to quarantine is made; finally some will successfully evade the quarantine.

Other predictions have also been made after generalising research into similar events to the (rare) event in question. Glass and Schoch-Spana (2002) used literature and reports on the behavioural response to natural and technological disasters to predict the response to bioterrorism. Supporting the work by Fischer (2002, 2003), they predicted the response would predominantly include cooperation, and that panic and behaviour such as rioting and looting would be rare. Glass and Schoch-Spana (2002), in agreement with the work by Pan et al. (2006), also predicted that pre-existing social relationships will continue to exert influence. Disaster shock and psychological dependency were shown to be rare, and people are able to assess and respond to information as they obtain it, taking charge of their own particular situations, or usefully participating in the response effort (Glass and Schoch-Spana, 2002).

In generating a model of the behavioural response to a nuclear incident Kanno et al. (2006) made predictions based on a variety of reports on disasters including floods, volcanic eruptions, fires, and industrial incidents. Their predictions were categorised according to the information, recipient and situation. Their predictions are shown in Table 2.3.

Table 2.3. Behavioural response to a nuclear incident (Kanno et al., 2006)

Category	Prediction
Information	Information from private channels is more likely to be acted upon than through public channels
	Trust in information will remain low if it is received from a smaller number of media sources
	Evacuation action is likely to occur after seeing others evacuating
	Evacuation instruction will increase recognition of urgency
	Seeing and hearing ambulances or fire engines will increase recognition of an incident
Recipient	Men are less likely to pass on obtained information than women
	People associated with others who need care or help tend to evacuate earlier
	More reactive behaviour is demonstrated by elderly people
	Farmers with land or animals tend to be reluctant to evacuate
Situation	Bad weather is likely to cause people to be reluctant to evacuate
	Distance from loudspeakers and weather will affect information acquisition

It should be highlighted that several of the papers presented focus on qualitative rather than quantitative data (Pan et al., 2006; Proulx., 2007) and provide predictions which can be relatively general (Tubbs and Meacham, 2009). Kuligowski's (2009) model is limited to direction of influence (increases or decreases) of factors on perception of cues and definition of the situation as fire and of risk to self/others. Proulx's (1993) stress model is devoid of any quantitative data; in fact the criteria for progressing through the model are very general, for example "the repeated perception of ambiguous information will eventually generate a state of uncertainty which will then induce a feeling of stress" (movement from first to second loop)(Proulx, 1993).

Proulx (2001) gave quantitative data for delays in evacuation, but an absence of specifics in a table of factors having an impact on human behaviour in fire. Concerning evacuation from an area affected by a chemical or biological agent, Fischer (2003) was only able to predict to a level of one third to one half of the population. Qualitative predictions may be sufficient; it depends on the purpose and use of the behavioural predictions.

Lack of data emerges as an apparent difficulty with this approach for predicting behaviour (Pan et al., 2006; Proulx et al., 2006). While Proulx et al. (2006) were able to generate a range of best and worst case evacuation times which were claimed to be reasonable, emphasis was made on the caution which must be exerted when using the results. Proulx et al. (2006) cautioned that the range of evacuation times could be exceeded and that more research was required to have greater confidence in the results. Similarly, Pan et al. (2006) specifically mentioned the lack of published research into non-adaptive crowd behaviours. Pan et al. (2006) also mentioned that the factors considered in theories of crowd behaviour in emergencies were incomplete. In generating the typology of response to disaster and threat, Mawson (2005) recognised that more research was needed to understand whether children are more likely to demonstrate attachment behaviours in fire.

It also seems apparent that different interpretations of the data are possible. Pan et al. (2006) presented several scenarios for crowd behaviours which do not contribute to individual or group safety; this appears to contradict studies promoting a predominance of altruistic behaviours (Drury et al., 2006). Considering repeatability, Aguirre et al. (1998) made several predictions from Emergent Norm Theory (EMT) regarding evacuation behaviour; it is unclear whether the same predictions would be drawn by another researcher. This point is confounded with the observation that sometimes predictions have not been referenced, for example Proulx (2001, 2007).

Kuligowski (2009) reported difficulty in identifying the influential factors on behavioural processes, reasoned because some factors could be indirectly influential on the behavioural processes, and also due to the large number of possible outcomes of the predictive model.

Theories in the scientific literature may not be entirely accurate in all situations. Aguirre et al. (1998) found some evidence in contradiction to their predictions based on Emergent Norm Theory from the WTC evacuation in 1993, particularly that receiving information and guidance from friends, office personnel and others near the evacuees did not delay evacuation. Sime (1980), in one of the earlier reviews of the concept of panic, recognised a tendency in incident reports and scientific literature to attribute apparently non-rational behaviour as panic, simplifying the behavioural response, and without systematic analysis of the experiences of those involved in the fire. Several recent studies have proven that panic is actually rare in emergencies (Proulx, 2001; Fischer, 2002; Glass and Schoch-Spana, 2002; Fischer, 2003; Drury et al., 2006).

Validation of predictions from scientific literature may also be difficult. Despite surveying 420 people, Aguirre et al. (1998) recognised that their data related to only one study on one incident, and that greater work needed to be done to test the Emergent Norm Theory (ENT). Fischer (2002) used mainly anecdotal evidence to prove the behavioural response to disaster model.

In summary, literature and technical reports have been used to generate and support theories of behaviour in emergency situations. These have often included qualitative and general predictions. The lack of available data to support predictions is an issue, and different interpretations of the data are possible. Concerns were also identified about the extent to which predictions can be generalised to different scenarios.

2.4 Simulation models

Computer-aided simulations, which represent the geometry of a building and contain digital representations of people, can be used by engineers and architects to evaluate the suitability of a building's design for emergency evacuation (Kuligowski, 2003). Simulations have potential advantages over other approaches during design and development and for proving building safety (Gwynne et al., 1999; Laughery, 2005). For example, simulations can easily be run many times to understand the distribution of evacuation times, giving best and worst-case scenarios (Gwynne et al., 1999). This is in contrast to timing a fire drill in a real building (Section 2.6) which provides a one-off instance of an evacuation; practicality is likely to prevent running repeated drills. Using simulation enables the assessment to be conducted during the design of the building, before any construction has taken place; this means any necessary changes are likely to be less expensive (Gwynne et al., 1999; Laughery, 2005). Simulations have also been used to test scientific hypotheses or to generate emergent behaviours on which to theorise (Pan et al., 2006), or to investigate the outcome of different evacuation strategies (Hsiung et al., 2009; Kinsey et al., 2009a). The characteristics of simulations models, as defined in this thesis, are shown in Table 2.4, in comparison to virtual environments (Section 2.5). The key difference is in the control of the movement and behaviour of the avatars, which are dictated by programming algorithms in simulation models, and controlled by participants in VEs.

Table 2.4. Defining characteristics of simulation models (and virtual environments) as the terms have been used in this thesis

	Simulation models	**Virtual environments (Section 2.5)**
Representation of building	Virtual/CAD	Virtual/CAD
Representation of people	Digital avatars	Digital avatars
Movement & behaviour of avatars	Non-player characters (NPCs) dictated by simulation model: user has little input during the simulated emergency event	Avatars' movements are controlled by users through keyboard, joystick or other user-interface

Simulation is reported to have advantages over the use of building codes, a traditional approach to proving building safety which gives prescriptive details such as number of exits, width of exits, travel distance and signage (Gwynne et al., 1999; Santos and Aguirre, 2004). Building codes do not address important factors which affect the outcome of an evacuation, including environmental effects (for example heat, toxic fumes and smoke), procedural aspects (for example training, prior knowledge of the building, management of the event), and behavioural aspects (including the response to an alarm, movement speed, and social processes) (Gwynne et al., 1999; Santos and Aguirre, 2004). Simulation models enable these aspects to be investigated during the design of the building (Gwynne et al., 1999). Despite being the most difficult and complicated factor, human behaviour has been simulated in increasing levels of detail in evacuation models (Gwynne et al., 1999; Kuligowski 2003).

Review papers of simulation models have provided details of the different modelling techniques used in the tools (Gwynne et al., 1999; Kuligowski, 2003; Santos and Aguirre, 2004; Kuligowski and Peacock, 2005). The findings from these reviews are summarised below to give an understanding of the methods used in this approach to predicting human behaviour.

One important distinction between the various simulation tools is the degree to which human behaviour is modelled. This has led to categorisation of the tools as behavioural, movement and partial-behaviour models (Kuligowski, 2003). In behavioural models the digital humans are capable of decision making and/or actions in response to the conditions of evacuation, as well as movement towards an exit (Kuligowski, 2003; Kuligowski and Peacock 2005). In movement models, digital humans are moved through the building from one point to another without modelling behaviour (Kuligowski, 2003; Kuligowski and Peacock, 2005). Partial behaviour models are mainly based on movement models, but have some limited simulation of human characteristics, for example a distribution of movement time among occupants or simulation of overtaking behaviour (Kuligowski, 2003; Kuligowski and Peacock, 2005).

The method of modelling behaviour itself varies within the behavioural models (Gwynne et al., 1999; Kuligowski, 2003). Some model behaviour implicitly, which means the behaviours are implied through features such as response delays (Kuligowski, 2003). Another approach is functional analogy, in which occupants' behaviours are governed through equations applied to all avatars, which results in all individual avatars demonstrating exactly the same response to any particular event (Gwynne et al., 1999; Kuligowski, 2003). Some models have used rule-based behaviour, in which the occupants make decisions based on specified rules. An example would be "If I am in a smoke filled room, I will leave through the nearest available exit" (Gwynne et al., 1999). Models can use deterministic processes or processes incorporating random selection of rules, or a combination of these. More recent models have been developed based on artificial intelligence approaches in which individual avatars are designed with the aim of mimicking human intelligence (Gwynne et al., 1999; Kuligowski, 2003). Behaviours have also been shown to emerge as crowd phenomena such as herding, crowding, and queuing behaviour (Pan et al., 2006).

In addition to differences in the method for modelling behaviour in the models, there have been several methods for modelling occupant movement (Kuligowski, 2003). These have included:

- Movement speed based on the density of their particular location
- Speed and flow specification by the user of the model
- Speed based on distance of the avatar to other avatars or building features.
- Movement based on an electric potential map, with exits assigned a value of 0 and potential increasing across the building as distance from the exit increases. Movement is dictated by avatars trying to lower their potential at each time step.
- Movement determined by the emptiness of the avatar's surroundings
- Movement based upon the condition of the simulated fire and smoke
- Function analogy (as described above), often using rules derived from non-human applications e.g. magnetism, fluid dynamics
- Links to another model of human movement
- Using only unimpeded flow rates (Gwynne et al., 1999; Kuligowski, 2003).

Another important distinction in simulation models is the method by which the building is represented (Gwynne et al., 1999; Kuligowski, 2003). The main approaches have been reported as either coarse or fine networks (Gwynne et al., 1999; Kuligowski, 2003). In coarse network modelling, the building is created as nodes, each of which represents a building feature such as a room, corridor or stairwell (Gwynne et al., 1999). Movement of occupants is calculated from node to node, which represents movement between the associated building features (Gwynne et al., 1999). This approach is therefore limited in its ability to model events within a room, although some calculations can be made based on the area of a room, and by incorporating an adjustment for furniture (Kuligowski, 2003; Santos and Aguirre, 2004). In fine network modelling, the floor plan is overlaid with nodes or a grid network (Gwynne et al., 1999; Kuligowski, 2003). The agents move between the nodes or cells in the grids based on the movement and behavioural methods described above. The size, distribution and connectivity between the nodes/cells vary from model to model, although as a rough guide typical spacing is approximately 0.5x0.5m^2 (Gwynne et al., 1999; Kuligowski, 2003). This approach enables more accurate positioning of an agent than coarse network modelling, and events within rooms can be simulated (Gwynne et al., 1999; Kuligowski, 2003).

A further difference in the simulation models is whether they adopt a global or individual modelling approach (Gwynne et al., 1999). With the global perspective the same attributes (e.g. movement speed, behaviour) are given to all members of the population (Gwynne et al., 1999; Kuligowski, 2003). To model any reduced walking speed in an individual (e.g. for a slower moving occupant) the effect must be calculated as a reduction in the walking speed for all avatars (Gwynne et al., 1999). With a global view, specific avatars cannot be investigated during the analysis (Gwynne et al., 1999). In contrast, individual modelling approaches enable the avatars to be individually assigned personal attributes. For example, different decision making strategies and movement speeds can be randomly or directly assigned to different avatars (Gwynne et al., 1999). Another aspect of this distinction is that occupants can view the building globally or individually (Kuligowski, 2003). In individual modelling the occupants do not know all exits paths and routes in advance of evacuation. However, with global modelling they know the entire building in advance and the most efficient route with which to exit (Kuligowski, 2003).

In addition to these modelling approaches, some of the models have been reported to offer what have been described as “special features” (Kuligowski, 2003). These include the ability to model features such as: the movement of occupants against the direction of the crowd; slow moving occupants or those with disabilities; and pre-evacuation times (Kuligowski, 2003). These features are included to increase the realism of the simulations, although the data to support their use is sometimes missing (Kuligowski, 2003). Recent related work has included a framework and prototype system which displays emergent social behaviours such as queuing, herding and competitive evacuation (Pan et al., 2005).

Despite the advanced features and increased use of evacuation models in building safety, they still have several limitations (Kuligowski, 2003).

Firstly, insufficient validation of the evacuation models has been reported as one of the biggest concerns over their use (Munley et al., 1996; Gwynne et al., 1999; Shields and Proulx, 2000; Kuligowski, 2003). Gwynne et al. (1999) described the lack of a validation work as the most important issue with simulation tools following a review of 22 evacuation models. This has been partly attributed to the lack of data available for validation (Gwynne et al., 1999; Gwynne et al., 2005), and also the lack of agreement on the meaning of validation and validation protocols in the context of simulation modelling (Shields and Proulx, 2000). One reason given for the lack of validation data is that evacuation trials are most often conducted to prove the safety of a building, rather than to generate data which are suitable for validation studies (Gwynne et al., 2005). Data required for validation must address several attributes, including the type of building, the characteristics of the population, and the nature of the environment. Moreover, evacuation studies generally give no indication of the distribution or range of possible evacuation times (Gwynne et al., 2005).

There have been some attempts to validate simulation models, however. Gwynne et al. (2005) attempted to validate the buildingEXODUS simulation against two sources of evacuation data, which were chosen for investigation as they were reported to have been commonly used for validation analyses. The first study involved the evacuation of 100 police cadets from a small room through a variety of door widths. Despite being conducted specifically with the aim of obtaining evacuation movement information, Gwynne et al. (2005) found this dataset to have several key omissions relevant for a validation study, including details of the participants' gender and age, the method by which the order to evacuation was made, the presentation order of the door widths and the distribution of evacuation times for each condition. Despite these limitations, Gwynne et al. (2005) claimed high agreement of buildingEXODUS with the results from this relatively simple evacuation scenario.

Limitations were also found with the second part of the validation exercise (Gwynne et al., 2005) which used data from an evacuation drill of 381 people from a multi-storey office building. The limitations were reported as a lack of information regarding: location of some occupants at the start of the evacuation drill; age, gender and ability level of the participants; delay time to evacuate due to inadequate self-report procedures; stairwell geometry, exit dimensions and obstacles; and exit paths

(Gwynne et al., 2005). These omissions are so important for complex evacuation studies that the authors questioned any attempt to use this data source for quantitative validation, and claimed only qualitative results, for example comparing the effects of an exit route becoming blocked and queues forming at a particular exit (Gwynne et al., 2005).

The lack of data for validating evacuation modelling has been reported by other authors. Munley et al. (1996) reported a lack of quantitative data regarding the influence of various factors on evacuation, which are necessary for development and validation. After a review of evacuation modelling, Muhdi (2006) concluded that more experiments should be designed specifically for the purposes of validating evacuation models.

Meacham et al. (2004) began to address the lack of data using Monte Carlo simulations, which use repeated random sampling to address uncertainty in input variables. They first assigned distribution values for occupant characteristics from experiments reported in the scientific literature. Where data was lacking they addressed this by increasing the variation in these values, but recognised the need to investigate these aspects further. They then recreated models for two evacuations for which the data were known. Meacham et al. (2004) ran the Monte Carlo simulations, selecting random values from the distributions mentioned above to evaluate the relationship of these with overall evacuation time. They claimed that this approach can be used to evaluate uncertainty within evacuation models.

Other researchers have conducted fire drills and evacuation studies specifically to generate data for validating simulation models. Olsson and Regan (2001) conducted fire drill evacuations from university buildings to obtain data which could be compared to those produced by the commercially available Simulex simulation tool. They conclude that the travel times obtained from each approach were similar. Ko et al. (2007) recorded exit times from industrial buildings to further investigate the validity of Simulex. They also investigated the EvacuatioNZ simulation tool which was being developed at the University of Canterbury. They report that exit flow rates were quicker in Simulex than those recorded in the trial, whereas the results from EvacuatioNZ were generally more comparable (Ko et al., 2007). Purser and Boyce (2009) ran experiments on merging behaviour in stairwells, which they report validated output from the GridFlow simulation software. Xu and Song (2009) also ran an evacuation study to obtain data on evacuee movement in stairwells which was used to develop and validate a simulation model for staircase evacuation. They conclude that there was “close agreement” between the results of the two methods.

Despite these generally positive reports, Kuligowski (2003) questioned the applicability of validation studies based on evacuation drills or non-emergency movement observations to real fire scenarios. Other authors (Pauls and Jones, 1980a; Proulx, 2001) have presented different views, claiming that fire drill scenarios are similar to the situation experienced by building occupants when the fire occurs in a different part of the building. Pauls and Jones (1980a) cited specific examples in which behaviour demonstrated in real emergencies was similar to that of drills. As another example, Kinsey et al. (2009b) argued that behaviour on escalators in rush-hour conditions may reflect emergency evacuation behaviour, given a similar desire to exit as quickly as possible.

Fire drill data are less convincing in validation studies when used as input data for the simulation model. For example, Sharma et al. (2009) ran a fire drill study in the headquarters of an engineering consultancy. They used pre-movement time as one of the input parameters to their SMART Move simulation tool. They claim from the results that the tool is “reasonable to use by a fire safety practitioner” (Sharma et al., 2009). However, this method only proved the predictive ability of the tool once these parameters were known; its predictive power without them is unknown.

Regardless of the conflict of opinion about validating simulation tools with fire drill data, this method still presents problems as described by Gwynne et al. (2005) and reported above. Kuligowski (2003) presented other means of validation, such as validation against building codes, using data from previous experiments and comparability checks with other models. However, each of these has limitations. Using building codes does not incorporate environmental, procedural or behavioural aspects (Gwynne et al., 1999); using literature can be problematic (Section 2.3); and comparison with other models will only test the similarity between them, not the accuracy with which they represent a real event (Kuligowski, 2003).

Related to this last concern, differences have been found in the evacuation times predicted by different simulation models when applied to the same scenario (Kuligowski, 2003; Christoffersen and Soderlind, 2009). Kuligowski (2003) found differences in simulated times in a hotel scenario (EXIT89 was 25-40% lower for evacuation times than Simulex, depending on the origin of the simulated fire), emphasising concerns about the validity of the models. Christoffersen and Soderlind (2009) also compared the EXIT89 and Simulex evacuation models to an unannounced evacuation drill from a high-rise office building. Again EXIT89 produced shorter evacuation times (19% shorter than the drill) than Simulex (2% shorter than the drill). They concluded that the results were “reasonably close” to the drill, although they also recognised the limitations of only conducting one drill rather than several to understand the range of evacuation times (Christoffersen and Soderlind, 2009). In a review paper Kady et al. (2009) found differences in movement speeds of occupants in the simulation models, even when they were based on the same data source, due to differences in the modelling approaches taken (Kady et al., 2009).

Zhao (1999) attempted to validate a model of the occupant response to fire in a building through comparison with a combination of other methods, including hand calculation, a commercially available simulation tool, and a Monte Carlo simulation approach. Interestingly, the results from the methods used for validation differed, most notably between hand calculation and the commercially available software (Zhao, 1999).

Other limitations in the use of evacuation models include asking the user for input for which there is little data available, such as specifying values for patience or drive in the avatars during a fire scenario (Kuligowski, 2003). This problem has been attributed to the lack of data on behaviour in fire (Muhdi, 2006), and can lead to developers and users implementing invalidated values (Kuligowski, 2003). In the comparison of hotel fires mentioned earlier Kuligowski (2003) had to base pre-movement evacuation times on data from apartments as it did not exist for hotel fires. Gwynne (2009) reported an attempt to make existing data more accessible by developing a standardised repository of human egress data. This online portal was intended to store data according to standardised headings, including for example, "date of data collection"; "nature of event"; and "methods used to extract data". It was being developed in recognition that existing data had been derived from a variety of sources, often extending back several decades and were sometimes difficult to comprehend and use. The portal aimed to address these problems, and aimed to be useful for simulation and other predictive approaches (Gwynne, 2009).

Simulation tools cannot accurately model many factors associated with human response to fire situations. Some of the shortfalls have been reported as: including the perception of fire according to the sense of threat, the effects of proximity to the fire, and the presence of other occupants or the forming of groups (Kuligowski, 2003). The simulation tools are lacking in the modelling of previous experience, the occupants' familiarity with the building, their alertness or state prior to the evacuation and commitment to any previous activity (Kuligowski, 2003). Insufficient modelling of behaviour during the early stages of an emergency was also reported by Kuligowski (2003), although some subsequent work has begun to address this (Pires, 2005; Kanno et al., 2006). Pires (2005) presented an approach for modelling human cognitive behaviour in the beginning of the evacuation process, based on scientific literature and logic diagrams. However, further work was reported as necessary to populate the model with accurate data (Pires, 2005). Kanno et al. (2006) generated a simulation model of the input, situation assessment and response of residents in a nuclear disaster. The model was derived through reports from other types of disaster and investigated the effects of various factors on behaviour. Kanno et al. (2006) attempted a validation study against reports from a critical nuclear incident in Japan in 1999. The authors concluded that many aspects of the behaviours from the real incident were evident in the simulation. However, data concerning the number of occupants who actually evacuated and sheltered were used to set the parameters of the simulation model (Kanno et al., 2006), therefore there were limitations in this approach to proving the validity of the predictive power of the simulation tool.

The lack of social process in evacuation tools was further reported in Santos and Aguirre (2004), who argued that the movement of an occupant can be largely dictated by their group. The insufficiency is obvious in models which represent only movement of the agents as these inherently assume no social behaviour, but is also apparent in the models which assume homogeneity in the population, as this does not realistically represent the diversity present during group decision making in an emergency evacuation. This restricts the realism of the evacuation tools as they cannot demonstrate emergent social behaviours. Part of the problem relates to the issue reported above that current data is unavailable regarding behaviour in emergencies. Santos and Aguirre (2004) recommended that the models incorporate agents who can assess the state of other occupants and generate a collective definition of the situation as it unfolds over time. They argued that actual evacuation movement occurs in groups, and that this must also be considered. Santos and Aguirre (2004) also proposed that social science research could help increase the realism of the simulation models. Even recent developments, such as simulations demonstrating herding, queuing and competitive behaviour (Pan et al., 2005; Sharma, 2009) have revealed that more work is needed. There was no validation work reported in these works, which may be necessary given the evidence that competitive behaviour is rare (Pauls and Jones, 1980b; Sime, 1995). Pan et al. (2005) specifically recognised the need for further work on collating data on individual and social behaviours.

Incorporating research on behaviour into simulation models can also prove difficult (Silverman et al., 2001; Pan et al., 2005). An approach used by several developers of simulation models was to review published models of human behaviour, then implement these as algorithms in their simulation tools (Cornwell et al., 2002; Kanno et al., 2006; Pan et al., 2006; Silverman et al., 2006). However, this approach has been described as difficult for the following reasons: published models of behaviour can be unspecific, un-quantified or incomplete; insufficient integration exists between different areas of research; developers have insufficient knowledge and understanding of psychology and behaviour; and poor communications exist between people working in social sciences and computer programming (Silverman et al., 2001, Silverman et al., 2006). Furthermore, human behaviours are complex and difficult to code in simulation tools, which has often resulted in an over-simplification of the behaviours (Pan et al., 2006). Muhdi (2006) argued that the incorporation of realistic behaviour, travel speeds, and occupant characteristics is necessary for evacuation modelling, yet remains one of the biggest challenges for researchers. Sharma and Gifford (2005) used Radio Frequency Identification (RFID) technology for automated tracking of evacuees which is amendable to simulation tool development, although it gives no insight to the decision making process. Nilsson and Uhr (2009) proposed complex systems as an approach for modelling human behaviour in fires. This holistic approach to modelling focuses on agents, artefacts and the interactions between them, and can be used to demonstrate emergent behaviours. The emergent phenomenon, if validated with observable data from real life, can indicate probable outcomes of an emergency scenario (Nilsson and Uhr, 2009). Concerning movement, Xu and Song (2009) found that the typical cell size in simulation tools of approximately $0.5m^2$ was too large to accurately model movement on the staircases, due to higher crowd densities.

Another concern reported by Kuligowski (2003) was that not all simulation tools offer visualisation of the evacuation. Computational simulations exist which do not enable the user to view problem areas during the simulated evacuation and offer only text or numerical output. Some visual simulations offer 2D animation; more advanced tools offer 3D interactive simulation (Kuligowski, 2003; Li et al., 2004).

In conclusion, a variety of methods have been used to create simulation tools which represent human behaviour in emergencies. These tools have been used to prove the safety of buildings during emergency egress. However, concerns have been raised over the validity of the predictions. Furthermore, they do not model all facets of human behaviour in an emergency.

2.5 Virtual environments

Virtual worlds have been described as offering exciting opportunities for researchers, partly because evidence exists of similarities between behaviour in virtual and real worlds, but also because research can be conducted in situations which may be impossible or too dangerous in the real world (Mol et al., 2008; Jarrett, 2009). Virtual environments provide the opportunity for high levels of experimental control and, because of their use of computers, provide the opportunity for capturing rich data regarding the behaviour of the controlled avatar (Jarrett, 2009). Users have been reported to find them captivating and convincing, displaying responses to events which would be expected in the real world (Jarrett, 2009).

Virtual environments have been investigated for use in evacuation training and planning (Shih et al., 2000; Mantovani et al., 2001; Gamberini et al., 2003; Murakimi et al., 2005; Mol et al., 2008; Ren et al., 2008; Chittaro and Ranon, 2009; Smith and Trenholme, 2009). Some of these studies have used software development kits (SDKs) supplied by computer games manufacturers which allow the creation of new scenarios, without requiring the user to develop a new software platform (Mol et al., 2008; Smith and Trenholme, 2009). Thus, the VE developers have benefitted from pre-existing functionalities for aspects such as modelling fire, smoke, movement, gravity and allowing collaboration. Moreover, the SDKs can be accessible merely upon purchasing a game or are even free to academics (Mol et al., 2008; Smith and Trenholme, 2009).

Smith and Trenholme (2009) investigated the use of gaming engines for generating virtual environments for evacuation drills. They found that the computer game technology supported rapid development of VEs, with one developer building a virtual representation of a university building in three weeks. They tested 12 participants in the VE in three different scenarios to investigate evacuation time and behaviour through a verbal protocol approach and a post-trial questionnaire. The authors concluded that while the time to evacuate followed a similar pattern to that in real life, it was generally longer in the virtual environment. No inferential statistics were provided, however. Time to evacuate was also affected by computer gaming experience, with self-reported experts taking less time to evacuate than non-gamers. Participants gave high ratings to attention/focus when completing the evacuation tasks and also to building and task realism, although there was greater variation in responses for this last point. Ratings for navigation difficulty were varied, which may have been due to differences in gaming experience (Smith and Trenholme, 2009).

A similar approach was used by Mol et al. (2008) to investigate virtual environments for emergency planning in nuclear facilities. Mol et al. (2008) also created a virtual environment using a gaming engine which represented a real building on a nuclear plant. They included modifications such as incorporating timers for evacuation time, and increased the realism of the walking speed. They presented evacuees with an evacuation scenario in both the real and virtual environments, with the same starting point and gathering point outside the building. These were conducted in two conditions: single person; and three people at a time through a networked system in which people could see the avatars of the other participants. While no statistical analysis was conducted, Mol et al. (2008) presented the exit times, which were shown to be similar between the VE and real scenarios.

Ren et al. (2008) investigated a virtual reality system for simulating emergency evacuations in fires. This was developed due to the cost, and inherent danger of predicting behaviour using fire drills, and the difficulty of predicting correctly human behaviour in simulation tools. Their system used a head mounted-display, with the participants controlling navigation with a mouse. Participants were able to pick up fire extinguishers and use them on the fire. The system incorporated computational fluid dynamics models to accurately represent flames and smoke within the VE. While no specific experimental results were presented, Ren et al. (2008) described their system as powerful and easy to use, and an inexpensive and safe method for evaluating building designs and for training and running fire drills. They also proposed future work to allow several participants to simultaneously participate in a drill (Ren et al., 2008).

Shih et al. (2000) used virtual reality to investigate the evacuation times and routes of evacuees in comparison to those derived from traditional calculation methods. Participants were asked to evacuate from a building in four conditions: with both signage and smoke; signage only; smoke only; and no signage or smoke. They found differences in the participants' routes and evacuation times in the VE compared to those predicted from the traditional calculations. They concluded that using a 3D scenario (particularly one with smoke) is beneficial over 2D drawings for investigating emergency behaviour during building design. Meguro et al. (1998) also found benefits of virtual environments over 2D plan drawings. They used a head-mounted display of a virtual maze to train one group of participants to evacuate from a corresponding real maze. Another group was given a plan drawing to study for 30 seconds. When participants experienced the real maze the average time for those trained with the head mounted display was shorter, although no inferential statistics were presented (Meguro et al., 1998).

Virtual environments have also been used to extract rules for evacuation behaviours (Murakimi et al., 2005). This study involved recording the position and orientation of an avatar controlled by the human in an evacuation situation. Murakimi et al. (2005) described the behaviour in operational rules and logic, using input from the participants, which were used to explain their behaviour.

Other training applications include the work of Chittaro and Ranon (2009) who developed a 3D (monoscopic) serious game for training evacuation procedures. They reported the potential of VEs for increasing motivation of training, and reducing the costs associated with fire drills. The training was implemented in increasing levels of complexity, from a small fire in an office, to an emergency in a larger laboratory. The system implemented many interactive features, such as the ability to pick up objects, press alarm and lift buttons, and make telephone calls (Chittaro and Ranon, 2009).

Mantovani et al. (2001) used a head mounted display to investigate signage design in a virtual environment. They found that arrows on the floor which moved with participants resulted in faster evacuation times than traditional signage, although the differences were not significant. They reported several errors demonstrated by participants during navigation in the virtual environment, such as colliding with obstacles, and being obstructed by features in the VE (Mantovani et al., 2001). In a development study Gamberini et al. (2003) used the same system to investigate participants' responses to fires in the VE. Participants were asked to navigate to a certain area in the VE and were told to respond naturally to the event which would occur. Upon arrival, participants were exposed to one of two fire scenarios, which had either high or low intensity, signified by the density of smoke, height of flames and level of noise. Participants were noted taking time to interpret the situation and determine the course of action. Their movements were seen to be more urgent thereafter (particularly in the high intensity condition), moving rapidly towards exits. Participants combined actions (e.g. turning and forward movement) which had been conducted separately before the incident. They also collided more frequently with objects in the VE. The authors conclude that the responses by the participants were similar to those they would expect in a comparable real-life scenario. They conclude that VEs are suitable for research and training in emergency scenarios (Gamberini et al., 2003).

As shown, the resources required to build virtual environments can be low - as little as three weeks to construct a university building (Smith and Trenholme, 2009). However, the skills required to conduct research in virtual worlds may differ to those required in real worlds, for example technical knowledge will be required (Jarrett, 2009). Furthermore, concerns about the validity of this approach emerge, given the differences in evacuation time between real and virtual environments shown in Smith and Trenholme (2009). However, this study did not include any inferential statistics for this comparison, thus providing an opportunity for further work. The results by Mol et al. (2008) indicate similarity between the evacuation times in the VE and real world, but also include no statistical analysis. Gamberini et al. (2003) concluded that VEs were suitable for training and research, but provided no empirical data to prove the validity of the behaviours demonstrated by participants in their study.

Smith and Trenholme (2009) also raised concerns about some of the behaviours demonstrated by the participants, including failure to attempt to exit through windows and a general willingness to open doors with smoke coming from underneath them. The former point was attributed to participants' expectations of virtual environments, whereas the latter was hypothesised to be caused by the absence of heat, fire and noise which would be present in a real fire. Smith and Trenholme (2009) described future work to increase the realism of the virtual experience, specifically incorporating working fire extinguishers and investigating a multi-user scenario due to the anticipated influence of other participants on the evacuee. The latter point was partially investigated by Mol et al. (2008), although they also reported yet further work involving autonomous agents to investigate crowded environments. Meguro et al. (1998) also recognised the lack of sensory cues such as smell and touch in the virtual environments which are present in real environments. Trainees using Chittaro and Ranon's (2009) simulator complained about movement speeds appearing too slow, even when based on realistic data. They also called for accurate physiological models to be input to the simulation. A small number of criticisms were made about the lack of emotional involvement and stress raised through the simulator (Chittaro and Ranon, 2009).

Mol et al. (2008) mentioned the necessity for training participants to control the avatars due to the difficulty of using mouse and keyboard actions and combinations. They suggested that a joystick may be a more user-friendly interface. They also found difficulties moving through doors, either caused by participants causing blockages, or the opening doors pushing the avatar back or trapping them between the door and the wall (Mol et al., 2008).

In conclusion, VEs are a quick and effective means of predicting behaviour in virtual environments. However, the validity of the behaviours demonstrated is not fully understood for emergency situations.

2.6 Fire drills and experimental evacuation studies

Despite criticisms of the use of fire drills for representing valid human behaviour in fires (Gwynne et al., 1999; Kuligowski, 2003), they have been widely used, with some authors providing specific justification for doing so (Pauls and Jones, 1980a; Proulx, 1995; Pauls, 1999; Proulx, 2001). Proulx (2001) argued that they are representative of the situation occupants will face in a real fire if it originates in a different part of the building and their only cue is the alarm. Pauls and Jones (1980a) argued that there are sufficient similarities between drills and real emergencies to make the former useful for studying human behaviour. These include building occupants treating the threat of emergency too lightly, and communication of the threat through ambiguous cues, such as alarms, which may be interpreted as a drill without giving an indication of the seriousness of the threat. Pauls and Jones (1980a) mentioned similarities between behaviours reported in a building fire (by evacuees and fire wardens) with those demonstrated in a drill they had previously held in the building.

Proulx (1995) argued the importance of drills for understanding potential problems, educating and training occupants, and for obtaining data which can be fed into evacuation simulation models. Pauls (1999) described some of the important contributions of fire drills to the study of human behaviour in fire, including the identification of a linear relationship between egress path width and flow capacity, and that evacuation time includes non-evacuation behaviour (i.e. people do not move directly towards an exit). Perry and Lindell (2003) recognised the value of drills for identifying and resolving problems with emergency planning procedures.

An approach to studying behaviour in a fire is to video the evacuees during a drill, then to supplement this by asking them to complete a post-evacuation questionnaire (Proulx, 1995; Shields and Boyce, 2000; Gwynne et al., 2003; Xudong et al., 2009). The questionnaire can be used to elicit non-observable factors, such as perception of the alarm (Proulx, 1995). This approach was used to investigate behaviours in an announced evacuation from a retail store in China (Xudong et al., 2009). The researchers created artificial smoke which reduced visibility and a broadcast message was used to instruct customers to leave. The video analysis was used to obtain total evacuation time (490s), and the total number of customers who evacuated from each exit. The main findings from the questionnaire are summarised in the following text. Of the occupants who were shopping with an accompanier, the most common action was leave immediately (55%); the second largest specific action was to search for their accompanier and then leave together (15%). Of the shoppers without accompaniers, only 19% left immediately; 24% assisted other customers and 16% told others about the fire. Concerning pre-movement time, 36% of occupants reported starting evacuation within 60s; 42% started within 60-120s; 13% within 120-180s and 9% reported taking over 180s to start evacuation. The majority of evacuees (61%) reported their first cue of the fire as an alarm, and 23% were told by staff. 11% were uncertain of what was happening and followed other customers. 33% of evacuees chose to exit by the most familiar exit; the largest category (40%) was the exit directed by staff; 20% chose the nearest exit (Xudong et al., 2009).

Shields and Boyce (2000) also used video and questionnaires to investigate evacuation from retail stores in a study conducted in the UK, although this study was unannounced to shoppers and store staff. Some of the key findings were that the majority (57-70%) of customers had little or no commitment to the activity being undertaken at the time of the alarm, and did not complete the activity. As found by Xudong et al. (2009) the alarm and staff warnings composed the majority (76%) of participants' first cues that there was an emergency. 35.1% of the shoppers who had been separated from an accompanier searched for them before evacuating together. The most commonly cited reasons given for the exit choice were: familiarity (19.5%); proximity (50.1%); and direction by staff (25.2%). The mean pre-movement times for the four stores ranged from 25 to 37s, and the maximum pre-movement times ranged from 55 to 100s. Total evacuation times ranged from 131 to 240s. Other interesting findings were that some of the exits were not used at all, which was attributed to staff not directing people to them. Also, evacuees demonstrated a reluctance to pass disabled evacuees which caused delays in some areas (Shields and Boyce, 2000).

The video and questionnaire approach to studying fire drill behaviour was also used by Proulx (1995) in an investigation of four mixed occupancy residential buildings, in which approximately 20% of the occupants had some movement limitation. Proulx (1995) reported that the video cameras were "invaluable" for capturing occupant movement data. Some of the main findings are summarised in Table 2.5. The longer times for evacuation in buildings 2 and 3 were attributed to several of the occupants being unable to hear the alarm in their apartments. Despite this, the movement time (speed on stairs) did not vary significantly. Other interesting findings were that occupants tended to evacuate as groups, and also demonstrated a tendency to use familiar stairwells (Proulx, 1995).

Table 2.5. Data from fire drills in four mixed occupancy residential buildings (Proulx, 1995)

Building	Mean time to start evacuation	Pre-evacuation actions	Mean time to evacuate	Speed on stairs
1	2min 30	Find pet Gather valuables Get dressed	3min 05	0.52m/s
2	8min 22	Have a look in corridor Move to balcony	9min 36	0.54m/s
3	9min 42	Gather valuables Have a look in corridor	10min 57	0.62m/s
4	3min 08	Get dressed Find children	4min 38	N/A

Olsson and Regan (2001) also investigated pre-movement delays and total evacuation times, although this study was conducted in three university buildings. Their findings are presented in Table 2.6. The authors highlighted that the Law and Commerce buildings had pre-recorded evacuation messages as part of the alarm, which may have contributed to the faster times to start evacuation. They also attributed the differences in pre-movement times to the different tasks and environments of the evacuees. The results of this study were then used as part of a validation study of a simulation modelling tool (Olsson and Regan, 2001).

Table 2.6. Evacuation study in university buildings (Olsson and Regan, 2001)

Building	Area	Mean time to start evacuation	Total time to evacuate
Lecture theatre	Theatre 1	38s	90s
	Theatre 2	28s	
Law	Computer lab	20s	170s
	Library	27s	
Commerce	Computer lab	19s	220s
	Classroom	24s	

Gwynne et al. (2003) used video and questionnaires to investigate pre-evacuation times and behaviours, in a university and also in a hospital. They found that in the hospital patients did not respond to the alarm until instructed to do so by a member of staff. In the university over 50% of the students begun to evacuate after hearing the alarm. The students' pre-movement times were influenced by the number of activities conducted prior to evacuation (shutdown computer, disengage socially, collect item, investigate), which Gwynne et al. (2003) cited as an "index of engagement" with prior activities. The total evacuation took 8m 42s, or 5m 13s without two occupants who walked through the building to check for any other remaining occupants. The summary results from the university building are shown in Table 2.7; the hospital analysis was cruder and for brevity has been omitted from this summary.

Table 2.7. Pre-movement time in a university fire drill evacuation (Gwynne et al., 2003)

Category:		n	Mean (sec)	SD[a] (sec)	Range (sec)
Role	Staff [b]	17	70.8	60.0	0-246
	Students	228	73.7	37.4	8-200
No. of actions completed prior to evacuation	≤1	62	56.9	38.4	8-141
	2	121	71.0	31.6	14-167
	≥3	41	104.0	34.4	17-200
Prompting	None	119	64.8	-	10-200
	By student	22	91.6	-	38-196
	By staff	87	81.4	-	8-147

[a] frequency distributions were often found to be skewed, bimodal or demonstrate kurtosis

[b] all other results are for students only

Other studies have used video recordings of people evacuating in fire drills and unplanned evacuations to obtain data on more specific behaviours, such as merging and deference behaviour in stairwells (Boyce et al., 2009; Melly et al., 2009). These studies used discretely positioned video cameras to capture the movements of people during the evacuation drills. The videos were used to record measures such as the flow rates of people merging from the relevant floor level and from the stairs above (Boyce et al., 2009; Melly et al., 2009). Boyce et al. (2009) concluded that overall merging was approximately 50:50 from the stair and floor, although variations existed at different stages in the evacuation. They noted deference behaviour (such as allowing people with babies to pass; a male evacuee pausing to let five women pass) affecting the merging (Boyce et al., 2009). Melly et al. (2009) noticed similar deference behaviour, for example a security guard pausing to let sixteen (mostly female) evacuees pass; a group of female evacuees paused to let an entire floor evacuate. Both studies found evidence of an increase in floor flow rate when the entrance to the stairwell was located adjacent to the incoming staircase (Boyce et al., 2009; Melly et al., 2009).

Fire drills and experimental studies have been used to study emergency behaviour in transport applications. Boer (2005) investigated driver behaviour in a tunnel when a truck fire was simulated in the road, blocking their exit. They found that motorists often stayed in their cars until an announcement describing the nature of the incident was made. There was also evidence of herding behaviour with participants copying others who they saw exiting (Boer, 2005).

A series of experiments have been conducted to investigate the factors influencing aircraft evacuation behaviour (Muir, 1996; Muir et al., 1996). Muir (1996) recognised that to investigate safety features the behaviours demonstrated in trials must be realistic, although ethical and practical concerns prevent causing fear in participants. Making reference to literature (although without specific citations) Muir (1996) stated that in serious emergencies, and with limited opportunity to escape, individuals compete to survive. To re-create this competitive behaviour Muir (1996) (*see also* Muir et al., 1996) implemented a technique in which an incentive payment was made to the first 50% of participants to evacuate from a plane. This technique was first used to investigate the effects of bulkhead apertures and seating configurations on evacuation rates (Muir et al., 1996). In a second phase, the effects of smoke on evacuation rate were investigated, although the incentives were given to the first 75% of the evacuees, to increase the data available for analysis (Muir, 1996). In a third condition, the effects of assertiveness and presence of cabin staff were investigated. Muir (1996) claimed that the incentive payments resulted in a procedure suitable for generating the behavioural data necessary for analysing design features or procedures. It was claimed that the behaviours had a high degree of realism, without causing "unacceptable levels of injury". Some of the behaviours which were reported to be evident in real emergencies, included stepping on or climbing over others and climbing over the backs of the seats to evacuate (Muir et al., 1996). Another reported behaviour was that some of the occupants grouped with friends and family before evacuating (Muir et al., 1996).

In an even more extreme scenario, Brooks et al. (2001) submerged a helicopter fuselage to investigate the breath-holding ability of the occupants in comparison to the time required for evacuation. They found that occupants were unable to evacuate without using emergency breathing apparatus which had been issued to them at the start of the trial.

Other transport-related studies have included investigating egress times from rail carriages (Jong-Hoon et al., 2009; Kinsey et al., 2009b) and from a school bus (Kady and Allen, 2009). Jong-Hoon et al. (2009) used 50 students to investigate movement time between carriages, and from the carriage to the trackside. Kady and Allen (2009) presented a framework and methodology for studying evacuation from a school bus, including orienting the door at various angles to simulate an over-turned bus. Kinsey et al. (2009b) used observations of commuters leaving a train in an underground station to predict escalator behaviour. They argued that behaviours in rush-hour approximate emergency evacuation behaviour. Following analysis of 4787 commuters' movement, they concluded that in non-congested conditions, 77% of pedestrians prefer to use the escalator, regardless of whether they approached the escalator or stairwell side. In congested conditions the ability to choose was limited due to the crowd density, resulting in 35% of the commuters using the escalator. They also found that the first pedestrians arriving on the escalator prevented those behind from walking (Kinsey et al., 2009b).

Laboratory studies have been used to investigate human behaviour in fire (Muhdi et al., 2006) despite concerns about the ethics of such an approach (Edelman et al., 1980). Muhdi et al. (2006) investigated differences between maximum and normal speeds for walking and crawling to increase knowledge on occupant characteristics for use in evacuation models. 26 students were timed either walking or crawling across 100ft. They found significant differences between normal walking and each of the other types of movement (Muhdi et al., 2006).

Kobes et al. (2009) conducted an experiment to determine the effects of smoke and low-level signage on navigation behaviour. 83 people were captured on video as they evacuated from a hotel following a call from the receptionist. Occupants were studied in one of three conditions: no smoke, with smoke, and with smoke but also with exit signs at floor level. They concluded that without smoke the majority (55%) of evacuees used the main exit; with smoke the majority (64%) used the nearest exit; and with smoke and the lowered exit signs an even higher percentage (75%) used the nearest exits (Kobes et al., 2009).

Thus, while several studies have used fire drills and experimental evacuations, the approaches are not without limitation. Causing participants any distress is generally considered unethical (Muir et al., 1996), and therefore the behaviours demonstrated may differ to those in real emergencies (Gwynne et al., 1999; Boyce et al., 2009). Ethical considerations may necessitate informing building occupants prior to a drill (Proulx, 1995; Xudong et al., 2009) which would be unlikely in a real emergency. Participants are unlikely to be exposed to adverse environmental conditions which may occur in a real fire (Muhdi et al., 2006). Boer (2005) emphasised the importance of participants having the same "sensitivities, concerns, and states of mind as the target population" in behavioural tests. A distinction was made between behavioural tests and emergency worker drills. In the former, the participants are likely to retain and engage in day to day activities whereas in the latter they are focussed on evacuation activities. Moreover, in emergency worker drills, the attention is generally on the rescue workers' performance rather than the behaviour of the public (Boer, 2005). A further problem is that conducting drills several times to investigate multiple scenarios has been described as difficult (Kanno et al., 2006).

Muir (1996) claimed that using an incentive generated competitive behaviour demonstrating a high degree of realism (Muir 1996; Muir et al., 1996). However, little evidence was provided to support this claim. In Muir (1996) no evidence was provided other than unreferenced mention of competitive behaviour in emergencies in the literature. Muir et al. (1996) referenced a laboratory study from the fifties (Mintz, 1951) as evidence of competitive, counter-productive behaviour in a low-threat environment. Muir et al. (1996) also mentioned that the behaviours demonstrated represented those reported in actual emergencies, and specifically stated that this was confirmed after some of the survivors of a major accident reviewed video footage of the experiment. However, no further details of this validation exercise were provided.

Fire drills and evacuation experiments also require considerable preparation and resources. These include pre-trial medical examinations and questionnaires, and possibly having medical personnel and fire-fighters on standby, due to the potential hazards (Muir, 1996; Muir et al., 1996; Brooks et al., 2001). Even in a hotel evacuation with fake smoke, ethics approval and pre-test health questionnaire screening was required Kobes et al. (2009). Brooks et al. (2001) actually required the use of emergency breathing apparatus. Despite these precautions, participants were reported to have withdrawn from these studies due to distress (Muir et al., 1996) or to have admitted experiencing anxiety (Brooks et al., 2001). Furthermore, the actual environment of interest (or a physical replica) was required. In the studies reviewed above these included a Trident aircraft parked on an airfield (Muir, 1996; Muir et al., 1996), or a helicopter and suitable apparatus to submerge it in water (Brooks et al., 2001). Concerning building fires, Pauls and Jones (1980a) provided quantitative estimates of the working time lost due to drills. They estimated that for a building of 16 storeys, and with a reported 1526 occupants taking part in the drill, the cost was approximately 1000 person-hours. The cost was even higher if consideration is given to post-drill discussions amongst evacuees (Pauls and Jones, 1980a).

In some of the experiments only limited segments of the population were deemed eligible to participate (fit participants aged between 20 and 50: Muir et al., 1996; highly experienced instructors and Navy clearance divers: Brooks et al., 2001). Thus, the participants' skills, abilities and aptitudes did not necessarily reflect those of the target population of end users (Muir et al.,1996; Brooks et al., 2001). This problem was also experienced by Muhdi et al. (2006) who limited their study to a small range of young participants: the authors recognised the need for further investigation with a more diverse range of participants. Participants were given gloves and knee pads in the crawling conditions which may have further affected the results (Muhdi et al., 2006).

Generalisability may also be a concern for a fire drill study. Boyce et al. (2009) highlighted that the results from the merging behaviour study in stairwells may be different for different building geometries. Thus, a particular study may not be relevant under different conditions. Boyce et al. (2009) also made comment on repeatability, mentioning differences in merging patterns and ratios for the same stair design. Kinsey et al. (2009b) highlighted that their conclusions regarding escalator behaviour may be dependent upon the direction of travel, and vertical distance travelled; culture may have also influenced behaviour. Xudong et al. (2009) anticipated cultural differences in evacuation behaviour, and found that pre-movement time in a Chinese retail store was considerably longer than those found by Shields and Boyce's (2000) studies conducted in the UK. Proulx (1995) found significant differences in the time to start evacuation and total time to evacuate between four different residential buildings.

Sometimes the drills and laboratory conditions rely on observations or performance measurements, and therefore omit a deeper understanding of motivations and behaviours (Muhdi et al., 2006; Boyce et al., 2009). Muhdi et al. (2006) recognised the absence of behavioural considerations in their lab experiment, such as determining the point at which an occupant would crawl and the impact of crawling on the decision making process. Boyce et al. (2009) relied upon video analysis, and recognised the need for further work to investigate factors such as the influence of occupant characteristics and motivations on merging behaviour in stairwells. One approach to gaining an understanding of participants' thought processes is to implement post-evacuation questionnaires (Proulx, 1995; Shields and Boyce, 2000), although this is not without faults, including the recall problems discussed in Section 2.2. This would be less of a problem if the survey is conducted immediately after the event, than after a period of time. Melly et al. (2009) reportedly gave evacuees a questionnaire in their study of deference behaviour in stairwells, although they were still unable to determine whether some of the behaviours were due to gender roles or authority role. Munley et al. (1996) described the quality of data from participants' self-reports of their experiences in fire drills as "coarse". They proposed electronic tracking as a solution (Munley et al., 1996) although this would not address the concerns about obtaining a deep understanding of motivations discussed above.

In conclusion, fire drills and experimental evacuations have limitations as an approach for predicting human behaviour in fire, such as concerns about the validity of participants' behaviour and the risk of physical danger to evacuees. However, they have also been shown to provide useful data which is representative of behaviour in real emergencies.

2.7 Participant predictions

Human behaviour in fire has also been investigated using approaches which rely on predictions by experimental participants. Heyes and Spearpoint (2009) used a combination of approaches involving participant predictions to investigate the choice of lift or stair use in an evacuation. One approach involved issuing a questionnaire immediately after a fire drill, based on the assumption that the evacuees were better placed to predict their behaviour following a related event. Another approach involved presenting images related to an evacuation scenario to students in a PowerPoint presentation, and asking for their response. The final approach was to use an online survey to question attitudes towards lift and stair use. In addition to these predictive approaches, evacuees in two buildings in bomb scare evacuations were interviewed in order to identify their choice of lift or stairs, and the reasons for their choices (Heyes and Spearpoint, 2009).

Based on the online and classroom surveys the authors found that with a higher floor level, more people predicted they would choose to use the lifts. There was no significant difference between the results from these methods. However, a comparison with the published literature revealed that while this trend has been noted in other investigations, the values vary notably, which Heyes and Spearpoint (2009) attributed to differences in the contexts of the evacuation scenarios.

The fire drill, online survey and classroom survey all demonstrated a decline in the number of participants willing to use the lifts as waiting time increased. However, when plotted as the percentage of participants willing to take the lift divided by floor level against time (a useful measure for engineering calculations) high scatter was found in the results, and the authors specifically recommended caution in their use (Heyes and Spearpoint, 2009).

Finally, analysis of the online survey and the classroom survey revealed that exit choice decisions were based on the evacuees' perceptions of the quickest possible escape route. This differed from the evacuation in the bomb scare, for which speed of exit was not commonly cited as a factor in evacuation choice. Heyes and Spearpoint (2009) suggested these differences were caused by differences in the scenarios. For example, the predictive approaches did not include fire or smoke and therefore participants were unclear about what was happening; in the bomb scare the event was much less ambiguous. Moreover, in the bomb scare several evacuees mentioned the actions of others in their evacuation choice, highlighting the importance of social factors on the outcome of an evacuation (Heyes and Spearpoint, 2009).

The concern highlighted above is whether people act in accordance with their predictions. This was also raised in Boer's (2005) study of motorists' behaviour in a simulated fire, as they found that 24% evacuated via the main roadway; they contrast this to previous research in which 60% of participants who were asked stated that they would evacuate via the main roadway. .

This review has shown that few studies have analysed participant predictions of behaviour in emergency situations. Those which have been conducted indicate limitations in the accuracy of the predictions.

2.8 Expert predictions

Another approach for understanding behaviours without putting participants at risk is expert prediction. Knowledge elicitation from experts is a recognised technique within the discipline of ergonomics (Shadbolt, 2005) and has been used in many applications, for example understanding expertise in railway controllers (Farrington-Darby et al., 2006). This approach has been used to predict behaviour in emergency situations in the disaster response and fire safety literature (Dombroski et al., 2006; Zachary Au, 2009; Groner, 2009). While experts were involved in all of the other approaches presented in chapter, this section differs in that the primary information resource was the knowledge, experience and skill of the expert, rather than relying on an additional resource such as interview responses or scientific literature.

Dombroski et al. (2006) presented an iterative approach for obtaining predictions for public compliance with official orders to either evacuate or shelter following an attack from a radiological dispersion device. The experts used were academics and emergency coordinators. They first developed a risk model which was generated into a hypothetical scenario. Experts and emergency coordinators made best guess, lower bound and upper bound predictions of compliance rates with the official orders, investigating the effects of media reports (supportive or sceptical), the ability to see or hear the explosion and location (at home or at work) (Dombroski et al., 2006).

The results showed predicted compliance rates of approximately 70-80% with an order to evacuate, and 60-70% with an order to shelter in the current location. Higher compliance was expected with an order to shelter at home or evacuate from work; lower compliance was expected with an order to evacuate from home or shelter at work. A 10% reduction in compliance was predicted if the media were sceptical. The experts predicted that seeing or hearing the explosion would not have a great effect on compliance rates. They predicted that compliance could be improved by 10-15% through any of four preparatory procedures: media training programmes to reduce their scepticism; improving communication of the protective measures taken at schools to parents; workplace drills for sheltering and evacuation; and implementing first-responder risk communication programmes (Dombroski et al., 2006).

Groner (2009) presented a situation awareness requirements analysis to understand the information necessary for people in various roles to improve the safety of lift use in an emergency. The approach was based upon an abstraction hierarchy in which roles and goals were decomposed into levels at which analysis was possible, similar to the established technique of hierarchical task analysis (Kirwan and Ainsworth, 1992). The first stage was to describe a general scenario, in this instance “a fire occurs on an occupied upper floor, remote from the elevator lobby” (Groner, 2009). Thereafter, the relevant roles were described, as well as the goals and responsibilities of each role. These made up the top level of each abstraction hierarchy. At the next level, the decisions were described which each role must take to achieve their goals. For example, a building occupant will need to decide “Do I need to take any action?” The information needed to make the decision was listed. A further level included possible sources of information, which could be used to design methods for displaying and communicating information to facilitate decision making. The approach was conducted primarily by academic researchers involved in fire safety, although the model was refined by the American Society of Mechanical Engineers Task Group on Use of Elevators for Occupant Egress. No information was provided on the members of the task group (Groner, 2009).

Zachary Au (2009) presented an approach incorporating hierarchical task analysis (HTA) and HAZOP (Hazard and Operability Study) for analysing the potential human error and behavioural issues in, and their impact on, the emergency response process. The approach involves first preparing the planned response to an emergency, which could be supported through the use of HTA. Then, a panel of 3-5 task experts review each stage of the plan, considering any possible deviations, and the causes and consequences of these deviations. Thus, the analyst can identify any weaknesses in the emergency response plan. Zachary Au (2009) claimed the approach has been successfully applied in a nuclear facility, and for offshore scenarios.

Dombroski et al. (2006) provided useful guidance for further application of the use of experts for behavioural prediction, for example the necessity to provide a sufficiently detailed description to the experts such that they are able to make predictions. However, there were also limitations to this approach. In Dombroski et al. (2006) the experts’ quantitative predictions demonstrated great variation, with best guess predictions of population compliance with official orders ranging from 5 to 80% of the population in one scenario. They did compare results from the different groups of experts (academics and emergency coordinators) claiming no statistical differences, although they recognised that the small samples sizes (10 and 32 participants, respectively) would only be sensitive to large differences. They concluded that large differences in sheltering predictions between the groups were apparent (Dombroski et al., 2006).

Furthermore, validity testing was limited in the study by Dombroski et al. (2006). The authors found that the expert predictions were roughly accurate when compared to compliance rates for Hurricane Katrina, but no statistical analysis was shown, and they recognised that there were many differences between the circumstances of Katrina and the hypothetical scenario (Dombroski et al., 2006). Further concerns regarding validity may be raised with regards to the prediction that seeing or hearing the explosion would not greatly affect compliance rates. Although this scenario was different to a fire evacuation, the finding appears in contrast to the findings of Gershon et al. (2007) who found cues such as these important for initiating and progressing evacuation movement in the 2001 WTC attacks.

The use of hierarchical task analysis in Zachary Au (2009) worked well for emergency response plans, particularly in highly procedural industries such as nuclear and petro-chemical as shown. However, it may be less easy to generate for less procedural situations, as faced by the public. Moreover, while the HAZOP technique was good for identifying possible issues with response plans, it relies on the knowledge of the experts. Zachary Au (2009) argued that the experts should have knowledge of how the process is intended to work, however these people may not necessarily know how people will behave in a fire. Groner (2009) recognised the importance of input from people involved in each role, to incorporate their input to the abstraction hierarchy.

In conclusion, expert predictions have not been widely analysed for predicting behaviour in emergency situations. Of the studies that have been conducted, quantitative predictions have shown large variation in the results. Task analysis based approaches have been used in situations dominated by procedure or for specific aspects of an emergency, but their generalisability to other scenarios is unclear.

2.9 Chapter summary

A number of different approaches have been used to predict human behaviour in emergencies, including reports by survivors, use of scientific literature, simulation, virtual environments, fire drills, and expert and participant predictions. These approaches have been reviewed, and the outcomes of the studies have been presented to give the reader an understanding of the behaviours anticipated in emergency situations.

Of the reviewed literature, few studies present a comprehensive analysis of the approaches used to make the predictions. This problem is confounded by differences between the approaches used in their type, application, study design and data collected, making direct comparison between them difficult. The following chapter explains the general methodology used within the research conducted for this thesis to address these issues. It also introduces the criteria used to evaluate the quality of the predictive approaches.

3. General methodology

3.1 Chapter overview

This chapter provides a description of the general methodology used to analyse the approaches for predicting behaviour. It presents the criteria against which they were evaluated and explains how these criteria were used to judge their quality.

Justification is made for the selection of approaches which were studied in greater depth in the research work conducted for this thesis. The selection was based on analysis of the reviewed literature to identify opportunities for further research.

3.2 Criteria for analysing the approaches

In order to explain the research methodology adopted for this thesis, it is pertinent to first restate the aims, which are summarised below:

Aim 1: To analyse approaches for predicting human behaviour in emergencies against established criteria for assessing their quality.

Aim 2: To develop new approaches for predicting human behaviour in emergency situations.

Aim 3: To make a systematic comparison of approaches for predicting human behaviour in emergencies.

Aim 4: To develop recommendations and guidance for HF professionals responsible for behavioural predictions in emergency situations.

To achieve these aims, it was necessary to select criteria against which the approaches would be assessed. These criteria were explicitly mentioned within Aim 1, but were considered during all research activities.

There were several criteria and possible means for evaluating ergonomics and human factors methods which could have been used (Stanton and Annett, 2000; Stanton et al., 2004; Stanton et al., 2005). Citing Annett's (2002) criticism that ergonomists do not always pay sufficient attention to the methods they use, Wilson (2005) presented a list and description of selection criteria which can be used "to judge the adequacy and quality of a method". These were as follows:

- Validity
- Reliability
- Generalisability

- Non-reactivity
- Sensitivity
- Feasibility of use
- Acceptability and ethics
- Resources (Wilson, 2005).

Wilson (2005) hints at pragmatism in the use of these criteria, stating that meeting all the criteria will be rare. He also points out that it will rarely be necessary to do so.

The criteria which were used to evaluate the predictive approaches in this PhD were based on those presented above (Wilson, 2005). These were chosen as they were specifically presented by Wilson (2005) as criteria for evaluating methods, which corresponded closely with aim 1. They were also chosen due to their relevance to human factors applications, which again matched the focus of this research. However, the specific aspects of the criteria were reviewed with consideration given to other relevant publications (e.g. Annett, 2002; Stanton et al., 2004; Stanton et al., 2005), with consideration of the specific subject area, and in recognition of the scope of the PhD. The various facets of each criterion are discussed in the following sections, highlighting those which were used most often for the research conducted for this thesis. These criteria were often used to evaluate *approaches* in the work conducted, as defined in Section 1.5. For example, the validity of an *approach* (combination of a *setting* and a *method*) for producing a *measure* of evacuation time was assessed. However, where appropriate, the criteria were used more specifically to evaluate a *method* or *measure*.

3.2.1 Validity

Validity is essentially the extent to which a method provides the results that it is supposed to (Wilson, 2005; Howitt and Cramer, 2011). However, there are several different types of validity, with a certain amount of overlap between them, which should be assessed with consideration of the intended purpose of the test/method (Annett, 2002; Wilson, 2005, Howitt and Cramer, 2011). The main types of validity include:

- **Face validity.** This aims to determine whether a method appears valid, simply on face value. While it has been argued that this is a weak measure due to its reliance on the subjective value of the experimenter (Banyard and Grayson, 2000), it can however contribute to an overall picture of the validity of an approach. For example, it may prove a useful check of whether a test appears to measure what it is supposed to (Howitt and Cramer, 2011).

- **Content validity.** This describes the extent to which the content of a method covers the concept under investigation (Howitt and Cramer, 2011). It can be enhanced through use of experts, investigation of relevant theory, and use of the scientific literature, to ensure a sufficient number and breadth of items are included in the method (Howitt and Cramer, 2011).

- **Concurrent validity.** This type of validity determines the extent to which the results of a method correlate with those from another which is used concurrently (Howitt and Cramer, 2011). It is a form of criterion validity, as the measure under investigation is compared to criteria which are accepted as valid. Banyard and Grayson (2000) suggest that, for example, if a researcher was developing a test of extraversion, they would correlate the results against those derived concurrently from an established personality inventory (the criteria).

- **Predictive validity.** This refers to the ability of a method to predict future events (Howitt and Cramer, 2011). This is also a type of criterion validity, with the future event treated as the accepted criteria. While described as predictive, this type of validity is also used to test the ability of a method to predict a current or past event (Banyard and Grayson, 2000). This strategy was often used in this thesis to evaluate the results obtained from a predictive approach against an existing reference scenario.

- **Construct validity.** This refers to the extent to which a test measures the theoretical concept it intends to, such as a model of human performance (Banyard and Greyson, 2000; Annett, 2002). Construct validity is proven through a variety of measures (including the other types of validity listed), and is used to develop an understanding of the underlying construct of investigation (Howitt and Cramer, 2011).

- **Convergent validity.** This is the extent to which several different measures of the same concept converge. Howitt and Cramer (2011) give an example that measures of honesty should converge, regardless of how the measures are taken.

- **Ecological validity.** This is the extent to which the results of a research study relate to those which would be obtained in the real world, i.e. whether a lab-based study would represent findings from everyday life (Banyard and Grayson, 2000; Dunbar, 2005).

- **External validity.** This measures whether the results of a study can extend to other scenarios (Dunbar, 2005). It is also termed generalisability (see

Section 3.2.6), and indicates whether the same pattern of results would be obtained with different participants or in different conditions. This was one of the criteria specified by Wilson (2005) for judging the adequacy of a method.

The focus in this PhD was often on predictive validity, in which the behaviours obtained from the investigated approaches were evaluated for representation of those reported in reference studies. For example, the predicted behaviours in the research conducted for this thesis were often correlated against results from Canter et al.'s (1980) retrospective analysis of behaviour in fire. However, other aspects of the validity are also commented on where appropriate, in particular the concurrent validity of the approaches when applied to the same (standardised) scenario (Chapter 8).

The cautionary note made in Section 1.7 should be re-emphasised: validity in this thesis often refers to the extent to which an approach or method predicts data in an existing data source and under the set of conditions in which it was tested. The limitations of the approaches and methods for behavioural prediction, and means by which their validity has been assessed, must be understood before they are used for any subsequent application.

3.2.2 Reliability

This important criterion determines whether the same results are obtained upon repeated application; this minimises the possibility of drawing conclusions from results which have occurred by chance (Wilson, 2005). It may concern whether the same results are obtained by different experimenters, at different times; or by the same experimenter under different conditions (Annett, 2002).

Measures of reliability include: internal measures, inter-rater reliability, test-retest, alternate forms and the associated replicability. Internal measures of reliability are mainly applied to questionnaire design and indicate the extent to which all internal items measure the construct or concept under investigation (Howitt and Cramer, 2011). This can be tested by splitting the questions in half (either first half/second half or odd/even numbers) and correlating the two halves against each other. Alpha reliability (Cronbach's alpha) is a statistical approach in which all possible halves of scores are compared against one another (Banyard and Grayson, 2000; Howitt and Cramer, 2011). Internal reliability is an important determinant of the overall reliability of the method (Howitt and Cramer, 2011).

Inter-rater reliability is a measure of the consistency between results from different administrators of a test (Banyard and Grayson, 2000). This is particularly important when using subjective measures, such as behavioural coding schemes from video footage, to ensure reliability in the results.

Another measure of reliability is test-retest, in which an approach or method provides the same results from two applications (Banyard and Grayson, 2000; Dunbar, 2005; Wilson, 2005). The results are correlated to indicate the strength of association between the two readings, thus indicating the reliability (Banyard and Grayson, 2000). If the same participants are used, one drawback is order effects, as the participants may remember their answers or responses in the second application (Howitt and Cramer, 2011). This may be resolved by alternate-forms reliability, in which a test measuring the same construct is administered in two similar, but different forms (Dunbar, 2005; Howitt and Cramer, 2011).

Replicability is associated with reliability, and is an important notion which refers to the extent to which a study can be reproduced (Banyard and Grayson, 2000; Howitt and Cramer, 2011). Replicability is similar to test-retest reliability, although in a replicability study some elements of the method or execution are likely to differ, often changed deliberately to investigate further, or to develop some aspect of the approach. To demonstrate replicability, the same pattern of results should be achieved on each application (Banyard and Grayson, 2000; Howitt and Cramer, 2011).

The research for this thesis investigated reliability using mainly replicability approaches (Howitt and Cramer, 2011). Test-retest reliability (Wilson, 2005) was limited by access to the data sources or the resources required to conduct repeat applications. Inter-rater reliability was also investigated where subjective ratings were used to ensure the coding schemes were applied consistently.

3.2.3 Resources

The financial costs of using the approaches and methods are presented: quantifiably where possible; where this was not possible qualitative judgments were made. Considerations made when evaluating the resources included:

- Training costs or specialist knowledge which must be acquired to apply an approach or method (Stanton, 1999).
- The cost of purchasing any specialist equipment
- The people required, both researchers and participants, to implement an approach or method (Wilson, 2005)
- The time taken to execute the study (Wilson, 2005).
- The cost of any logistics required to set up an approach. Section 2.2 demonstrated that this can add significant costs to a study, as seen by the

visa requirements, hotel costs, and admin associated with contacting participants in the WTC study by Galea et al. (2007).

- Resources required for analysis. For example, Galea et al. (2007) also reported that the necessary resources to analyse interview data can be considerable.

- The costs to third parties. For example, Pauls and Jones (1980a) discuss the costs associated with a loss of working time caused by building evacuees leaving their workplaces in a fire drill.

- Any potential re-use, which may be set against the costs described above. For example, Gwynne et al. (1999) discuss the ease by which simulation tools can be re-used to provide a distribution of evacuation times; this would not be feasible with fire drills.

3.2.4 Sensitivity

Methods and approaches should have an appropriate level of sensitivity. Wilson (2005) provides an example of a wooden ruler being inappropriate to measure changes in stature caused by vibration, as it would be insensitive to this level of change. This criterion was used in this thesis to analyse what type of data were produced by each method, and at what level of detail. It was also used to investigate whether the outcomes of the particular approaches were appropriate for use by human factors professionals for predicting behaviour in emergencies. This judgement was made based on the specific requirements of human factors professionals when working in this area as described in Chapter 1 in addition to more general requirements for behavioural predictions in emergency situations as reviewed in Chapter 2. The criterion was used to consider detail such as whether the data obtained was quantitative or qualitative, but also whether the measures were useful for supporting HF applications, including design, training, and the development of emergency response procedures. These measures often included what acts people did in the emergency, time to evacuate, perception of danger and exit choice in an evacuation.

3.2.5 Ethics

As seen in the literature review (Chapter 2) ethics requires particular consideration when researching human behaviour in emergencies. In particular, the avoidance of physical or psychological distress must be ensured (Banyard and Grayson, 2000; Banyard and Flanagan, 2005; Howitt and Cramer, 2011), despite consideration of situations which pose a risk of harm to those involved. Ethics considerations included:

- **Physical harm.** As mentioned above, the methods and approaches were reviewed to determine any risk of physical harm to participants.
- **Psychological distress.** Again, the methods and approaches were reviewed to identify the risk of causing participants distress. While this may occur concurrently with a risk of physical harm, it may also occur when remembering a distressing situation (Gershon, 2009), or discussing a hypothetical event.
- **Deception.** It is sometimes important in fire drill type studies that participants have the same goals as the target population (Boer, 2005). This may result in deception, as it may be preferred not to pre-announce the fire drill (e.g. Shields and Boyce, 2000; Purser and Bensilum, 2001). The ethics of deception require careful consideration and it is not recommended without strong justification (Banyard and Flanagan, 2005).
- **Simulator sickness.** When conducting research in virtual environments, the potential exists for participants to experience any of the symptoms associated with simulator sickness (Cobb et al., 1999). This also required consideration during the research conducted for this thesis which used VEs.

Other ethics considerations were made of the right to informed consent, confidentiality of the information provided by participants, de-briefing to explain fully the background to the research, and the right to withdraw from the trial (Banyard and Grayson, 2000; Banyard and Flanagan, 2005; Howitt and Cramer, 2011).

3.2.6 Generalisability

Generalisability received particular attention as to whether the results from the methods and approaches could be applied to different emergencies or different scenarios. As mentioned above, generalisability is also described as *external validity* although given the emphasis on this criterion by Wilson (2005), and importance to predicting behaviour in a range of emergency situations, it was reported under a separate heading and in addition to the other aspects of validity. However, the research work conducted for this thesis generally focussed on building evacuation, and further work is required to fully investigate the generalisability of the methods.

3.2.7 Non-reactivity, acceptability and feasibility of use

Consideration was also given to the other criteria listed by Wilson (2005) and commented on where appropriate, although these were often analysed and reported within the headings listed above. For example, feasibility of use was often associated with resources; acceptability was often related to ethics considerations.

3.3 Analysis of approaches used in previous studies

The criteria described above were used to analyse the previous applications of the approaches presented in Chapter 2 in order to identify their strengths and weaknesses. This analysis also aimed to reveal the approaches which had not been sufficiently evaluated against the criteria in previous work, or which did not provide sufficient information for an analysis to be made. Thus, the gaps identified during the review of approaches (presented below) determined those which were selected for further analysis.

Reports by survivors of emergency situations

The use of reports by survivors as an approach for predicting behaviour in emergency situations is presented in Table 3.1. Although this approach has face validity, concerns can be seen with the predictive validity of post-event reports, due to their reliance on survivors' memories. Furthermore, the reliability and generalisability of this approach have not been investigated in depth in previous research. However, a necessary resource to implement this approach is access to survivors of a recent emergency situation. This is unlikely to be available to a human factors professional on demand. Therefore, this approach was considered infeasible, and no further studies were conducted to investigate its performance against the criteria.

Table 3.1. Review of "reports by survivors of emergency situations" against criteria for judging the quality of an approach

Criteria	Comments
Validity	Often uses post-event questionnaires and interviews: survivors' recall of past events may not be accurate (Edelman et al., 1980; Wood, 1980; Aguirre et al., 1998; Fahy and Proulx, 2005; Gershon et al., 2007).
Reliability	Little empirical data on reliability, although Drury et al. (2006) found unexplained differences in their data between events.
Sensitivity	The approach can provide rich insights into behaviour in emergencies, including actions taken, perceptions of danger, and estimates of evacuation times (Fahy and Proulx, 2005; Proulx and Reid, 2006; Gershon et al., 2007; Averill et al., 2009).
Ethics	Ethics requires consideration as participants may have experienced a traumatic event, which they are required to remember (Gershon et al., 2007; Gershon, 2009).
Resources	Considerable resources are required to conduct interviews and analyse data (Pauls and Jones, 1980b; Wood, 1980; Gershon et al., 2007; Galea et al., 2009). Access to survivors of an emergency event is a necessary requirement for this approach.
Generalisability	Predictions may be limited to the scenario from which data were derived (Edelman et al., 1980; Aguirre et al., 1998; Drury et al., 2006).

Predicting behaviour from scientific literature

An overview of the use of scientific literature as an approach for predicting behaviour is shown in Table 3.2. Concerns were raised about this approach against several of the criteria, including validity, reliability, sensitivity, and generalisability. The ethics considerations were deemed acceptable, as it requires no participants. The resources required to implement the approach for behavioural prediction in emergency situations had not been empirically reported. However, given the importance of literature in almost all scientific research, this approach was selected for further analysis.

Table 3.2. Review of "predicting behaviour from scientific literature" against criteria for judging the quality of an approach

Criteria	Comments
Validity	Lack of data is a concern (Mawson, 2005; Pan et al., 2006; Proulx et al., 2006), which reduces the content validity.
Reliability	Some concerns raised about replicability (Aguirre et al., 1998)
Sensitivity	A focus on qualitative data/general predictions (Proulx, 1993; Pan et al., 2006; Proulx, 2007; Tubbs and Meacham, 2009) or more specific aspects (Sime, 1995; Ozel, 2001; Proulx, 2001).
Ethics	Not raised as a concern – there are no participants involved.
Resources	No empirical data.
Generalisability	Predictions from literature may not be generalisable to all situations (see page 30).
	Different interpretations of data are possible (see Drury et al., 2006 vs. Pan et al., 2006; Kuligowski, 2009).

Simulation models

A review of simulation models can be seen in Table 3.3. This approach has been extensively researched and analysed, with consideration given to similar criteria to those listed (e.g. Gwynne et al., 1999; Kuligowski, 2003; Santos and Aguirre, 2004). Therefore, no further studies were conducted specifically to investigate the performance of this approach against the criteria.

Table 3.3. Review of "simulation models" against criteria for judging the quality of an approach

Criteria	Comments
Validity	The predictive validity of the tools is one of the main concerns with this approach (Gwynne et al., 1999; Shields and Proulx, 2000; Kuligowski, 2003; Gwynne et al., 2005; Muhdi, 2006).
	Regarding the concurrent validity, differences have been found between simulation models when applied to the same scenario (Kuligowski, 2003; Christoffersen and Soderlind, 2009).
Reliability	Gwynne et al. (2005) demonstrated reliability through test-retest evaluation of the buildingEXODUS simulation tool.

Criteria	Comments
Sensitivity	Can be used to investigate building designs before a building exists; has the advantage of being able to obtain a distribution of evacuation times (Gwynne et al., 1999; Kuligowski, 2003). Most models offer visual simulations to enable the user to identify bottlenecks or other problems in the evacuation. Computational simulations exist which provide no visualisation, and are therefore limited to numerical/textual descriptions of the simulated evacuation (Kuligowski, 2003; Kuligowski and Peacock, 2005).
Ethics	Advantageous over other approaches in that the effects of smoke, fire and toxic gases can be investigated on the outcome of an evacuation without concern for any participants (Gwynne et al., 1999; Santos and Aguirre, 2004)
Resources	Re-use is high - evacuation scenarios can be run many times to understand the distribution of times (Gwynne et al., 1999).

Virtual environments

A summary review of the use of virtual environments for predicting behaviour in emergencies against the criteria for judging the quality of an approach is shown in Table 3.4. It can be seen that concerns were raised about the predictive validity of evacuation times obtained through this approach, and of the face validity of some of the behaviour demonstrated in VEs. No empirical data has been identified on the reliability of the approach. No notable ethics considerations were presented, and the resources required to implement this approach were low.

Recent studies (Mol et al., 2008; Smith and Trenholme, 2009) have commented positively on the use of VEs for studying behaviour in emergencies. It was therefore decided to investigate this approach further. In particular, research was required to understand in greater detail the validity of the approach. Further work was also required to review VEs against all the criteria to provide further guidance on their use for predicting behaviour.

Table 3.4. Review of “virtual environments” against criteria for judging the quality of an approach

Criteria	Comments
Validity	Some concerns about differences in evacuation times between real and virtual worlds; some unrealistic behaviour demonstrated in VEs (Meguro et al., 1998; Mantovani et al., 2001; Gamberini et al., 2003; Smith and Trenholme, 2009).
Reliability	No empirical data.
Sensitivity	Have been used to investigate evacuation times and behaviour in the VE (Mol et al., 2008; Smith and Trenholme, 2009)
Ethics	Can incorporate fire and smoke, which is too dangerous to study in the real world (Mol et al., 2008; Jarrett, 2009; Smith and Trenholme, 2009).
Resources	Low – virtual environments can be built quickly (Smith and Trenholme, 2009).

Fire drills and experimental evacuations

Table 3.5 summarises a review of fire drills/experimental evacuations as an approach for predicting and analysing behaviour. This demonstrates conflicting views on the predictive validity of the behaviours derived from this approach. The reliability and generalisability of the behaviours to other scenarios of the approach also required further investigation. This approach was therefore selected for further analysis.

Table 3.5. Review of "fire drills and experimental evacuations" against criteria for judging the quality of an approach

Criteria	Comments
Validity	Concerns over the predictive validity of behaviours (Gwynne et al., 1999; Boer, 2005; Boyce et al., 2009); other authors have supported their use (Proulx, 1995; Pauls, 1999; Proulx, 2001).
Reliability	No empirical data regarding reliability.
Sensitivity	Approach has been used to study human behaviour in fire, including egress times (Proulx, 1995; Shields and Boyce, 2000; Gwynne et al., 2003; Xudong et al., 2009).
	Observations may need supplementing with questionnaires to fully understand behaviours (Proulx, 1995; Shields and Boyce, 2000).
Ethics	Concerns about participant well-being (Muir, 1996; Muir et al., 1996; Brooks et al., 2001; Muhdi et al., 2006; Kobes et al., 2009).
Resources	Conducting several times is difficult; considerable preparation and resources are required (Pauls and Jones, 1980a; Muir, 1996; Muir et al., 1996; Brooks et al., 2001; Kanno et al., 2006; Kobes et al., 2009).
	A physical representation of the environment of interest is required (Gwynne et al., 2003; Muhdi et al., 2006; Boyce et al., 2009; Kobes et al., 2009; Melly et al., 2009).
Generalisability	Results may be specific to the scenario from which they were derived (Boyce et al., 2009; Kinsey et al., 2009b; Xudong et al., 2009).

Participant predictions

Insufficient data existed in previous research to comprehensively evaluate participant predictions as an approach for predicting human behaviour in emergency situations (Table 3.6). However, this approach could potentially avoid some of the ethics considerations of, for example, running an evacuation experiment as participants are not exposed to any physical danger. For these reasons, the approach was selected for further investigation in the research conducted for this thesis.

Table 3.6. Review of "participant predictions" against criteria for judging the quality of an approach

Criteria	Comments
Validity	Some concerns about the predictive validity (Boer, 2005; Heyes and Spearpoint, 2009).
Reliability	No empirical data
Sensitivity	May provide insights into behaviour in fire, such as choice of lift or stair use in an evacuation (Heyes and Spearpoint, 2009).
Ethics	Avoids the risk of physical injury which is present in approaches that attempt to physically re-create aspects of the emergency, as seen in Heyes and Spearpoint (2009).
Resources	No empirical data

Expert predictions

Table 3.7 shows an analysis of expert predictions against the criteria for judging the quality of an approach. As for participant predictions, it can be seen that analysis of the criteria was incomplete. Therefore, expert predictions were also identified as an approach worth investigating in greater detail.

Table 3.7. Review of "expert predictions" against criteria for judging the quality of an approach

Criteria	Comments
Validity	Some concerns about predictive validity (Dombroski et al., 2006).
Reliability	No empirical data
Sensitivity	Task analysis based approaches useful in situations dominated by procedure (Zachary Au, 2009); usefulness in other situations less clear
Ethics	No empirical data regarding ethics, but no study participants are involved.
Resources	Requires access to experts (Dombroski et al., 2006; Groner, 2009; Zachary Au, 2009).

Summary

In summary, the approaches which required further analysis against the criteria for judging their quality included the use of literature, virtual environments, fire drills, participant predictions and expert predictions. This led initially to studies to analyse these approaches in greater detail against the criteria, reported in Phase I. This is shown graphically in Figure 3.1.

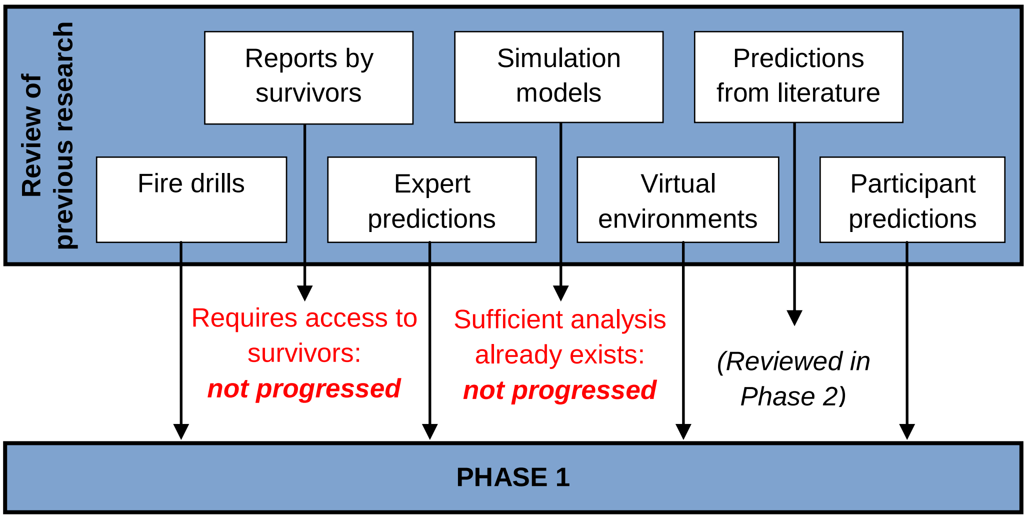

Figure 3.1. Approaches selected for further analysis

3.4 Research methodology

Section 3.3 justified the selection of approaches which were investigated in further detail. These either demonstrated potential for predicting behaviour, or had not been fully analysed against the criteria for judging the quality of the approach. The general research methodology is described below; the specific methodologies will be presented in the relevant chapters.

Phase 1 included analysis of virtual environments, fire drills, participant predictions and expert predictions. The studies within Phase 1 were conducted specifically for this PhD, and therefore differed from previous research as insight could be gained into the performance of the approaches against the criteria for judging their quality. For example, a more detailed analysis of the resources required to implement each approach was possible, which was not often reported in previous research. Phase 1 also included research work which was either not conducted specifically for this thesis, or with notable contribution from others. However, the author had a sufficient level of involvement with the use of the predictive approaches to allow for analysis against the criteria. Within Phase 1 new approaches were also developed, with the aim of improving their performance against these criteria.

The approaches which showed the greatest potential success for predicting human behaviour in emergencies in Phase 1 were continued to Phase 2. Here, a standardised comparison was conducted which allowed for a more controlled analysis of the approaches, again with consideration of the criteria for judging their quality. For example, it was possible to have an even more detailed analysis of the resources required to run a fire drill, or use of virtual environments for predicting emergency behaviours. This was conducted on the basis that few previous comparative analyses had been conducted of approaches for prediction. Phase 2 also included analysis of literature as a predictive approach, as a benchmark against which the results of the other approaches were compared.

Thus, the methodology for evaluating the approaches was to use them to make a behavioural prediction or study of behaviour, then analyse the outcome of that prediction or study with reference to the criteria for judging the quality of an approach.

3.5 Chapter summary

This chapter has introduced and described the criteria which were used to judge the quality of the predictive approaches. Justification has also been made of the selection of approaches for further analysis. This chapter explained the general approach taken for analysing the predictive approaches against the criteria; the specific analyses will be described in the subsequent chapters.

Phase 1: Tests of approaches

Phase 1 includes tests of the approaches which demonstrated potential for predicting human behaviour in emergencies following the review of previous research in Chapter 3. It includes the analysis and development of approaches which had not been fully evaluated against the criteria for judging their quality.

The selection of approaches included in Phase 1, based on the review of previous research, is shown in Figure I below. This also indicates the chapter structure within Phase 1.

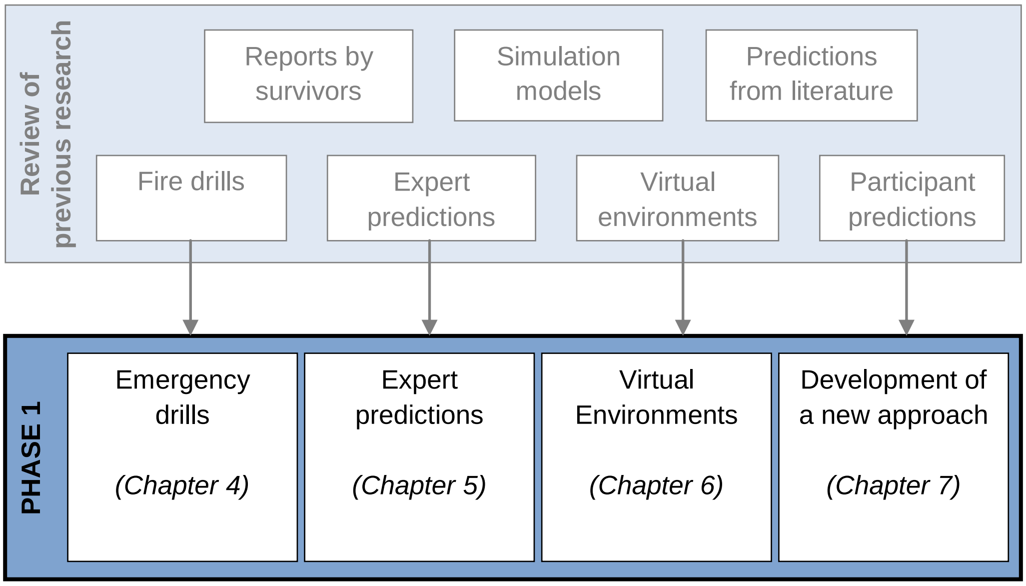

Figure I. Selection of approaches from previous research which were investigated further in Phase 1.

4. Emergency drills

4.1 Chapter overview

This chapter presents research into the use of emergency drills for predicting human behaviour in fire. This includes investigations of large-scale emergency response drills set up to train emergency services to respond to a Chemical, Biological, Radiological or Nuclear (CBRN) terrorist attack. It also presents Study 1: an analysis of behaviour demonstrated during a fire evacuation of a hotel.

4.2 Introduction

The literature revealed that drills and evacuation experiments have provided useful data about human behaviour in emergencies (Section 2.6). The main limitations of the approach were concerns about the validity of the behaviours demonstrated, ethics considerations, and the resources required to conduct this type of investigation. This section describes research into emergency response drills as further investigation of this approach for predicting human behaviour in emergencies. The research work aimed to investigate some of the criticism of this approach, such as the inability to obtain valid human behaviour from a drill (Gwynne et al., 1999; Kuligowski, 2003). It also aimed to identify the type of behavioural predictions and data which can be obtained from the use of drills.

The work included observation of large-scale emergency response exercises, and Study 1: an investigation of a hotel fire evacuation in response to an unannounced (false) fire alarm. The selection of researched exercises was determined by their availability to the author. For the emergency response exercises, access was arranged through contacts the University of Nottingham had with a local emergency response team, whereas Study 1 was opportunistic research following the author's chance involvement with a fire alarm evacuation. The findings from the investigations were compared against the criteria for judging the quality of a human factors approach presented in Section 3.2, taken from Wilson (2005).

4.3 Large-scale emergency response drills

4.3.1 Introduction

Large-scale emergency response drills were investigated for predicting behaviour in emergencies. The drills were conducted by Fire Resilience teams across the UK as a requirement of the New Dimension programme. This £330m government-funded initiative aimed to improve the Fire and Rescue Services' ability to respond to a threat such as a Chemical, Biological, Radiological or Nuclear (CBRN) incident. It was initiated after the 11th September attacks on the World Trade Center with the intention of improving training, equipment and response procedures in the UK Fire and Rescue Services (Communities and Local Government, 2009a).

As part of the New Dimension programme, several mass-decontamination exercises were organised to train emergency responders and test equipment and procedures in the emergency services. These procedures may be required following a CBRN incident, as it may be necessary to decontaminate members of the public if they have been exposed to harmful agents (Cabinet Office Civil Contingencies Secretariat, 2005; Home Office, 2009; PSCA International, 2009). During the exercises volunteers underwent decontamination activities, including getting undressed then showering in specially constructed tents, before re-dressing in disposable outfits (South Yorkshire Fire and Rescue, 2007; Lancaster and Morecambe College, 2008; Ambulance HART, 2009; Communities and Local Government, 2009b).

The aim of this research was to investigate the drills as an approach for predicting human behaviour in emergencies. While some authors have argued the benefit of drills for understanding human behaviour in fire (Proulx, 1995; Pauls, 1999; Proulx, 2001) these studies were investigations to provide more information on the approach.

4.3.2 Method

Three mass-decontamination exercises were observed in England between September 2007 and September 2008. These were mainly conducted to test the mass-decontamination equipment and procedures in the emergency response teams. For example, the fire service aimed to confirm the time taken to erect the tents, and the time taken for participants to pass through them. The participants included emergency service personnel and civilians.

The research involved observation of the exercises from a sufficient distance to avoid causing any interference, in accordance with human factors best practice (Kirwan and Ainsworth, 1992). Video footage and photographs were taken for subsequent reference. The behavioural analysis was based on approximately 50 participants involved in the decontamination exercise. The videos from the exercises were reviewed, and participants' gross-level behaviours were qualitatively analysed against the coding scheme shown in Appendix I and descriptions of human behaviour in emergencies from Chapter 2.

The investigation aimed to provide an initial understanding of the value of dry-run exercises as an approach for predicting human behaviour in emergencies. Thus, during the data collection consideration was given to the criteria for judging the quality of a human factors approach (Wilson, 2005).

4.3.3 Results

Some of the behaviours demonstrated by the participants in the mass-decontamination events are listed below. These were notable behaviours which contributed an understanding of the validity of emergency drills for predicting human behaviour. Details of the procedures have been kept to a minimum due to the sensitive nature of this type of drill.

- Participants often exhibited herding behaviour; they generally moved in groups or in pairs from the waiting area to the decontamination tents.
- Grouping was also seen while waiting, although individuals were seen leaving the group to attend to their belongings, or to look around.
- Participants spent a significant portion of the waiting time talking with each other. Their discussions appeared relaxed and casual.
- Participants were generally observed to listen to and follow the instructions given by authority figures.
- Participants spent some time chatting casually with the Police guards while they waited to be decontaminated. However, the guards held sufficient authority to keep the participants in the waiting area until the correct time to be decontaminated.
- A pair of participants investigated the police guard’s backpack at one point in the drill. This involved touching the backpack, apparently without permission or invitation from the guard, who did not seem concerned.
- Participants often removed aspects of their personal protective equipment, such as their facemasks or the hoods of their decontamination robes.
- While waiting to go through the decontamination tents, participants generally appeared relaxed, talking and showing few signs of anxiousness or distress. However, they often looked towards the decontamination tents, presumably in anticipation of the next stage of the process.
- At one point a participant was seen raising his hands and cheering in response to some verbal information from the guard.
- One participant pretended to lift the front of his robe while facing another participant. He also pinched the elbow of a fellow participant.
- Another participant was seen talking on a mobile phone while waiting to go through the decontamination tents.

- Participants were observed stamping their feet, apparently to keep warm.
- At one point a participant lifted the cordon tape on top of his colleague's head.
- Participants were seen toy fighting with their hands inside their robes.
- Moving towards the tents, one group of participants clapped and pointed at another participant who was walking in the opposite direction.
- One participant paused while walking towards the tent to talk to an observer.
- Participants entered the decontamination tents in an orderly fashion; there was no urgency or pushing. One participant was seen making an obvious gesture to allow another participant through before him.

4.3.4 Discussion

The behaviours described above, and other pertinent findings from the investigations of large-scale dry-run exercises are discussed below, categorised according to the criteria for judging the quality of an approach (Section 3.2, based on Wilson, 2005). The findings were mainly indications worth further investigation, rather than conclusive results about this approach for predicting behaviour.

Validity

The behaviour demonstrated by the participants who were decontaminated in the exercises varied in the extent to which they would represent behaviour in a real emergency. While it was difficult to accurately report the predictive validity of the behaviours without evidence from an actual decontamination event, attempts have been made below based on the literature surrounding other emergencies.

The behaviours demonstrated by the participants which would be expected in an actual decontamination exercise included:

- Grouping behaviour, and clustering during movement (Pan et al., 2006)
- Following instructions issued by authority figures (Fischer, 2003)
- Making telephone calls (Fischer, 2003), although the reason for the call was likely to be different to that in an actual decontamination.
- Orderly behaviour (Mawson, 2005; Drury et al., 2006)
- Social behaviours (Pan et al., 2006)

The observed behaviours which may not be expected in actual decontamination included:

- The range of joviality behaviours (e.g. dancing, pretending to lift the front of the disrobe cape, putting a cordon tape on another participant's head and toy fighting). While evidence exists for altruistic behaviours aimed to raise spirits during emergencies (BBC News 2001; Gershon et al., 2007), no evidence exists for such extreme joviality as seen during the drills.
- An absence of fear or anxiety (Drury et al., 2006), as evidenced by relaxed conversation between participants and removal of personal protective equipment.
- Investigating through touch the Policeman's respiratory equipment. While there is little evidence in the literature to dispute this behaviour, it is anticipated that it would not be tolerated in an actual emergency situation.

The impact of the unexpected behaviours on the outcome of the event was unclear. Further investigation was required to determine the validity of the behaviours demonstrated in dry-run exercises.

Reliability

Similar behaviours were demonstrated by the participants in all exercises with no extreme variation, although they were not compared in detail.

Resources

While it was not possible to put a figure on the resources required for the exercises, they were clearly considerable. The exercises involved several hundred emergency response personnel, participants, expensive equipment and large areas to conduct the exercises. They also took several hours to erect the tents, decontaminate the participants then collapse all the equipment. Two of the drills were all day events; the others took approximately four hours. Costs were expected to have run into tens if not hundreds of thousands of pounds. This was in agreement with the high resources reported in previous studies using fire drills and evacuation experiments (Pauls and Jones, 1980a; Muir, 1996; Muir et al., 1996; Brooks et al., 2001; Kanno et al., 2006; Kobes et al., 2009). However, if proved to be valid and generalisable, the results of the exercise would have high re-use, as the predictions could be applied to other behavioural investigations.

For the data analysis, the large-scale drill was based on a qualitative review of the video footage against published behavioural phenomenon. While the literature review was conducted over several months, the analysis itself took less than a week.

Sensitivity

The sensitivity of this approach for predicting human behaviour depends on the purpose of the investigation. If it is to obtain an indication of some of the behaviours expected during an emergency scenario then it may be appropriate, if proven to reveal valid behaviours; if measuring time for a procedure it can only be one sample from a range of possible times (Gwynne et al., 1999).

The method used (video observation) only provides a description of the behaviours; the method would need to be supplemented with interviews and questionnaires to provide an understanding of the motivations and acts of the participants.

Ethics

There was no notable evidence of distress or physical harm caused to any of the participants.

4.3.5 Conclusions

These investigations provided an insight into the use of dry-run exercises for predicting human behaviour in emergency situations. Some behaviours were demonstrated which may be anticipated in an actual mass-decontamination. However, behaviours were also demonstrated which may not have been expected – further research is required to understand the impact of these. The resources required to implement the approach in these scenarios were also found to be high due to the large number of participants and equipment involved. These could be reduced in smaller-scale fire drill evacuations, as demonstrated in a study of a hotel fire evacuation in Section 4.4 below.

4.4 Investigation of the behaviour demonstrated during a hotel fire evacuation (Study 1)

4.4.1 Introduction

The behaviour demonstrated during fire evacuations was further investigated when the author was involved in an unannounced evacuation of a hotel. The evacuation occurred in February 2010 from a hotel in Saariselkä, in the north of Finland. The author, and five colleagues, was staying in the hotel as part of a project meeting on the ManuVAR (CP-IP 211548) EU-funded research project. The fire alarm sounded early in the morning at around 6:30am. Only a partial evacuation of the hotel took place, with many guests remaining in their rooms. However, all of the ManuVAR participants evacuated. Approximately 15 minutes after the fire alarm had begun a Fire Officer entered the building, without any protective equipment. Soon after, the fire alarm stopped and the officer returned to the fire gathering point announcing that it was safe to re-enter the building. It was unclear what had caused the alarm, although it was rumoured to have been triggered by inappropriate use of one of the saunas in the hotel. It should be noted that outside temperatures were in the range of -9 to -24 degrees Celsius, necessitating the use of heavy winter clothing when outside. It was also dark at the fire gathering point.

The evacuation was actually a false alarm rather than a fire drill, although it can be used to provide insight into the behavioural analysis which is possible from retrospective analysis of a drill. It also provided information relevant to the analysis of reports by survivors of emergency situations. However, a key difference is that in the hotel fire in Saariselkä there was never any real danger to the participants, which is more comparable to a drill than an actual event. Thus, in this instance analysis of the approach is appropriate to the use of fire drills for predicting human behaviour in an emergency.

The method chosen for use in this study was selected to yield data comparable to the study of behaviour in real hotel fires by Canter et al. (1980), introduced in Section 2.2. This was selected as a reference study as it was the most detailed and relevant available data source regarding human behaviour in real fires. In the study, Canter et al. (1980) transcribed and coded interview data from survivors of real fires using a common taxonomy. They then analysed the sequences of behaviour using sequential analysis (Bakeman and Gottman, 1986), an established method for studying behaviour which provides an understanding of events as they unfold over time. It is therefore particularly suited to studying dynamic aspects of behaviour. Canter et al. (1980) used sequential analysis to generate decomposition diagrams which indicate the transitional relationships between the acts (e.g. Figure 2.1). These diagrams were shown for multiple-occupancy fires (which included hotel fires) and were used to validate the behaviours obtained from this study.

4.4.2 Method

All five of the author’s colleagues who had taken part in the hotel evacuation were contacted and asked if they were willing to be interviewed regarding their experiences. They all agreed, although three asked to complete an electronic questionnaire rather than be interviewed due to constraints on their time. All data were collected within three weeks of the event, at the earliest convenience of the participants.

As was the case in the reference study, participants were asked to describe, in order, the actions they took after realising something unusual was happening. They were told to be very detailed in their responses. Participants were also asked to state where they performed the actions (i.e. hotel room, corridor, or foyer). The information was captured and entered into a spreadsheet.

Participants were also asked the following questions:

1. Please rate your perception of danger when you first heard the alarm:

No danger				**Great danger**
1	2	3	4	5

2. Please rate your perception of danger when you decided to leave your room:

No danger				**Great danger**
1	2	3	4	5

3. Please rate your perception of danger as you exited the hotel:

No danger				**Great danger**
1	2	3	4	5

4. How long do you think it took you to leave your room?
5. How long do you think it took to evacuate the hotel?
6. Which of your colleagues were already outside when you left the hotel?
7. What influenced your choice of evacuation route?

Figure 4.1. Questions used in the study of evacuation behaviour, Saariselkä hotel evacuation, February 2010

4.4.3 Results

Frequency of acts

The frequencies of the acts reported in this study are shown in Table 4.1 within the taxonomy of acts reported in the reference study (Canter et al., 1980). The frequencies are shown as a percentage of the total number of comparable acts, for both this study (N=96) and the reference study (N=1703). For clarity, acts less than 1% for both studies are not shown in the table.

Note that act 27a “End of involvement” was included to support analysis of the sequence data, but was not included in the taxonomy of acts in the reference study. Therefore, all frequency data have been omitted from this act to facilitate the comparison. Acts 12c and 28a were reported in the Saariselkä hotel evacuation, but could not be mapped to any act in the reference study and therefore new action categories were created. Act 1a was included by default, as all participants were engaged in some form of pre-event action.

Table 4.1. Frequency of acts, as a percentage of the total number of acts per study. Acts are shown in descending order of frequency for Study 1.

		Frequency (%)	
Code	**Action Category**	**Study 1**	**Canter et al.**
6a	Seek information and investigate	17	10
18b	Note behaviour of others (unambiguous)	15	3
9a	Dress/gather valuables	10	5
4b	Arrive at conclusion	8	1
21a	Experience uncertainty	8	1
23a	Enter area of minimum risk	7	5
24a	Leave immediate area	6	4
1a	Pre-event actions	5	6
10a	Evasive	5	5
12c	Return to room	5	0
5a	Incorrect interpretation	3	3
2a	Perception of stimulus (ambiguous)	2	7
28a	Return to hotel	2	0
3a	Perception of stimulus (unambiguous)	1	5
4a	Correct interpretation	1	1
16c	Experience negative feelings	1	0
17a	Note persistence of stimulus	1	1
18a	Receive information (verbal)	1	3
2b	Note behaviour of others (ambiguous)	0	1
3b	Note worsening of immediate situation	0	3
3c	Note fire development	0	3
7a	Disseminate warnings/information	0	3
7c	Raise the alarm	0	1
11a	Coping (self-related)	0	4
12a	Securing environment	0	3
14a	Give instructions	0	1
14b	Receive instructions	0	2
15a	Give assistance	0	1
15b	Receive assistance	0	3
15c	Note arrival of assistance	0	2
15d	Seek assistance	0	2
16a	Experience movement/breathing difficulties	0	3
18c	Note people who need to be rescued	0	1
20a	Duty related	0	2
27a	End of involvement	*	

* This act was reported 5 times in this study, but no value was provided in Canter et al. (1980). Therefore, it was removed to improve the accuracy of the comparison.

Both sets of data demonstrated non-normality in the Shapiro-Wilk test (this study: W=0.582, df=52, p<0.001; Canter et al.: W=0.838, df=52, p<0.001). The frequencies of the acts demonstrated in the Saariselkä fire drill correlated significantly with the frequencies of acts obtained from the reference study of behaviour demonstrated in real hotel fires (r_s=0.414, N=52, p<0.01) and showed a medium effect size based on Cohen's (1988) standard effect sizes. A scatter plot of the results, again shown as a percentage of the total number of acts, is shown in Figure 4.2.

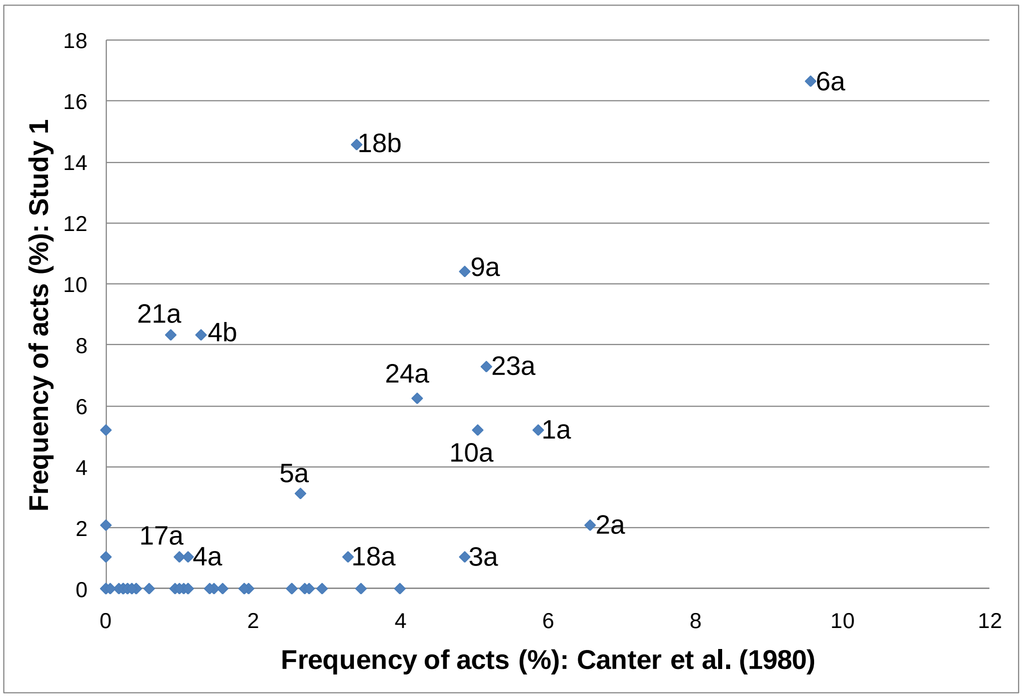

Figure 4.2. Frequency of acts for Study 1 and the reference study, shown as a percentage of the total number of acts per study. Labels (shown for all acts for which both values>0) refer to the codes in Table 4.1.

Sequence of acts

Standardised residuals (Bakeman and Gottman, 1986) were calculated for the transitions between each group of acts, to match the calculation which Canter et al. (1980) appear to have used for their "strength of association values". Canter et al. (1980) explain that these values indicate the extent to which occurrence of an act increases the likelihood of the following act occurring at this point in the sequence. Standardised residuals place an emphasis on the influences of cause and effect by indicating the extent to which a transition deviates from that expected by chance alone, rather than focussing on the transitional probability of moving from one act to the next. The latter measure may simply reflect differences in the probabilities of the acts occurring, whereas standardised residuals take into account the base rate for acts through calculation of the expected frequencies (Bakeman and Gottman, 1986). Furthermore, standardised residuals indicate how far each observed transition is above or below the expected frequency. Thus, large standardised residuals show the transitions with greater deviance from the expected result, which can be identified for further analysis (Colgan and Smith, 1978). The standardised residuals for the Saariselkä fire drill were calculated using the following formula:

$$\text{Standardised residual} = \frac{\text{observed frequency - expected frequency}}{\sqrt{\text{expected frequency}}}$$

$$\text{where expected frequency} = \frac{\text{row total *column total}}{\text{grand total}}$$

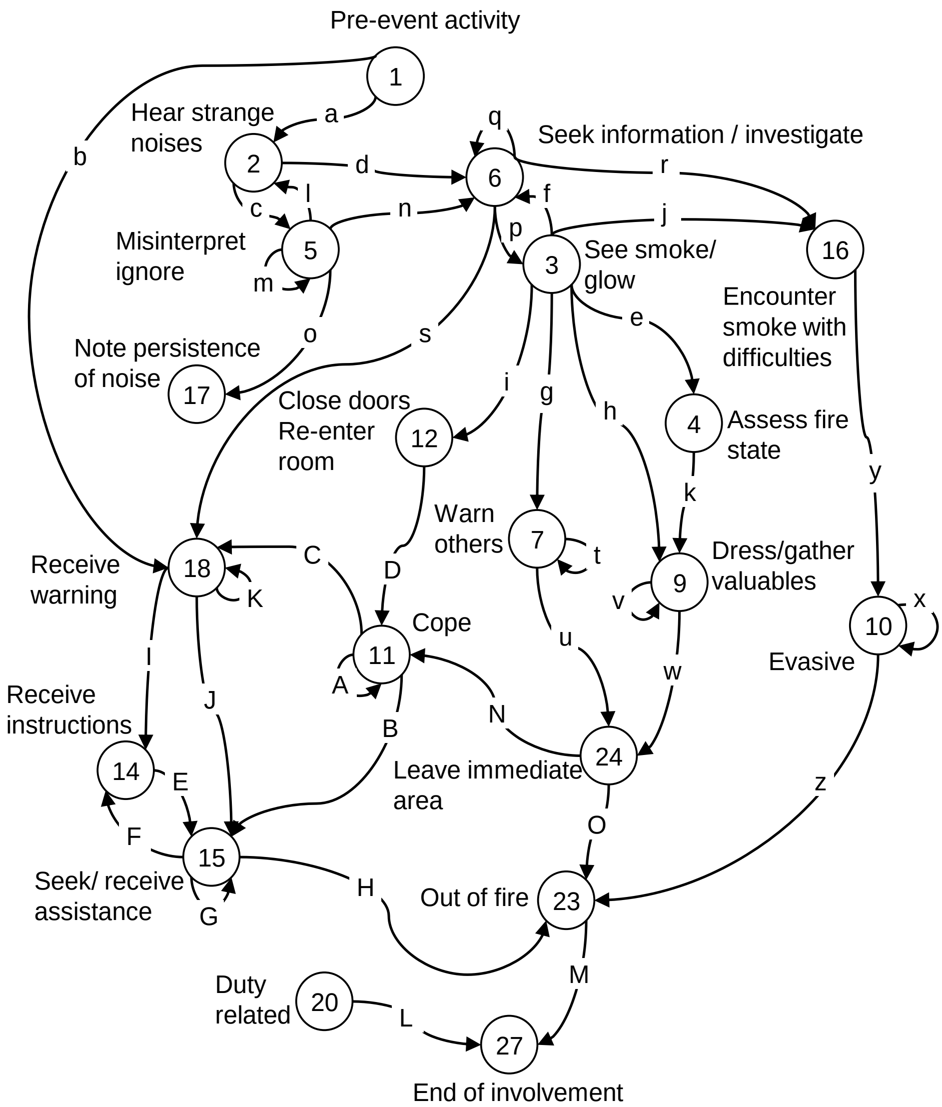

Figure 4.3. Transitions investigated in Study 1

The standardised residuals were investigated for all the transitions for which values were provided in Canter et al. (1980). These are shown diagrammatically in Figure 4.3. Each numbered node represents a group of actions: the numbers of the nodes correspond to the numerical component of the codes in Table 4.1. The arrows represent transitions between the nodes, and point to subsequent acts. There is no meaning in the relative distance between the act nodes, or their position. The standardised residuals for each of the labelled arrows from this study and Canter et al. (1980) are shown in Table 4.2. As seen in Table 4.1, in several instances acts occurred in Canter et al. (1980) which were not reported during the fire drill investigation. This resulted in some transitions for which the standardised residuals were not calculable, as either the row or column totals were 0, and therefore the expected frequency could not be obtained. These incalculable standardised residuals were removed from the analysis.

Table 4.2. Standardised residuals for transitions between acts, shown in descending order for Study 1.

Transition (Figure 4.3)	Study 1	Canter et al. (1980)
a	5.87	24.70
z	4.36	6.70
c	3.75	14.75
w	3.00	4.92
h	2.78	2.91
O	2.36	4.92
v	1.92	4.02
d	1.15	5.20
k	1.10	1.73
K	1.08	6.37
M	1.05	25.49
n	0.71	3.09
q	0.20	1.81
j	-0.10	2.96
o	-0.18	2.68
i	-0.23	5.36
y*	-0.23	1.78
l	-0.25	2.68
e	-0.31	4.92
m	-0.31	4.02
f	-0.41	1.38
p	-0.41	11.21
r	-0.41	1.96
x	-0.51	4.02
b	-0.88	0.34
s	-0.95	3.20
g	**	0.84
t		8.05
u		1.78
A		2.23
B		4.85
C		1.57
D		4.47
E		1.34
F		1.78
G		6.37
H		11.77
I		3.13
J		2.60
L		6.70
N		3.57

* Note: in Canter et al. (1980), transition y was shown to move from “evasive” to “encounter smoke with difficulties”. However, in the example for domestic fires the transition was the other way round. This pattern was logically more likely, and therefore was assumed to be an error in the Canter et al. (1980) hotel fire scenario, and was treated as the direction for domestic fires.

**Blank cells are shown for transitions for which the row total or column total was 0 (i.e. one of the acts for the transition was not reported in this study)

Both sets of data demonstrated non-normality in the Shapiro-Wilk test (this study: W=0.845, df=26, $p<0.01$; Canter et al.: W=0.664, df=26, $p<0.001$) and therefore a Spearman's rho correlation was run. This indicated a significant relationship between the transitions calculated in this study, and those found in Canter et al. (1980) ($r_s=0.459$, N=26, $p<0.05$). The results demonstrated a medium effect size based on the categories provided by Cohen (1988).

Time taken to evacuate

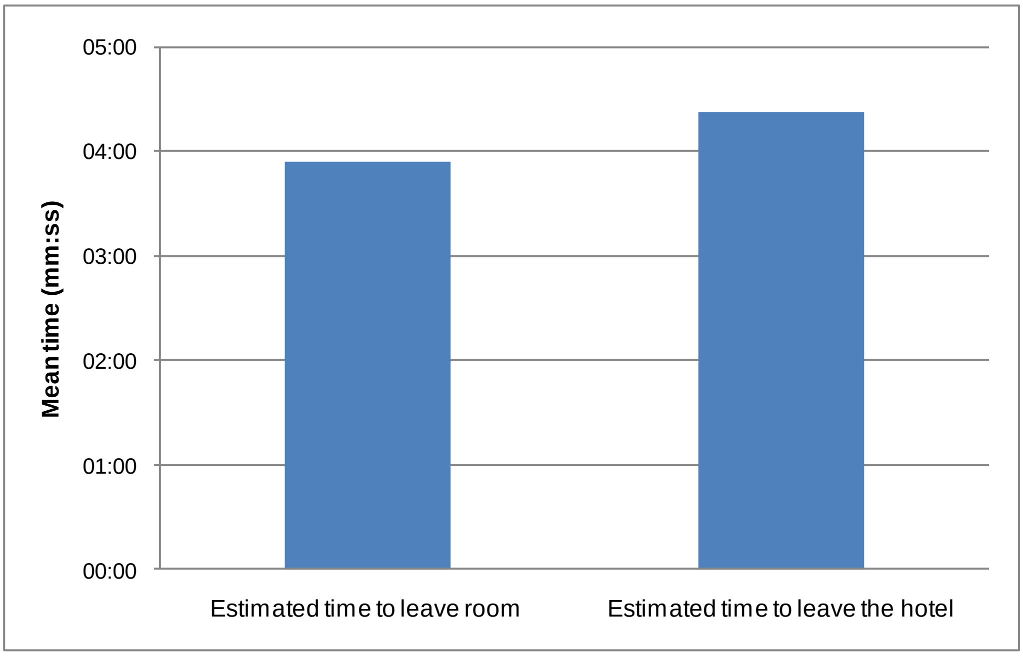

Figure 4.4. Participants' mean estimated times to evacuate.

The times taken to evacuate, based on participants' estimates, are shown in Figure 4.4. Participants were asked how long it took them to leave their room and how long it took them to evacuate the entire hotel (i.e. time to leave room plus travel time to the exit). The times are shown in Table 4.3 in comparison to other research work on pre-movement and total evacuation time. While it is recognised that these times were obtained from different building structures, and the method used to calculate them varied, they serve as a guide to the accuracy of the participants' estimated times in the Saariselkä hotel evacuation.

Table 4.3. Comparison of evacuation times to other research work

	Context	Method	Pre-movement time		Total evacuation time	
			Mean (min:sec)	Range (min:sec)	Mean (min:sec)	Range (min:sec)
Study 1: Saariselkä evacuation	Hotel	Participant estimates	3:54	2:30-5:00	4:24	3:00-5:30
Proulx (1995)	Mixed occupancy residential	Timed fire drills in four buildings	2:30-9:42		3:05-10:57	
Proulx and Reid (2006)	High rise building	Reports from evacuees	5:00	0:00-30:00		
Proulx et al. (2006)	Single family house	Predictions based on scientific reports		2:00-16:10		

Perception of danger

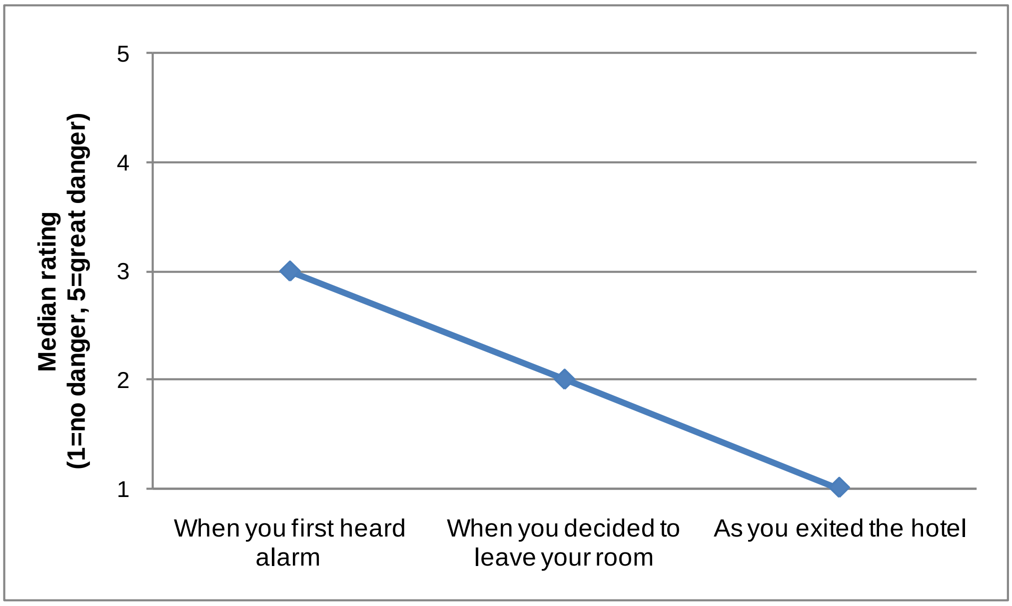

Figure 4.5. Median ratings for perception of danger

Figure 4.5 shows the median ratings for participants' subjective ratings of perception of danger. It demonstrates a decline in the level of perceived danger from when the alarm was first heard to exiting the hotel. In a study of a fire in a nursing home, Edelman et al. (1980) found that occupants who were certain that there was a fire left their rooms immediately; those who were less certain took longer to investigate or returned to their rooms to prepare to evacuate. Evidence can be found for similar behaviour in the Saariselkä hotel fire evacuation, as the lower perception of danger (only reaching a middling score when the alarm first sounded) may be responsible for a less urgent response, shown in the perceived evacuation times (Figure 4.5) and high number of investigative actions shown in Table 4.1.

Route choice

The reasons for participants' route choices are shown in Table 4.4. It can be seen that familiarity dominates, with all participants choosing to evacuate via the main exit, rather than through an emergency exit. The use of familiar exits is reported in several studies of human behaviour in fire (Proulx, 1995; Shields and Boyce, 2000; Mawson, 2005; Xudong et al., 2009).

Table 4.4. Participants' reasons for route choice

Participant	What influenced your choice of evacuation route?
A	"I chose the way I was used to taking"
B	"There was no smoke, no panic, so I took the route I knew. I had checked the evacuation route the evening before, I could've used that one, but I didn't even think of it, just decided to go through the main exit."
C	"I chose the only route that I knew"
D	"Just automatically went back the way I came in – didn't occur to me to look for another fire exit."
E	"Normal way to exit"

Position in evacuation sequence

Finally, it is worth commenting on the participants' perception of their position in the evacuation sequence, i.e. their response to the question "Which of your colleagues were already outside when you left the hotel?" Three of the participants reported that they were the first to leave the hotel. While one of the respondents was new to the project and may not have recognised their colleagues, this result does indicate problems with recall, as found in other studies (Edelman et al., 1980; Aguirre et al., 1998; Gershon et al., 2007).

4.4.4 Discussion

The results of this study are discussed below with consideration to the criteria for judging the quality of an approach (Section 3.2, based on Wilson, 2005).

Validity

Despite a small number of participants, the correlation between the frequency and sequences of acts reported in this study and those from a study of behaviour in real hotel fires (Canter et al., 1980) demonstrated significant relationships and medium effect sizes (r_s=0.414, N=52, p<0.01; and r_s=0.459, N=26, p<0.05 for frequencies/ sequences of acts respectively). While variation was seen between the behaviours in the Saariselkä evacuation and the reference study (Table 4.2 and Figure 4.2), the relationships between them were greater than expected by chance. The results were based on non-parametric correlations, and therefore the association is limited to the rank order of the behaviours, not their absolute magnitudes. However, based on the effect sizes, the reported behaviours were considered indicative of those from real fires. This indication of predictive validity in the approach supported previous work in the literature which has found accuracy in evacuees' reports of emergency events (Wood, 1980).

Participants estimated evacuation times were within the ranges predicted through other research work (Table 4.3), which provides further evidence for the predictive validity. Furthermore, the fact that several hotel occupants remained inside the hotel confirmed reports by Proulx (2007) that an alarm by itself will rarely trigger evacuation without additional cues. Proulx (2007) cited the reasons for this as poor recognition of the alarm, lack of confidence due to false alarms, and inability to hear the alarm.

Participants' behaviour was also consistent with their perception of danger as found through other research work (Edelman et al., 1980). Route choice demonstrated a tendency towards the familiar, main hotel exit, as found in other evacuation studies (Proulx, 1995; Shields and Boyce, 2000; Mawson, 2005; Xudong et al., 2009).

Perhaps the most significant concern related to problems with recall, as indicated through confounding reports of the number of colleagues who were already outside the hotel. This is a problem with the method used (participant reports) and would likely be resolved with alternative methods, e.g. video observation.

Sensitivity

As mentioned in Section 4.3, the sensitivity of this approach is dependent upon the purpose of the evacuation. It did provide some indication of the frequencies and sequences of behaviours which would be expected in a hotel fire evacuation. It also gave an indication of the evacuation times although this was only one instance of an evacuation. As reported by Gwynne et al. (1999) the results from one event could not be used to understand the possible distribution of evacuation times.

Ethics

Ethical considerations are necessary with this approach to ensure minimal risk of injury and distress to participants. In the case of the Saariselkä Hotel drill one participant did experience some distress while deciding what action to take. However, the overriding emotion expressed by participants was one of inconvenience. Thus, in this study the ethics of the drill itself were less of an issue than in other applications of fire drills and experimental evacuations (Muir, 1996; Muir et al., 1996; Brooks et al., 2001; Muhdi et al., 2006; Kobes et al., 2009).

The ethics of the data collection method presented less risk of distress than questioning participants of an actual event (e.g. Gershon et al., 2007; Gershon, 2009) who may have experienced a more traumatic situation than a fire drill evacuation.

However, it was explained to participants in the Saariselkä Hotel drill study that they were able to pull out at any point if they felt distressed. Participants were also told not to participate if they suffered any mental ill-health as a further precaution to protect their well-being.

Resources

As for the large-scale emergency response drills, it was difficult to put a figure on the resources required for the exercise. However, given that the evacuation was mainly hotel guests, and that the only official involved was a single Fire Officer, whose involvement lasted around 20 minutes, the resources were less in this case. The cost of the drill was estimated in terms of hundreds of pounds, which was less than those required for the large-scale emergency response drills (Section 4.3.4). It was also less resource intensive than previous studies of this type (Pauls and Jones, 1980a; Kobes et al., 2009).

Regarding the resources of the data analysis, the method involved coding and quantitative analysis of the reported acts. However, given the small sample size, this was still completed within approximately one week.

Generalisability

The predictive validity of the behaviours demonstrated seemed higher in this instance than in the larger scale emergency response drills presented in Section 4.3. Therefore, the generalisability of using drills to predict behaviour may be limited to fire evacuations. Further work was required to understand whether the results can be generalised to all building types.

Concerning the participants, all were highly educated members of a research project which may have influenced their behaviour. Further work was also required to investigate whether a similar pattern of results would be obtained from participants from different backgrounds.

Feasibility

The feasibility of using fire drills to understand evacuation behaviour is reasonably high, given an obligation on businesses to run them annually as part of their fire risk assessment (Business Link, 2009). However, in other industries (such as tourism) it may damage businesses to have too many fire drills, particularly if they are run at night, as in the case of the Saariselkä hotel evacuation.

4.4.5 Conclusions

This study has shown that the reports from evacuees in a hotel fire evacuation in Saariselkä produced behaviours with medium effect sizes when correlated with those demonstrated in studies of behaviour in real fires. The evacuation times were representative of a real emergency, and the acts taken were realistic given the perception of danger. The resources required for the exercise were relatively low, and feasibility of use was high. Based on these conclusions, fire drills were studied in greater depth for *Phase 2: standardised comparison of approaches* when a similar approach to the Saariselkä fire drill study was used to investigate behaviour in an evacuation from a university building (Chapter 8).

4.5 Chapter summary

This chapter has described research into the use of large-scale emergency response drills for predicting behaviour. Study 1 of this thesis has also been presented, which contained analysis of the behaviours demonstrated in an unannounced evacuation from a hotel. The large-scale drill required significant resources, and questions were raised about the predictive validity of some of the behaviours demonstrated. However, the behaviours demonstrated in the hotel evacuation proved indicative of those from real evacuations. It was concluded that fire drills should be investigated further as an approach for behaviour in building fires/evacuations.

5. The use of experts for predicting human behaviour in fire (Study 2)

5.1 Chapter overview

This chapter presents a study to investigate expert predictions of human behaviour in fire (Study 2). The approach involved nine Fire Officers predicting the likelihood of various statements by using the paired comparisons technique (Sinclair, 2005). The results were compared to a reference study of behaviour in real fires (Canter et al., 1980).

5.2 Introduction

As presented in Section 2.8, expert knowledge has been used to create predictions about how people will behave in emergencies (Dombroski et al., 2006; Groner, 2009; Zachary Au, 2009). However, the approach has not been widely investigated, and is limited in these instances to asking for quantitative predictions of people complying with evacuation orders (Dombroski et al., 2006) or task analysis based approaches (Groner, 2009; Zachary Au, 2009). The latter can be used to predict interaction with a specific part of the system (Groner, 2009) or are suitable for highly procedural interactions (Zachary Au, 2009). Thus, more investigation was required to judge the quality of this predictive approach.

This study (Study 2) arose from an opportunity to conduct research with a group of Fire Officers. The aim was to obtain information about the use of experts as an approach for predicting behaviour in emergencies. This was achieved by asking them to make predictions about human behaviour in fire, in a scenario for which the actual behaviours were known. Work based on this study was published by Lawson et al. (2009a).

5.3 Method

The experts used in this study were nine emergency response personnel each with several years' experience as operational Fire Officers. They all had experience of building fires and of other large-scale emergency incidents. A two hour appointment was allocated to conduct the experiment as part of a regular meeting of the officers. The meeting was held in a typical office environment.

The method used for the experts' predictions was based upon the paired comparisons technique, due to its strength as a technique for scaling opinion. This strength comes from the use of comparative judgements, for which people are "notoriously accurate", particularly when compared to absolute judgments (Nunnally, 1967). People also generally find comparative judgments easy to make (Sinclair, 2005). In paired comparisons more data are collected than necessary to create a scale of the entities being compared, and thus the extra data can be used to check the validity of the results (Sinclair, 2005). It is recognised as a more thorough means of scaling response than other approaches such as ranking, which is generally accepted to provide reliable data only for the first and last two or three items due to the limited capacity of humans to rank order larger lists (Nunnally, 1967; Miller, 1994; Sinclair, 2005).

In paired comparisons, participants are asked to make a judgement between two entities on a given dimension. The entities are taken two at a time from a selection, until all possible comparisons have been made. In this instance, they were taken from acts demonstrated in real fires, as published in the study of human behaviour in fires (Canter et al., 1980) presented in Section 2.2. This reference study therefore provided detailed and quantified data on human behaviour in real fires, which were used to validate the predictions of the experts.

Ten of the acts were randomly selected from the reference study of human behaviour (Canter et al., 1980). A limit was set at ten, as with the paired comparisons technique this required 45 individual comparisons, which was considered to be a reasonable limit of what would be achievable within the allocated time. The acts were transformed into statements, for example the act reported as “feel calm/unconcerned” in Canter et al. (1980) was re-written as “I felt calm, despite what was going on around me”. This was to improve ease of comprehension. The re-written statements were checked by an independent researcher to ensure they accurately represented the original acts. The list of re-written statements can be seen in Table 5.1.

Table 5.1. List of statements used in the paired comparisons exercise

A	I was told that there was a fire in the house
B	I went to look for someone even though the landing was starting to fill with smoke
C	I felt calm, despite what was going on around me
D	I went to find out more about what was going on
E	I closed the door to prevent the fire from spreading
F	I got dressed and gathered my valuables
G	I tried to fight the fire
H	I moved away from the room where the fire was
I	I tried to get through the hallway, but couldn't because of the fire and I was choking in the smoke
J	I got out of my house

A list was made in which each statement was paired with every other statement. The order of the list was randomised, and each pair of statements was written on a PowerPoint slide. On each slide, a statement was written to give participants an option if they could not choose between those related to the human behaviour in fire. This was included as it was deemed undesirable to force the Fire Officers into a decision if they expected that someone in a domestic fire would have a similar likelihood of saying either of the previous statements. Forcing them to choose in this situation would at best extend the time taken for the test, and at worse could have reduced their willingness to participate. A typical slide is shown in Figure 5.1.

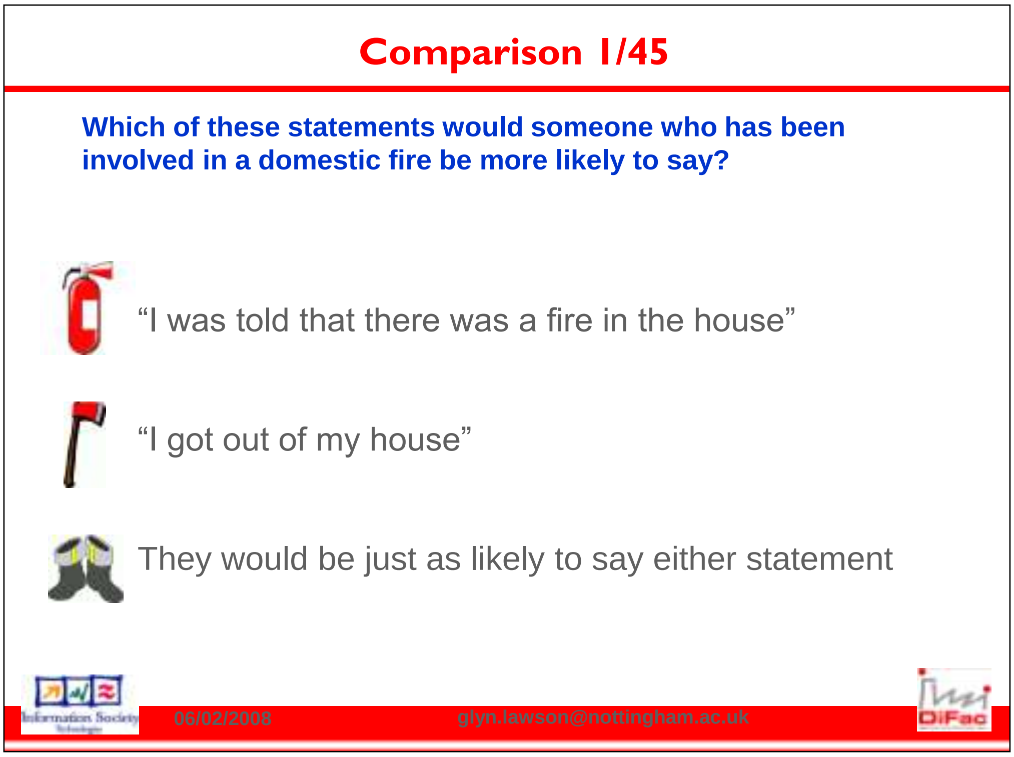

Figure 5.1. Example slide used for the paired comparisons

During the meeting, a presentation was first made explaining the background to the research. Consent forms were issued and collected. Each participant was then given three A4 sheets, one with a clipart symbol of a fire extinguisher, one with an axe and one with a pair of boots. These symbols were also shown next to the statements written on the comparison slides, as shown in Figure 5.1. Participants were asked to read each slide, and then raise the sheet with the symbol which most accurately reflected their opinion. They were asked not to confer, or to look at others' selections. The experimenter progressed through the 45 comparisons, recording the total number of fire extinguisher, axe and boot symbols held up by the experts for each comparison.

5.4 Results

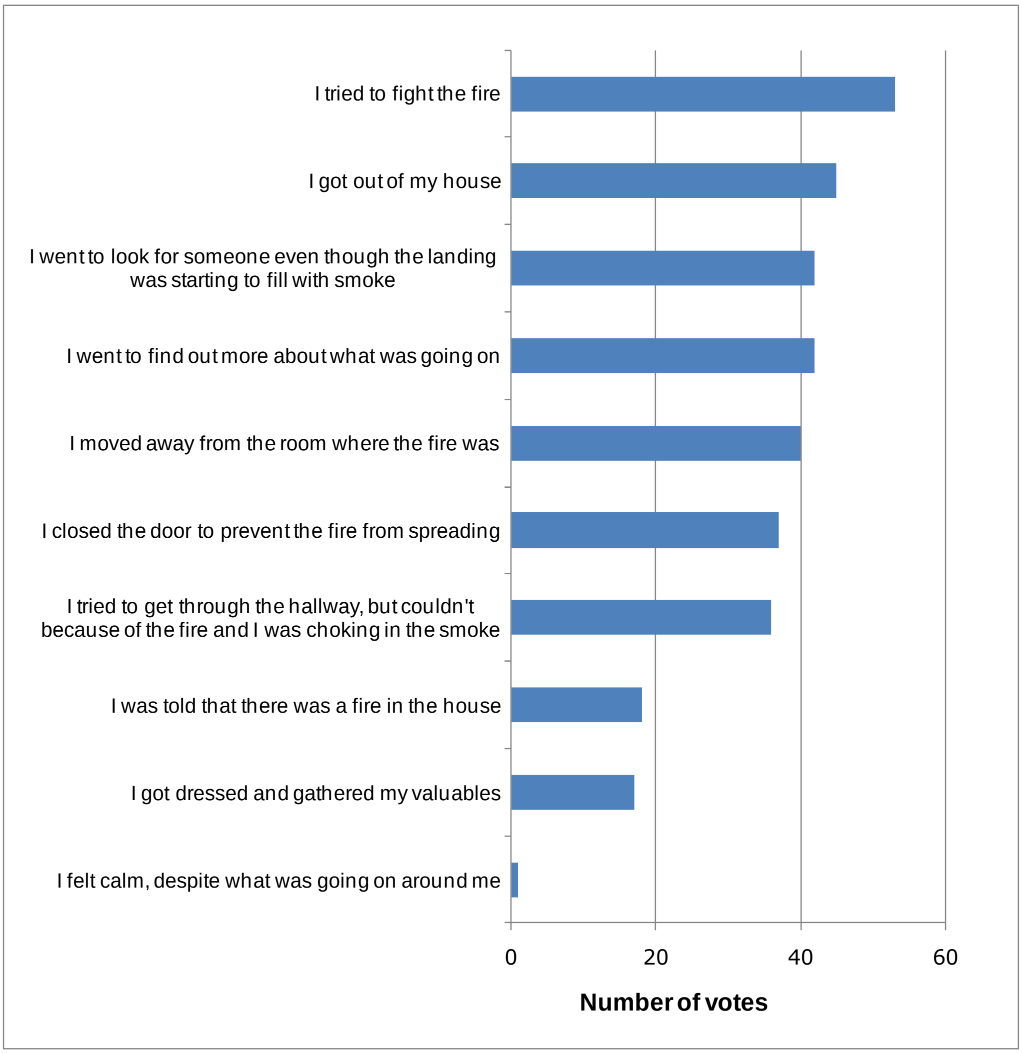

Figure 5.2. Number of votes for each statement

The total number of votes for each statement was counted across all comparisons. This included only the values for statements which were judged as more likely than another. Indecisive votes, for which the Fire Officers could not predict which statement was more likely, were removed from the analysis. The total numbers of votes per statement are shown in Figure 5.2. This indicates the likelihood of each statement being made following a domestic fire, according to the experts.

The frequency of acts predicted by the experts is plotted against the acts reported in the reference study in Figure 5.3. Paired comparisons typically yield high-quality data (Sinclair, 2005): both sets of data were investigated and found to be normally distributed according to the Shapiro-Wilk test (W=0.916, df=10, p=NS and W=0.888, df=10, p=NS). A Pearson's correlation demonstrated no significant association between the predicted acts and reported acts from real fires (r_p=0.236, r^2=*0.055,* N=10, p=NS).

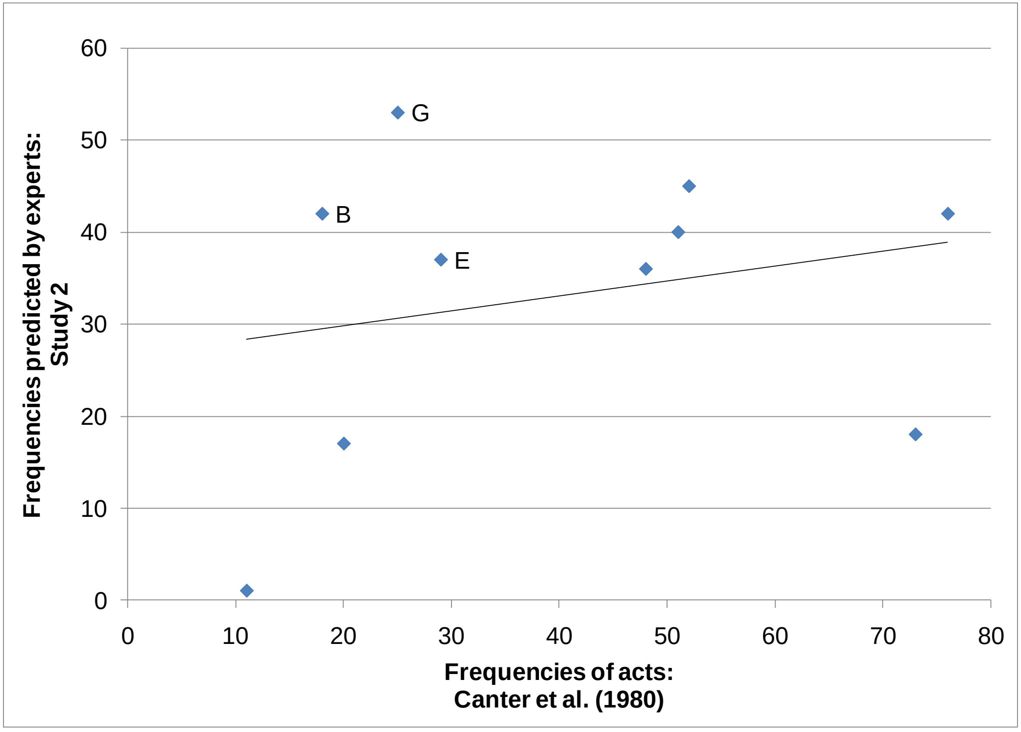

Figure 5.3. Scatter plot showing predicted acts against those demonstrated in real fires. Note data points B, G and E are related to activities associated with fire-fighters.

During discussions after the trial, the experts indicated that their training and experience might have biased their reporting of some of the behaviours. In particular, several experts anticipated over-reporting of the statements associated with the roles of fire-fighters, including: "I went to look for someone even though the landing was starting to fill with smoke"; "I closed the door to prevent the fire from spreading" and "I tried to fight the fire". These are shown as data points B, E and G, respectively in Figure 5.3. The correlation was re-run without these three values, and while it remained non-significant (r_p=0.641, r^2=*0.411*, n=7, p=NS), the correlation coefficient demonstrated a larger effect size.

To interrogate the data further, the method described by Sinclair (2005) for converting paired comparisons data to averaged z-scores was applied. This reduces the error in any single comparison by taking into account the values for each statement across all comparisons. The frequencies of votes on each comparison were first converted into z-scores, which were totalled for each statement before being averaged. Again, the data showed no significant correlation with those in Canter et al. (1980)(r_p=0.209, r^2=0.043, N=10, p=NS).

5.5 Discussion

The discussion below was based on a comparison against the criteria for judging the quality of a predictive approach (Section 3.2, *based on* Wilson, 2005).

Validity

The findings show that in this instance the predictive validity of the approach was low, as the experts' predictions of human behaviour in fire were not significantly correlated with the behaviours recorded from interviews with survivors of real domestic fires. The experts used in this study self-reported an anticipated over-estimation of behaviours which were typical of fire-fighters' responsibilities. When removed from the analysis the correlation coefficient increased, although the results still failed to show a significant relationship. These results demonstrate that this approach would require development in order to have use for predicting human behaviour in fire. They also highlight the importance of questioning and evaluating any method used to generate predictions of human behaviour in fire, even if it relies on expert knowledge.

One possible opportunity to develop the approach was to repeat the evaluation using a more detailed reference study. Several experts expressed difficulty choosing which statement was more likely due to insufficient detail in the scenario. This was in accordance with the recommendation made by Dombroski et al. (2006) that sufficient detail of the incident is required to obtain expert predictions. The Fire Officers in this study were only asked which statement would be more likely to be made by someone in a domestic fire – no other description was provided. This approach was chosen due to limited detail in the reference study. To introduce information, for example the size of the house or demographics of the person who made the statement, might have introduced confounding variables. The study by Canter et al. (1980) was, despite these concerns, one of the most detailed depictions of human behaviour in fire available.

The predictions in this experiment were improved following removal of the acts associated with fire-fighting. Further research could develop a method to adjust for these acts, as simple removal could influence the predicted model of human behaviour.

Resources

A benefit of this approach was relatively low resources. Although gaining access to the experts relied upon existing contacts, conducting the experiment took only around one hour, plus around a day for preparation and couple of days for analysis. Thus, once an expert panel has been arranged, a prediction of human behaviour could be obtained within a matter of days, if not hours.

Re-use of the information obtained is likely to be low, particularly if the scenario must be specific and detailed in order to increase the predictive validity. Thus, a panel of experts would need to be convened for each emergency situation under investigation. However, the analysis method could be re-used, providing some resource saving.

A final point worth mentioning is the choice of experts to use in the panel. This study used a panel of Fire Officers with experience of attending domestic fires. An alternative may be a selection of behavioural psychologists. However, the purpose of this research was to evaluate and develop approaches for ergonomists to use when addressing specific problems. It was anticipated that access to fire-fighters would be more practical than access to psychologists with expertise in behaviour during emergencies.

Sensitivity

A limitation of the approach was that it did not create an initial list of the behaviours; it only provided predictions for their likelihood. It would therefore need to be paired with another method or approach. There are many which could be used: brainstorming, literature analysis, or the talk-through approach reported in Chapter 7. It also did not produce any predictions for evacuation times, such as pre-movement or travel times.

Perhaps the most important limitation of this approach was that it did not address sequences of behaviours. These have been recognised as fundamental to understanding human behaviour in fire (Canter et al., 1980), yet only their predicted frequencies were revealed through this approach. Thus, even while the likelihood of any particular behaviour may be obtainable (with development), this approach will not reveal where in the development of the emergency such a behaviour is likely.

Ethics

This study did not put participants at any physical risk, as it was conducted in a typical office environment. It did present some potential for distress, had it reminded the experts of previous traumatic emergency situations. However, there was no indication of this during the study. It was emphasised prior to the study that participants could withdrawn at any point.

5.6 Conclusions

In this case study, the approach used for predicting human behaviour did not yield predictions which significantly correlated with behaviours reported by people involved in actual domestic fires. The correlation improved when tested without behaviours associated with the roles of fire-fighters, for which the experts reported a bias and likely over-estimation, although significance was still not achieved. Possible development work included re-testing using a more detailed validation scenario, as the experts stated that the comparisons were difficult without knowing the specifics of the fire for which they were making predictions.

In addition to the validity concerns, the approach did not provide information on the sequences of behaviours, or where they would occur within the development of a fire. This was a major limitation for understanding human behaviour in fire situations (Canter et al., 1980). Furthermore, the approach did not identify the behaviours which were to be used in the paired comparisons. Thus, although the study was conducted in a short period of time, and the method used offered the advantage of quantifiable data, the limitations were too great to progress the development of this approach within the scope of this research.

5.7 Chapter summary

This chapter presented a study into the use of experts for predicting behaviour in fire, based upon the paired comparisons technique (Sinclair, 2005). Nine Fire Officers were asked to predict which of two statements would be more likely in a fire, for a series of 45 comparisons. Despite producing quantifiable results in a short period of time, the frequency of the behaviours predicted by this approach did not significantly correlate with a reference study of behaviour in real fires (Canter et al., 1980). Furthermore, the approach could not be used to reveal sequences of behaviours, or identify the behaviours to be used in the comparisons. For these reasons, the approach was not included in Phase 2: standardised comparison of approaches in Chapter 8.

6. Virtual environments for evacuation studies

6.1 Chapter overview

This chapter presents research into the use of virtual environments for predicting human behaviour in emergencies. It reports development and evaluation work on an emergency training simulator for an EU-funded research project. It also reports the findings of an MSc student project which used a Second Life virtual environment for evacuation drills.

6.2 Introduction

The literature review introduced studies which have used virtual environments (VEs) as tools for investigating human behaviour in emergencies (Meguro et al., 1998; Mantovani et al., 2001; Gamberini et al., 2003; Mol et al., 2008; Chittaro and Ranon, 2009; Smith and Trenholme 2009). This section presents early work with VEs which justifies their inclusion in Phase 2: standardised comparison of approaches (Chapter 8). The research was composed of two main projects: the development of an emergency training simulator as part of the DiFac (Digital Factory for Human-Oriented Production System: IST-5-035079) EU-funded research project and a Masters Student dissertation project supervised by the author. Both pieces of work involved input from other people and therefore the author's contribution is clarified in both cases. The work provided initial analysis of VEs as an approach for predicting human behaviour. Furthermore, the influences of these projects on the research direction were important and were therefore summarised in this thesis.

6.3 Development of an emergency training simulator

6.3.1 Introduction

As mentioned above, an emergency training simulator was developed as part of the DiFac (Digital Factory for Human-Oriented Production System: IST-5-035079) EU-funded research project. A digital factory is a virtualised environment which can support manufacturing activities including factory layout planning, product design and training (DiFac Consortium 2006). The aim of DiFac was to develop technologies to increase the competitiveness of European small to medium enterprises (Lawson et al., 2007b). The author's role in the DiFac project was to provide human factors support to the development and evaluation of the tools developed during the project.

Another main objective of DiFac was to improve safety in factories, which was to be partly addressed through training employees in safe working practices (DiFac Consortium, 2006). A tool was developed which could simulate an emergency scenario, thus enabling training of the worker in response procedures without stopping production for a fire drill, and without exposing them to any actual danger (Lawson et al., 2007b, Lawson and D'Cruz, in press). The training simulator incorporated situated artificial intelligence technology. This enabled the virtual environment to be populated with autonomous agents. These are digital representations of people with their own motivations and goals affected by the stimuli presented in the virtual environment. The agents are thus able to demonstrate a dynamic response to unfolding events in the virtual environment, for example an explosion (Comptdaer et al., 2005). It was believed that with the autonomous agents the training experience would be enhanced, providing a more realistic scenario for the trainee (Lawson et al., 2007b). Thus, while the main development was on a virtual environment, this work also investigated simulation tools to a lesser extent.

Initial work led by the author included the development of a framework which was intended to support the implementation of behaviours and models published in the scientific literature into the simulation tool (Lawson et al., 2007a). Generating this framework raised important themes for behavioural prediction which will re-emerge in this thesis. These included the importance of considering the generalisability of behaviours to scenarios other than those from which they were derived. Also, the framework addressed validation of the predicted behaviours, which in this approach was recommended through the use of experts (e.g. Fire Officers) and established techniques such as DELPHI (Rowe and Wright, 1999). Finally, research for the deliverable report conducted at the development partner's site revealed that the simulation developers had a strong preference for quantitative rather than qualitative behavioural models as these were easier to implement as algorithms in the simulation (Lawson et al., 2007a).

During the development phase, two main evaluations were made of the emergency training simulator. These evaluations were designed and administered by the author, who also collated and analysed the results. They are presented here as the findings give indications regarding the use of VEs as an approach for behavioural prediction.

6.3.2 Evaluation of an emergency training simulator by Emergency Response Coordinators

At this stage in the project (early 2008), the training simulator was focussed on training incident commanders or safety managers in correct procedures in the event of a fire, explosion or other accident within a manufacturing company. The proposed method of training involved the trainee watching a simulation scenario on a 3D stereoscopic screen, and issuing commands to the responders (represented as autonomous agents) verbally. The trainee's commands would be implemented into the system by a trainer, which would affect the motivations and behaviours of the autonomous agents. A screen shot of the virtual environment is shown in Figure 6.1.

The evaluation aimed to elicit feedback on the realism and utility of the tool for training emergency response procedures.

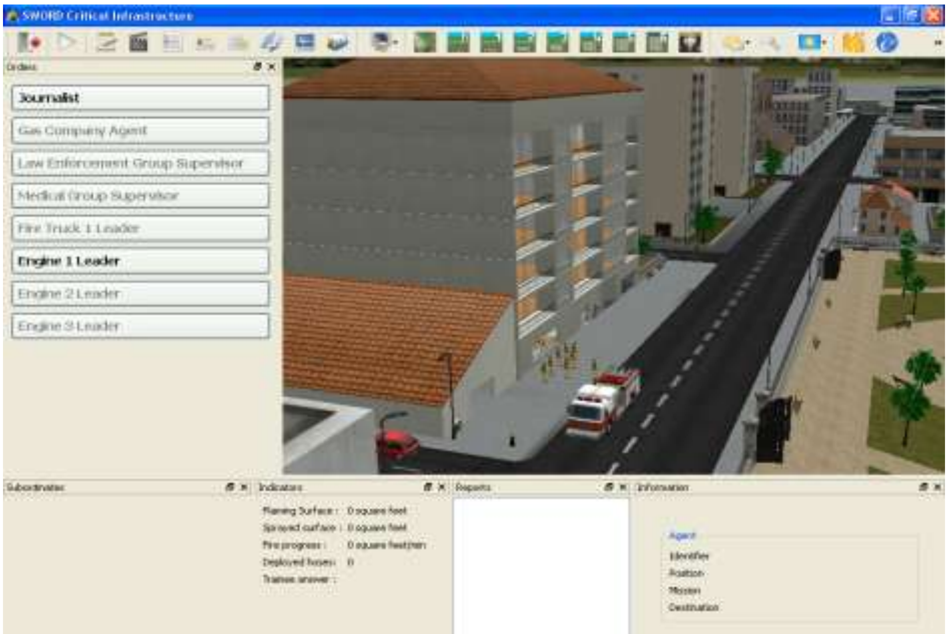

Figure 6.1. Prototype emergency training simulator (Lawson and D'Cruz, in press)

Method

The evaluation was conducted using nine Emergency Response Coordinators, who were all former operational Fire Officers, with several years' experience of participating and coordinating the emergency response to many incidents, in a variety of situations. The evaluation was conducted in a typical meeting room, as part of a regular meeting.

The Emergency Response Coordinators were given a presentation on the background to the DiFac project, and were shown a short video clip of a fictitious training session using the proposed emergency training simulator. During the clip, the various features were explained in more detail and questions from the audience were answered. After the presentation, a general discussion was held, and a semi-structured group interview was conducted, using the following questions:

1. *Could you imagine using this system in your everyday work?*
2. *If so, how?*
3. *Is there anything you particularly like about the system?*
4. *Is there anything you particularly dislike about the system?*
5. *Do you have any recommendations for improving the system?*
6. *Please provide feedback on the visual representation of the environment and people in it*
7. *Please provide feedback on the behaviour of the people within the simulation*
8. *How realistic do you think this behaviour is?*

Notes were taken during the group discussion and semi-structured interview, and are summarised below.

Results

In general, the Emergency Response Coordinators felt that the simulation tool could be useful for training responders to manage an incident. They described it as “very interesting” and “very impressive”. Furthermore, they also thought it could be useful for evaluating response procedures as these are well defined and could have been implemented easily in to the simulation.

Further answers which gave indications of the value of VEs for predicting human behaviour in emergencies are reported below. For the full results, see Lawson et al. (2008).

When asked if they could imagine using the system in their everyday work, the Emergency Response Coordinators indicated that they could, if it were developed to be more relevant to the types of emergencies for which they were responsible, which were generally larger incidents. Thus, they had some belief in the VE representing real-life for it to be a worthwhile training tool. They particularly liked the ability to manipulate the simulation dynamically, and the ability to change views.

Concerning the visual representation of the people in the environment, the Emergency Response Coordinators felt that it should be made more realistic. In particular, the civilians’ walking movements, and a scene when a fire-fighter throws a casualty on the ground were not felt to be realistic and detracted from the simulation. However, they felt that the clip was too short to evaluate fully the behaviour of the avatars. They also stated that when developing the behaviours it was important to recognise the purpose of the training exercise.

Discussion

From this study initial indications were that the Emergency Response Coordinators perceived value in the use of VEs for training. As the behaviours demonstrated by trainees in the VE needed to be realistic to make them suitable training tools, by inference they may also have use for predicting behaviours. In fact, the coordinators seemed particularly sensitive to behaviours in the VE which were not realistic. While the application example focused on training, it provided some support for the use of VEs as an approach for predicting behaviour in virtual environments.

6.3.3 Evaluation by European SMEs

For the next development iteration (mid 2008), the training simulator focussed on training factory employees in emergency evacuation procedures (Lawson and D’Cruz, in press). The intention was for participants to evacuate from a virtual representation of their factory in a fire, learning the correct evacuation routes, and procedures such as checking designated areas for colleagues (represented as autonomous agents) and helping any employees who had disabilities. The solution was executed as a desktop training simulator, running on a standard PC or laptop. It supported self-training, in which the trainee could run an evacuation themselves, then review their performance. The system was evaluated by small to medium enterprise (SME) employees during a project meeting and evaluation session in Turin, June 2008. Once again, the author was responsible for coordinating and managing the evaluation, and analysis of the results (Lawson and D'Cruz, 2008).

Method

The evaluation took place within the premises of one of the project industrial partners. Participants included 10 representatives of SME companies, all involved with manufacturing in Europe, and 12 other members of the DiFac consortium, including development partners and academics.

The evaluation was split into two sessions: a morning session involving SMEs closely associated with the project and the consortium members, and an afternoon session for the participants who were not within the DiFac project.

In the morning, the technologies were presented on an overhead projector. Then, the end-users and the consortium were asked to provide feedback on a questionnaire containing high-level, opened-ended questions, for example: whether the participant could envisage using the technology within their organisation; their likes and dislikes; and any recommendations (Lawson and D'Cruz, 2008).

In the afternoon, the SME representatives received a demo of the emergency training simulator, before being given the opportunity to try it for approximately 10 minutes. They were then asked to answer the same questionnaire as used for the DiFac partners.

Results

The comments made by the participants were collected and entered into an Excel spreadsheet. Similar comments were grouped and the total number of times comments were made within this category was recorded. This approach was based upon a simplified version of theme-based content analysis described in Neale and Nichols (2001). The results which were relevant to the use of virtual environments as a tool for predicting behaviour in emergencies are shown below.

1. Could you imagine this system being used within your organisation?

8 of the respondents stated that they could imagine this system being used within their organisations.

2. How would this system support your current working activities?

Responses included (number of comments shown in brackets):

- Avoid disruption/production losses caused by fire drills (4)
- Train fire procedures/location of exits (3)
- Train in a stressful environment (1)
- As one aspect of the factory planning phase (1)

Thus, indications were made by the respondents that interacting with the VE was sufficiently convincing to replace training in the real world. By implication, the behaviours demonstrated in the VE must have shown some level of realism for them to hold this opinion. However, after stating that the system would be useful for training, one participant commented that "It's like a game, not sure if I react as in reality: difficult to answer. In real situations I might panic. Here, I take time to look around."

3. Please state what you liked about the system

Responses included (number of comments shown in brackets):

- Usefulness for training (6)
- Realism (5)
- User interface/user-friendly (5)
- Good graphics/visuals (5)
- Game-like (2)

Comments relating to the usefulness for training and the realism indicated that the VE may elicit valid behaviours. The "game-like" comments were made in relation to familiarity with a similar interface used on games, rather than implying that the system was purely for entertainment.

4. Please state what you disliked about the system

Responses included (number of comments shown in brackets):

- Interaction with the system (2)
- Not realistic/strange looking avatars (2)
- Concerns about how realistic training is with simulation (1)
- Sickness (1)

All of the above comments reflected some of the potential pitfalls with the use of VEs for behavioural predictions. Interaction with the system referred to difficulties navigating through the VE; for example some participants had difficulties manoeuvring the avatar through doors in the factory. Other concerns about the realism of the avatars' appearances and training using simulation needed to be investigated further. Finally, sickness could be a problem when using dynamic VEs, as these and other health effects of simulators can be serious for some people (Nichols and Patel, 2002).

5. How do you think the system could be improved?

Responses included (number of comments shown in brackets):

- Increase realism (2)
- Better navigation in virtual environment (2)
- Use Wii interface (2)
- Increase realism of avatars (1)
- Make it more intuitive (1)
- More sounds (1)
- More realistic walking speeds (1)

Several comments related to increasing the realism of the system, thus raising concerns about the validity of a behaviour witnessed in an environment which is not perceived as realistic.

Discussion

Once again, the system was generally viewed positively and as a useful tool for training. However, a small number of comments were raised regarding the realism of training in a virtual environment. One of the participants specifically stated that they were unsure if they would behave in the VE as they would in real life, which needed to be investigated further.

Other concerns included the ease of interacting with the environment, which some participants found detracted from the experience. Simulator sickness was raised as an issue, which was a potential concern with VEs.

While this study also focussed on training rather than predicting behaviours, the positive overall feedback justified further investigation into VEs/simulations for behavioural prediction.

6.3.4 Conclusions from the development of an emergency training simulator

The development work on the emergency training simulator for the DiFac project provided initial findings regarding the suitability of VEs and simulation for behavioural prediction, plus broader aspects worth consideration in this research. First, the emergency training simulator VE was generally viewed positively by Emergency Response Coordinators and SME representatives who would ultimately use the system. Among other aspects, they liked the realism of the visualisation and that it could be used as a suitable alternative to fire drills in factories, which can cause disruption to the workplace. However, they were very sensitive to the behaviour of the agents within the simulator, identifying readily behaviour which they believed was not realistic.

Secondly, and in relation to the broader issues identified through this research, the issue of generalisability of behaviours was raised. This concern was identified when attempting to improve the realism of the avatars in the simulation software through implementing behavioural models from the scientific literature. The importance of checking the validity of the behavioural models was realised, as the simulation model would only be as realistic as the data on which it was based. Finally, the developers of the simulation software demonstrated a preference for quantitative behaviour models. Qualitative models were much harder to implement in the behavioural algorithms of their simulation software, an important issue to consider for the outcome of any predictive approach.

The research described above was motivated mainly by the development of a training simulator, and therefore further work was required to evaluate virtual environments as an approach for behavioural prediction, in addition to training applications. This justified a more formal evaluation of virtual environments against the criteria for judging the quality of a human factors approach, which were introduced in Section 3.2.

6.4 Comparison of a fire drill scenario in a virtual environment to real life

Further relevant research in this area included a comparison of a building evacuation in real life to one in a virtual environment. This work was conducted as the research project of an Interactive Systems Design MSc student at the University of Nottingham. It was supervised by the author, whose input included generating the project idea; fortnightly supervision meetings; direction for the development of the virtual environment; and guidance for the experimental design and data analysis. The student successfully completed the MSc in 2009. An overview of the research project is provided below, further details can be found in Yogesh (2009).

The research was composed of two main parts, both of which were relevant to the use of VEs as a tool for predicting human behaviour in emergencies. The first part involved the development of a virtual environment using the Second Life virtual environment (Linden Research Inc., 2009). The second involved an experiment which compared the evacuation times and routes taken from a building in the University of Nottingham in both the real and virtual worlds.

For the development of the virtual environment, Yogesh (2009) first investigated a selection of software development toolkits, based mainly on the Smith and Trenholme (2009) investigation into gaming technology to build virtual environments for fire drill evacuations. However, Yogesh (2009) chose Second Life to build her VE, rather than the gaming technologies proposed by Smith and Trenholme (2009). Second Life is an online virtual environment in which users are able to create and customise and avatar. Their avatar is able to communicate with those of other users through typing or voice-over IP protocol. Users are able to purchase land and construct their own buildings through the development toolkit (Rymaszewski et al., 2007). Yogesh (2009) selected it for this study as the software was free, the land was paid for by the University of Nottingham, a team of developers were available at the University to support students, and there was good availability of online video tutorial and documentation (Yogesh, 2009).

Yogesh (2009) constructed a replication of the Psychology building from the university in Second Life. The building represented an area of approximately 40m x 20m, including one main corridor, with rooms on either side. The original building plan showing the area re-created in Second Life is depicted in Figure 6.2.

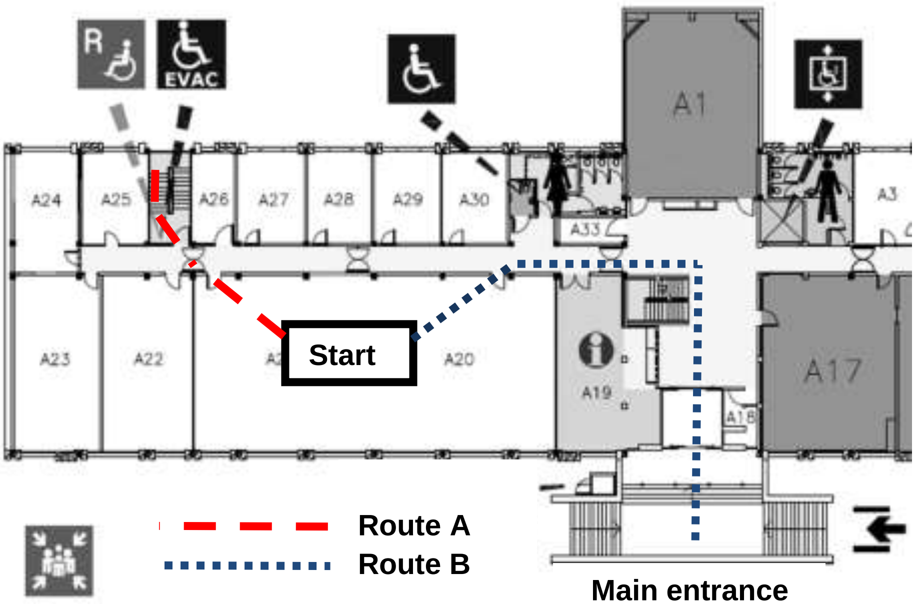

Figure 6.2. Plan view of the building area created in Second Life

A photo taken in one of the rooms in the real building is shown in Figure 6.3. A corresponding viewpoint in the virtual environment is shown in Figure 6.4. While differences are apparent, the VE was close enough to be recognisable.

Figure 6.3. Photo of a computer lab in the real building

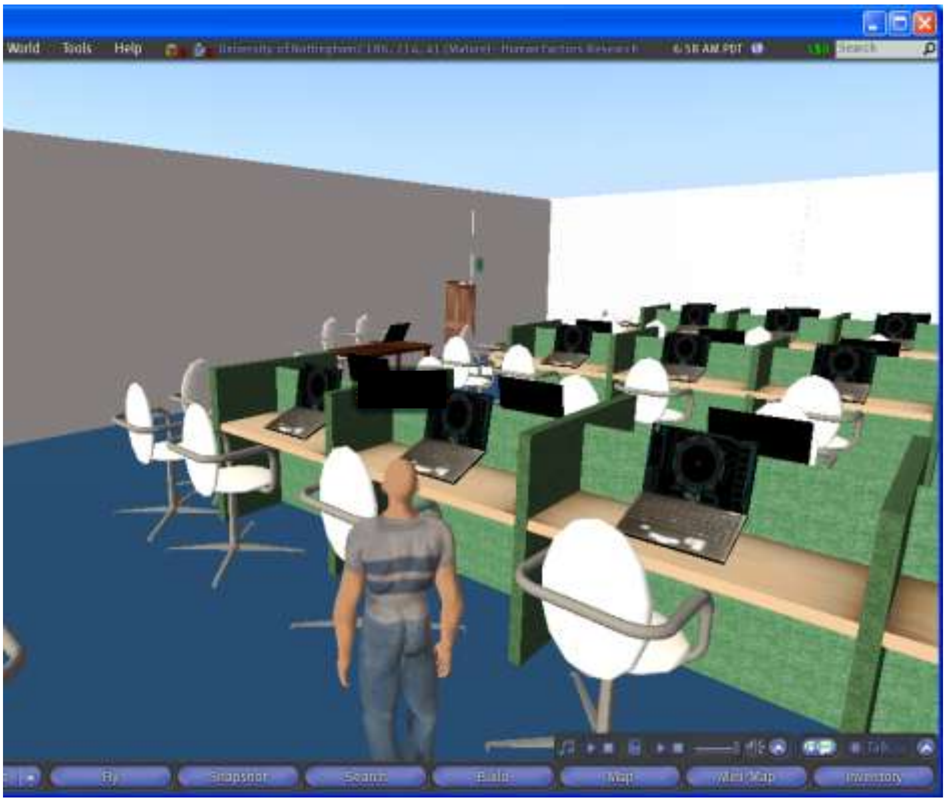

Figure 6.4. Screen shot from the virtual environment

Yogesh (2009) tested 20 participants in both the real and virtual environments in a counterbalanced within-subjects design. Each participant was told individually that when a signal was given by the experimenter, they were to evacuate from the computer room shown above. The signal was given almost immediately after the instructions had been read to the participant. Each participant evacuated by themselves. The same procedure was followed in the virtual environment, with participants controlling avatars out of the building after being given a signal by the experimenter. Again, this was done individually with one participant using the VE at a time. Yogesh (2009) recorded the time it took to evacuate, and the route participants chose in both the real and virtual environments.

Yogesh (2009) found that there was a significant difference in the evacuation time between real and virtual environments ($t=4.58$, $df=19$, $p<0.001$), with participants taking longer in the VE (mean=43.03 seconds, SD=17.91) than the real world (mean=21.79 seconds, SD=5.28). The dispersion was also greater in the virtual environment. Some of the reasons given for the increased time for evacuation in the VE included: getting stuck at doors; getting lost; falling down stairs and other navigation difficulties. Other comments made by participants included that their evacuation behaviour was based on the state of the emergency which was not at all represented in either scenario: there was no indication of fire, smoke, or other people.

Participants had three options for route choice, described as A, B and C. Routes A and B are shown in Figure 6.2; the one participant who took route C in the real building headed along route A, but missed the fire exit signs and moved further into the building. Yogesh (2009) reported high numbers of participants selecting route B in both the real and virtual environments (80 and 85% respectively). This was the main route into the building, and the one with which they were most familiar.

The author of this PhD thesis extended Yogesh's analysis to investigate the statistics of route choice. Initially, a chi-square test was run, and demonstrated no significance ($x^2=1.697$; $df=2$; $p=NS$), indicating no significant change in route selection in the different environments (real or virtual). However these results must be treated with caution as four (66.7%) of the cells had an expected frequency of less than five; moreover chi-square is normally used for a between-subjects design, in this instance participants completed the experiment in both virtual and real environments. To confirm the null hypothesis, McNemar's test was run, although as a requirement for this test is dichotomous data it was necessary to omit the one participant who selected route C in their evacuation of the real building: all other participants selected either routes A or B. The result was also non-significant ($n=19$, $p=NS$) which provided no evidence for a change of route choice between the virtual and real environments. Analysis of the descriptive statistics (Figure 6.5) supports these conclusions.

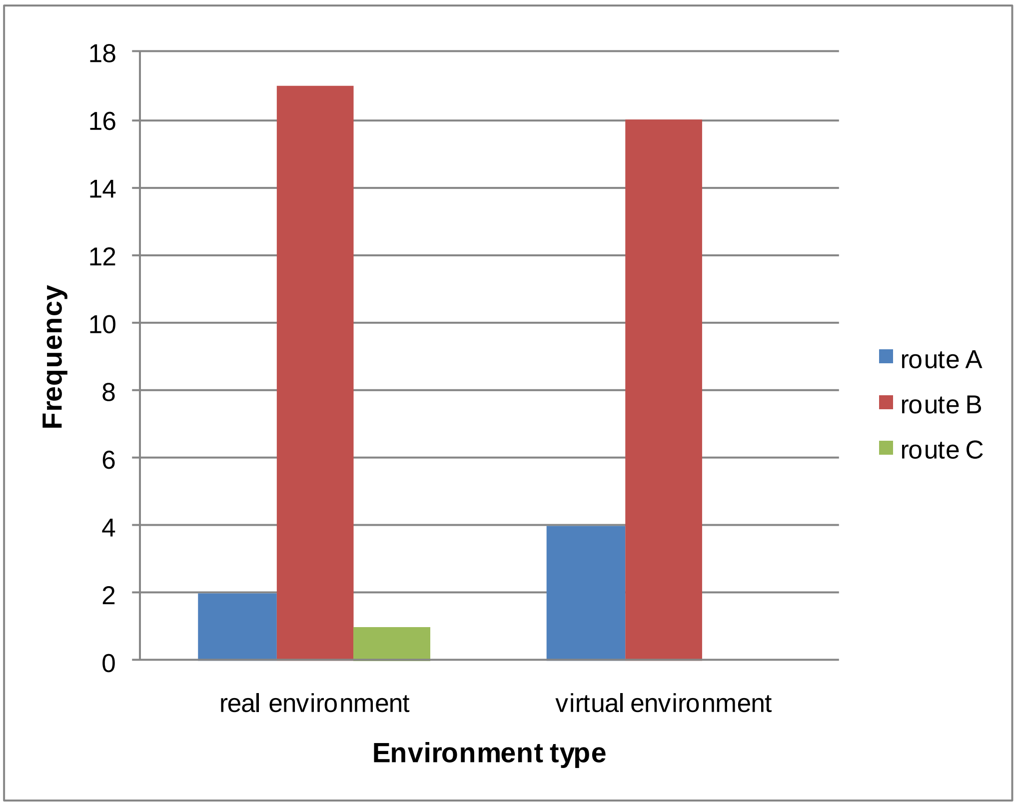

Figure 6.5: Route selection in virtual and real environments (re-created from Yogesh, 2009)

Yogesh (2009) concluded that while the evacuation in Second Life took longer, for most participants their evacuation route was based on the one they were most familiar with in real life. Yogesh (2009) suggested that the familiarity between real and VEs is therefore sufficient to train participants in fire drills using VEs. Yogesh (2009) also proposed that VEs have the benefit of enabling research in which participant route selection and behaviours are easily witnessed in the VE, which could be harder and more expensive through video capture in real environments.

6.5 Initial comparison of virtual environments against the criteria for judging the quality of an approach

From these studies it was possible to make some initial comment on virtual environments and simulation of behaviour against the criteria based on Wilson (2005) for judging the quality of an approach (Section 3.2).

Validity

Considering route choice in the VEs, the experiment in Section 6.4 demonstrated that egress behaviour was similar in the virtual and real worlds, thus indicating concurrent validity. However, this was the most common route in and out of the building, and it was unclear how more complex routes would have compared. Furthermore, Yogesh (2009) found a significant difference between evacuation time between the real and virtual environments, which supported reports by Smith and Trenholme (2009) that evacuation in the virtual environment took longer.

A review of the DiFac training simulator work provides mixed views on the face validity of VEs for behavioural prediction. Several participants felt that simulators have use for training in fire procedures, implying some level of face validity of their behaviours in the VE. However, some comments were specifically received that the training was not realistic, or that people were unsure about whether their behaviour in the VE reflects behaviour in a real emergency.

Reliability

Yogesh (2009) reported that the dispersion of the evacuation times was much greater in the virtual environment (mean=43.03; SD=17.91) than in the real world (mean=21.79; SD=5.28). While this does not provide an indication of test-retest reliability or replicability, the result indicated less consistency in the evacuation times in the virtual environments. Replicability is investigated in greater detail after further testing with virtual environments in Chapter 8.

Resources

Yogesh (2009) managed to construct the virtual environment and conduct the evaluation within three months, with no prior experience of the Second Life interface. Smith and Trenholme (2009) reported that a single developer constructed their virtual environment in three weeks. This type of study could be conducted within 1 to 2 months, once expertise has been gained in developing the virtual environments. For Second Life, sufficient experience could be gained within around 3 weeks.

The potential for re-use is high with this approach, as the VE setting could be used for subsequent testing, for example to investigate design modifications.

Sensitivity

The study by Yogesh (2009) indicated that route choice was the only valid measure to investigate using virtual environments; even that may not be the case in a more complicated scenario. However, if combined with a method such as a post-trial questionnaire then the reasons participants made decisions could also be determined.

Ethics

Participants interacted with desktop computers, in trials which lasted less than an hour, and were therefore at minimal risk of physical injury. One of the main concerns was simulator sickness, as reported by one participant in the DiFac study, although measures can be taken to minimise this (Nichols and Patel, 2002).

6.6 Conclusions

The use of virtual environments showed promise for predicting human behaviour in emergency situations, as demonstrated by a study using Second Life and through a training simulator developed on the DiFac (IST-5-035079) EU-funded research project.

Evaluation against the HF criteria for judging the quality and adequacy of an approach revealed that virtual environments were found to have high feasibility, require low resources and have acceptable ethics, although more work was required to investigate the validity the behaviours. Thus, it was decided to include virtual environments as a predictive approach in Phase 2: standardised comparison of approaches, reported in Chapter 8.

6.7 Chapter summary

This chapter has presented research into the use of virtual environments for predicting human behaviour in emergencies. An initial comparison has been made against the criteria for judging the quality of an approach, based on Wilson (2005). While some concerns were revealed regarding the validity of the evacuation times, the approach required low resources and may have revealed the evacuation routes which participants would take in a real environment.

7. A new approach for predicting the human response to emergency situations

7.1 Chapter overview

This chapter presents a new approach for predicting human behaviour in emergencies. It falls within the broad category of *participant predictions*, an approach which was shown to be under-analysed and infrequently used in previous research (Section 2.7). This type of approach was also noted to have the potential for obtaining predictions of behaviour without putting participants in any physical risk of danger. The approach developed in this chapter combined the talk-through method (Kirwan and Ainsworth, 1992) with sequential analysis (Bakeman and Gottman, 1986). This combination enabled predictions to be made without putting participants at risk of injury. It also permitted investigation of sequences of behaviour.

The chapter presents three separate studies: an initial investigation into the approach (Study 3a; Section 7.2,), development work including an investigation of the reliability (Study 3b; Section 7.3) and further investigation to review the generalisability of the approach (Study 3c; Section 7.4).

7.2 Initial investigation (Study 3a)

7.2.1 Introduction

This study (3a) was an initial investigation into a new approach for behaviour prediction. The approach drew from existing methods for studying and analysing behaviour. This study formed the basis of a publication by Lawson et al. (2009c).

Sequential analysis (Bakeman and Gottman, 1986) was introduced in the hotel fire evacuation (Study 1, Section 4.4) in which it was applied retrospectively to acts people had carried out during the evacuation. It was used again in this new approach for predicting the human response to emergency situations due to its suitability for studying dynamic aspects of behaviour. As introduced in Section 2.2, sequential analysis has already been applied to investigate patterns of behaviour in domestic, multiple occupancy and hospital fires (Canter et al., 1980).

The new approach for behavioural prediction developed in Study 3a aimed to yield the same type of sequential analysis as in the work by Canter et al. (1980), but used individuals' predictions of their likely behaviour in a fire rather than actual data from those who have experienced fires. This enabled comparison of the predicted data with those obtained from actual fires. However, sequential analysis must be applied to observable or recordable behaviours; as a technique it only describes and analyses behaviour, it does not generate any behavioural phenomena. For Study 3a, a talk-through method was selected for generating the behaviours of interest. In this method, participants are simply asked to describe their actions in response to a scenario or statement. It is a low-cost method, as no special equipment is required (Kirwan and Ainsworth, 1992). Furthermore, using the talk-through method meant that a dangerous situation could be described and participants' hypothetical responses, feelings and opinions could be recorded without putting them in actual danger. As it is low cost and quick to implement, it was chosen for Study 3a in lieu of observable or reportable behaviour in a real fire.

7.2.2 Method

Participants

Adverts were placed around the University of Nottingham, and were circulated by email. The adverts explained that research was being conducted into human behaviour in emergencies. They told people not to apply if they had been involved (or had a close friend or relative who had been involved) in a fire, or if they suffered from any mental ill health. These precautions were put in place to avoid causing participants any distress. 20 participants (10 male; 10 female; mean age=35.89, SD=9.75, range=20-52) were recruited from staff and students of the University of Nottingham.

Apparatus

A small meeting room was used for the study. It contained a laptop computer which was connected to an overhead projector.

Procedure

Participants were allocated a one hour appointment. They were asked to sketch a plan layout of their house, and identify who would typically be in their house at night-time. Participants were then asked to detail, in order, every action they would take after hearing a faint cracking noise coming from their kitchen, which on investigation they were told was a fire. They were told to stop their description when they reached the point where they imagined they would be completely out of danger, typically upon exiting their house. If their responses contained insufficient detail, or if they missed a logical stage, they were probed with statements such as "I think you're missing something there" or "I'd like more detail about the stages between those acts". Their stated acts were recorded using a laptop computer, which the participants were able to check as it was projected onto a display screen. The sketches of the participants' plan views of their houses were also recorded. The whole experiment was conducted with procedures which received approval from the University of Nottingham Human Factors Research Group Ethics Committee.

The example used was informed by a number of assumptions. First, it was assumed that the fire started in the kitchen as anecdotal evidence suggests that this is where most fires begin (Bosley, 2003). Participants were only told they could hear the smoke alarm upon entering the kitchen. This second assumption was unsupported by data from real fires, but prevented participants from stating they would exit the property as soon as they heard a fire alarm, a sequence of events which was experienced in the pilot studies but was not supported by Canter et al. (1980), probably due to an increase in the prevalence of smoke alarms since the original study. Furthermore, it was anticipated that many people do not have smoke alarms in their kitchens due to the risk it being set off by typical cooking activities (Haslam, 2010). If participants stated that they would fight the fire, they were told that despite their attempts, the fire would not be put out. Also, if participants reported that they would send someone else to investigate the noise, they were told that this person informed them that there was a fire upon their return.

The hypothetical actions described by the participants were coded using the taxonomy of acts generated by Canter et al. (1980). Where participants described an act which could not reliably be mapped to an act in the original study, new codes and activities were created. The codes given to the participants' behaviours were reviewed by an independent researcher to check for consistency.

7.2.3 Results

Initially, the 49 acts from the original Canter et al. (1980) study were scrutinised, and any which were not expected in this case study due to the example scenario were omitted from the analysis. These were mainly related to actions for which insufficient detail was reported in the original study, and therefore to use them in this study would have introduced inaccuracies. For example: any activities regarding involvement with smoke were omitted as there was no data on the spread of smoke in Canter et al. (1980); any act which arose as a result of an action by another person, aside from a partner returning after investigating the fire, was omitted as again no data was published in the original study on the actions of the other people; struggling with fire equipment and any acts relating to someone who was not from the house containing the fire were also omitted. After this process, 23 of the 49 act categories originally described by Canter et al. (1980) were deemed suitable for comparison. This corresponded to a total of 233 statements (or 65%) of all 361 statements reported in this experiment; 34 statements (9%) were excluded as a result of the process described above, the other 94 statements (26%) were reported in this experiment but could not be mapped onto acts in the original taxonomy. There were 617 comparable acts from the reference study (Canter et al., 1980). The frequencies for the comparable acts, hypothetical (this study) and based on behaviour in real fires

(Canter et al., 1980) are shown in Table 7.1 as a percentage of the total number of comparable acts per study.

Table 7.1. Frequency of acts, shown as a percentage of the total number of comparable acts per study. Acts are shown in descending order of frequency for Study 3a.

Code	Action Category	Frequency (%) Study 3a	Frequency (%) Canter et al.
12a	Seek information/investigate	12	12
8a	Advise/instruct/reassure	12	8
17a	Dress, gather valuables	10	3
25a	Enter area of minimal risk	9	8
3b	Note fire (development)*	7	5
22a	Warn	7	6
5b	Ask advice/request information	6	3
18b	Fight fire	6	4
18a	Fetch things to fight fire with	6	4
13b	Prevent fire spread	5	5
19a	Evasive	5	5
24a	Note/wait for fire brigade arrival	5	7
16b	Go to neighbour's house	3	2
21a	Pass through/enter fire area (investigate etc)	3	6
4b	Disregard/ignore prior stimulus	2	2
3a	Perception of stimulus (associated with fire)	1	2
4a	Interpretation (incorrect)	0	4
10a	Experience negative feelings	0	3
13a	Realise door to fire area open	0	1
14b	Wait for person/action to be completed	0	4
9a	Feel calm/unconcerned	0	2
10b	Experience uncertainty	0	2
10c	Feel concern about occupants	0	3

* This act was a prompt by the experimenter at an appropriate stage in the act sequences i.e. telling the participant they noted a fire when their act sequences led them to enter the kitchen.

A scatter plot of the results, showing the frequency of comparable acts from the original study against those from this experiment is shown in Figure 7.1.

The data did not demonstrate normality according to the Shapiro-Wilk test (this study: W=0.895, df=23, p<0.05; Canter et al.: W=0.878, df=23, p<0.01). A Spearman's rho test demonstrated that a significant relationship existed between the rank order of frequency of acts from the reference study and those from this experiment (r_s=0.694, N=23, p<0.001). The correlation coefficient also indicates a large effect size according the values given by Cohen (1988).

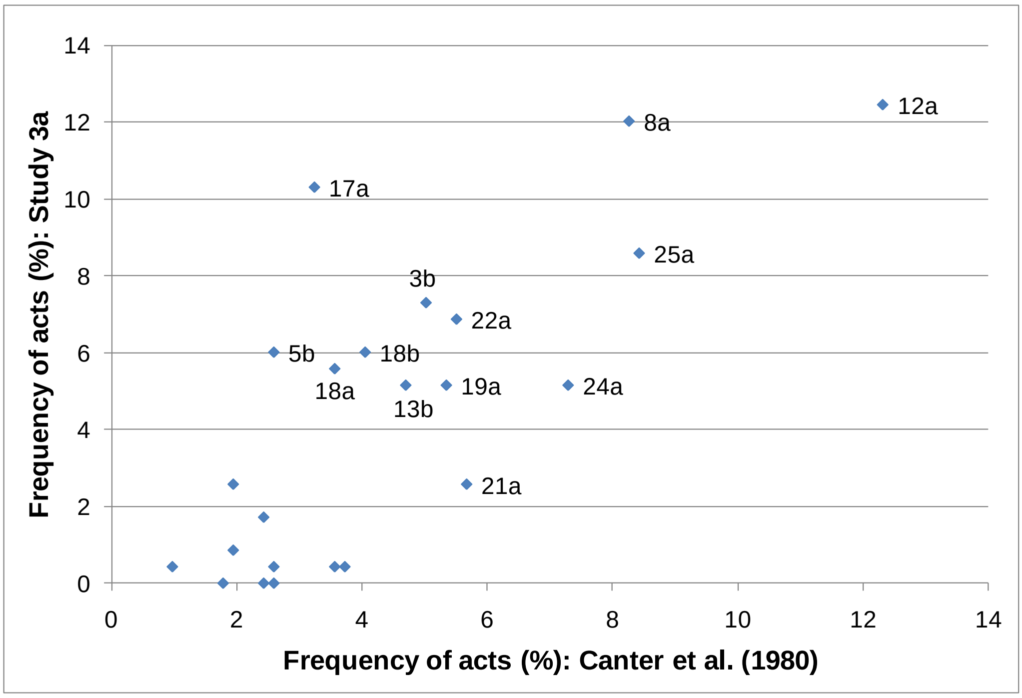

Figure 7.1. Frequencies of acts for Study 3a (hypothetical) and the reference study (reported), shown as a percentage of the total number of acts per study. Labels (shown for higher frequency acts) refer to the codes in Table 7.1.

A matrix was created in which the number of times any comparable act followed every other comparable act was recorded. The transitions were grouped by the numerical component of the code shown in Table 7.1, as per Canter et al. (1980). The matrix was used to generate standard residuals for each transition (observed frequency minus expected frequency[a], divided by square root of expected frequency). The standard residuals were identified for any sequences between the act groups for which values were reported in Canter et al. (1980). These standard residuals indicate the extent to which a transition occurs relative to its expected frequency, through calculation of the base rate for acts. Inclusion of the base rates emphasises cause and effect, i.e. the extent to which one act gives rise to another, rather than reflecting the probability of moving from one act to the next without correcting for the frequencies of the acts. The latter measure may simply reflect a high frequency of occurrence in the subsequent acts. Thus, with standard residuals larger positive values indicate transitions which occurred more often than expected; larger negative values indicate transitions which occurred less often than expected.

The values are shown on a decomposition diagram in Figure 7.2. The strength of association values from Canter et al. (1980) are labelled on arrows which point to subsequent acts; the corresponding values from this experiment are shown in brackets. There is no meaning in the position of the nodes. Values marked (*) are for

[a] Expected frequency = (row total*column total)/grand total

standardised residuals which could not be obtained due to a row total or column total of 0, which consequently rendered the expected frequency incalculable. These acts were not included in the analysis.

The values were tested using Shapiro-Wilk and the Canter et al. (1980) values were found not to be normal (this study: W=0.910, df=21, p=NS; Canter et al.: W=0.720, df=21, p<0.001). The correlation between the standard residuals reported in Canter et al. (1980) and from this study was found to be significant (r_s=0.457, N=21, p<0.05). The effect size was medium according to the classifications provided by Cohen (1988).

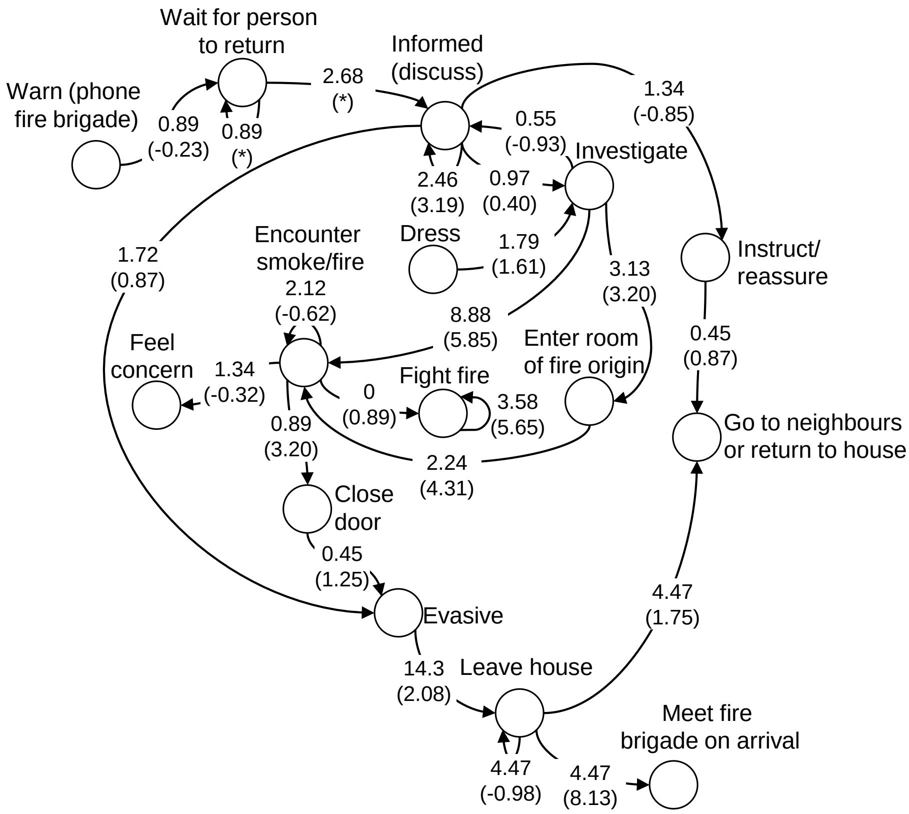

Figure 7.2. Decomposition diagram showing strength of association values from Canter et al. (1980), and those from Study 3a in brackets

7.2.4 Discussion

In this instance, the approach demonstrated a statistically significant relationship between the frequencies of the hypothetical acts reported in this study with those from the original interviews with survivors of real fires, with a large effect size. The sequences of the acts also demonstrated a statistically significant relationship with those from the original study, with a medium effect size. Therefore, in this example, the predictive approach was able to provide an indication of the behaviours shown in real domestic fires, based on rank order correlations. It was achieved with minimal resources; 20 people each gave up to one hour in a basic office meeting room with a laptop, projector and whiteboard.

There were aspects of the predictive approach which required further development. In particular, the validity needed to be assessed in greater detail. Although the face validity of this approach appeared to be low – sitting comfortably in an office environment during the day is very different to be being woken by a fire in a home at night – the approach produced some indication of what actions people will take when measured against the behaviour of people in actual domestic fires. However, this study only demonstrated comparability with another study. If access had been gained to a more detailed incident report from an emergency scenario, this would have allowed for a more fine-grained analysis of the behaviours, and would also have permitted inclusion of some of the acts omitted from this study, such as interaction with smoke. A more detailed account of human behaviour in fire would also have enabled assessment of the timing of acts, as well as frequencies and sequences.

Another limitation of the study was that performance influencing factors such as time-of-day effects, fatigue and training were not accounted for. It is acknowledged that these are likely to influence behaviour. However, the study used for comparison of results (Canter et al., 1980) had no information regarding these factors. Therefore including them in this study would have introduced confounding variables.

The approach only revealed what actions the respondents predicted they would do; it did not reveal the causes of their behaviours. That is, it attempted to answer the question “what would people do?” not “why would they do it?” The latter may be of interest to an ergonomist when attempting to fully understand the behaviours, and how they might be influenced by workplace design or training, for example.

The generalisability and reliability of the technique could not be evaluated based on this case study alone: the study needed to be repeated and applied to other types of emergency situations. These findings led to further work, in particular repeating the study to provide indications of the reliability (Section 7.3). Furthermore, the Canter et al. (1980) study provided results from different types of occupancies, which were investigated to study the generalisability (Section 7.4).

Despite the limitations and further work required, in this example the approach was able to predict many of the behaviours people have demonstrated in real domestic fires, and provided an indication of their sequences. Although it did not identify or provide an understanding of the reasons why people exhibited these behaviours, it was a low-cost and efficient approach, which did not rely on experts or people who have experienced a fire. Therefore, it could have use as a first-pass approach for predicting behaviour in an emergency, which would be supplemented by other more involved approaches at later stages in the development process. For example, the results from this study could be used to increase the realism of the computer representations of people in a simulation tool by incorporating probabilities of various behaviours. The simulation could then be reviewed and further refined through other methods to increase its predictive validity.

7.2.5 Conclusions

This case study examined the use of a new approach for predicting human behaviour in a domestic fire. The technique was low cost and was conducted quickly, yet in this example the frequency and sequences of the predicted acts still demonstrated statistically significant and medium to large correlations with those from another study based on past events. Therefore, it might have use as an approach for predicting behaviour in new situations for which there is no existing knowledge. However, further work was required, in particular to test reliability and generalisability and to develop greater confidence in the validity of the results. It was expected that with development this approach had the potential to offer a variety of uses for understanding how people will behave in an emergency situation. For example, in an industrial setting, the approach could be used to generate information that could help develop appropriate response plans, evacuation procedures, signage, training programmes or building layouts. It was anticipated that despite the short-comings, the approach could be used to quickly obtain indications of how people might behave, with minimal resources.

7.3 Development study (Study 3b)

7.3.1 Introduction

This section presents development work on the new approach for predicting human behaviour in emergency situations, based on the talk-through method (Kirwan and Ainsworth, 1992) and sequential analysis (Bakeman and Gottman, 1986). This work formed the basis of a publication by Lawson et al. (2009b).

The initial predictive study (Section 7.2) demonstrated significant relationships for the frequency and sequence of a selection of predicted actions with a reference study of behaviour from people involved in real fires (Canter et al., 1980). Reliability is also an important criterion for the success of an approach or method, which essentially requires that the same results are achieved upon repeated use (Wilson, 2005). This investigation aimed to investigate the reliability of the predictive approach through a replicability study with different participants. It was also decided to investigate in further detail the predictive validity of the approach by removing from the analysis only two acts. These were related to technology which has become more prevalent since 1980. This differs to the analysis used in the initial study (Section 7.2), in which categories not expected due to the chosen scenario or experimental protocol were removed. This change simplified the administration of the approach as a predictive tool, as users would be likely to find it difficult to identify which acts to remove from the analysis unless justification is obvious, as for the two acts relating to a change in technology over several decades.

7.3.2 Method

Participants

As in Study 3a (Section 7.2.2), recruitment was conducted using posters and email adverts. In addition to excluding people who suffered any mental ill-health or who had been involved in a fire, participants from the initial study (3a) were also excluded. This was because the aim was to investigate whether the results drawn from a different sample demonstrated a significant relationship with the study of behaviour in real fire (Canter et al., 1980).

20 participants (14 male, 6 female; mean age=31.32, SD=5.47, range=23-41) met the application criteria and were allocated a one hour appointment each. Each participant was asked to sign a consent form which emphasised that they could withdraw at any point if they felt distressed.

Apparatus

The study was conducted in a small meeting room, which contained a laptop computer and overhead projector.

Procedure

The methodology replicated that used in the initial investigation (Section 7.2.2). Firstly, participants were asked to sketch a plan layout of all rooms on all floors of their house. This provided a visual reminder to the participants, who were required to consider the layout during the trial. It also familiarised the experimenter with the layout, which helped them understand comments made by participants. Then, participants were asked to imagine that it was the middle of the night and that they, and everyone else who typically sleeps in their house, were asleep in their beds. They were asked what actions they would take if they were woken by a faint crackling noise coming from the kitchen. They were told that this noise was caused by a fire if their anticipated action sequence led them to approach the kitchen. They were told to be reasonably explicit, and were probed for more detail if not enough was given. Every hypothetical act they reported was recorded, in order, until they were told to stop, which was typically when they had exited their house, or reached a state that would remain unchanged until the fire brigade arrived. The predicted acts were recorded on a laptop and displayed on a projection screen so that participants could see their anticipated act sequences. The sketches of the floor plans were also recorded by the experimenter.

As in the initial investigation into this predictive approach (Section 7.2), the predicted acts were coded against a common taxonomy of human behaviour in fire, taken from Canter et al. (1980). Every effort was made to map the predicted acts onto the taxonomy, but if this was not possible, they were mapped onto categories generated from the initial investigation of the predictive approach (Section 7.2), or new categories were created. The frequencies with which each predicted act occurred were recorded.

7.3.3 Results

The taxonomy of acts and frequencies with which each hypothetical act was reported is shown in Table 7.2. This shows the frequency with which each of the acts from the taxonomy were predicted in this study, as a percentage of the total number of comparable acts (N=349). The last column shows the values from the study of behaviour in real fires (Canter et al., 1980), again as a percentage of the total number of comparable acts (N=1189). Note that acts 1-25 were from the original taxonomy (Canter et al., 1980); acts 27a, 28a and 39a were found in the first predictive study (Section 7.2); and act 42a was found in this study, but could not be mapped to any previous act. Only two of the predicted acts were removed from the analysis: 30a: *collect mobile/cordless phone* (10 occurrences) and 34a: *turn burglar alarm off* (1 occurrence). These were removed as the prevalence of these technologies was likely to have increased since the original study in 1980, and therefore the comparison of these acts would have been unrepresentative. For clarity, acts with values less than 1% for both studies have been omitted from the table.

The frequencies of acts were not normal according to the Shapiro-Wilk test (this study: $W=0.675$, $df=55$, $p<0.001$; Canter et al.: $W=0.907$, $df=55$, $p<0.001$). The frequency of predicted acts demonstrated a significant relationship with Canter et al's (1980) study of behaviour in real fires ($r_s=0.323$, $N=55$, $p<0.05$), with a medium effect size (Cohen, 1988). A significant relationship was also reported in the initial investigation (Section 7.2), albeit with a reduced set of acts. These findings show that for this scenario, the predictive approach can reliably produce a significant relationship for the frequencies of acts with those reported by people who have been involved with real fires. However, the predictive validity of the frequency of acts, as shown in the effect size, was at best indicative of those in real emergencies.

Table 7.2. Frequency of acts, shown as a percentage of the total number of comparable acts per study. Acts are shown in descending order of frequency for Study 3b.

		Frequency (%)	
Code	**Action Category**	**Study 3b**	**Canter et al.**
12a	Seek information/investigate	15	6
8a	Advise/instruct/reassure	8	4
18b	Fight fire	8	2
18a	Fetch things to fight fire with	7	2
25a	Enter area of minimal risk	7	4
17a	Dress, gather valuables	6	2
3b	Note fire (development)*	6	3
22b	Phone for assistance	5	1
24a	Note/wait for fire brigade arrival	5	4
28a	Return to bedrooms	5	0
13b	Prevent fire spread	4	2
39a	Wake someone	3	0
19a	Evasive	3	3
16a	Go/gain access to house with fire	2	3
21a	Pass through/enter fire area (investigate etc)	2	3
2a	Perception of stimulus (ambiguous)	1	3
5b	Ask advice/request information	1	1
16b	Go to neighbour's house	1	1
22a	Warn	1	3
4a	Interpretation (incorrect)	1	2
27a	Take/carry pet	1	0
42a	Take weapon/threaten/attempt to scare	1	0
5a	Receive warning/information/instruction*	1	6
10a	Experience negative feelings	1	1
2b	Alerted/awoken (ambiguous)	1	2
3a	Perception of stimulus (associated with fire)*	1	1
14b	Wait for person/action to be completed	1	2
4b	Disregard/ignore prior stimulus	0	1
10b	Experience uncertainty	0	1
19b	Leave immediate area	0	4
20b	Cope with smoke	0	1
23a	Rescued/assisted	0	1
1a	Pre-event actions	0	4
2c	Note behaviour of others (ambiguous)	0	1
3c	Encounter smoke	0	3
6a	Search for people (in smoke)	0	2
6b	Encounter person in smoke	0	1
7a	Observe rescue attempt	0	2
10c	Feel concern about occupants	0	1
14a	Indirect involvement in activity	0	2
15a	Rescue	0	4
20a	Forced back by/breathing difficulties/due to smoke/flames	0	4

* These acts were prompts by the experimenter at an appropriate stage in the act sequences i.e. for 3b, telling the participant they noted fire when they said they would enter the kitchen. For 3a, one of the recorded acts was a prompt by the experimenter; the other was reported unprompted by a participant.

A scatter plot of the frequencies, again shown as a percentage of the total number of acts per study, is shown in Figure 7.3.

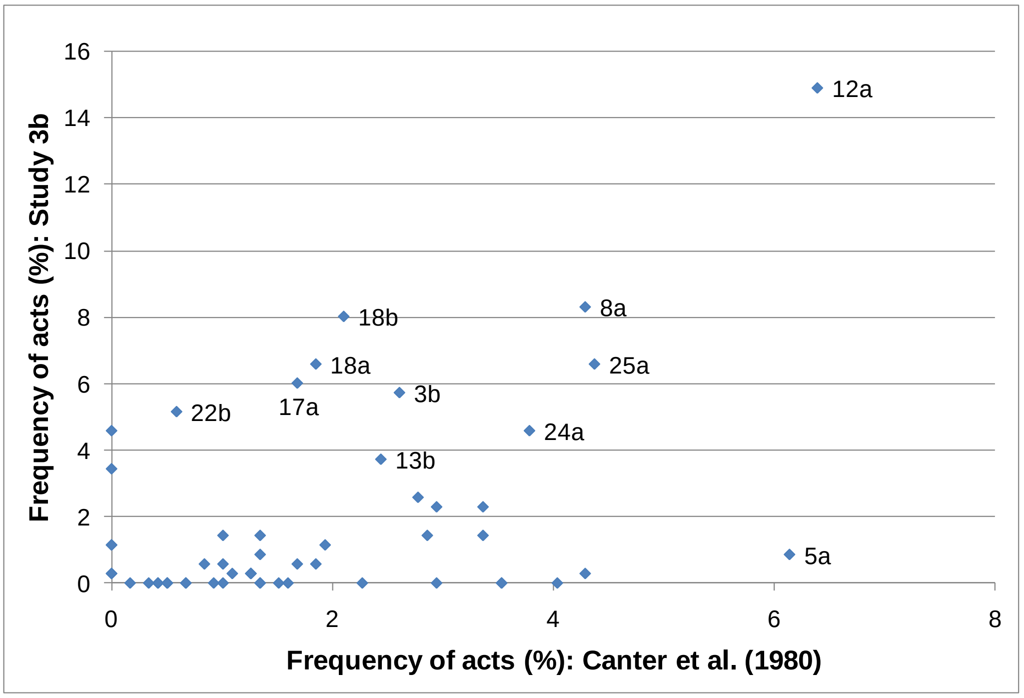

Figure 7.3. Frequency of acts for Study 3b (hypothetical) and the reference study (reported), shown as a percentage of the total number of acts per study. Labels (shown for higher frequency acts) refer to the codes in Table 7.2.

The method used to calculate standardised residuals was similar to that described in Section 7.2.3, although in this instance only two acts (*collect mobile/cordless phone* and *turn burglar alarm off*) were excluded from the transition matrix. The transitions investigated for this study are shown in Figure 7.4, which included all transitions reported for domestic fires by Canter et al. (1980). The standardised residuals for each of the labelled arrows from this study and Canter et al. (1980) are shown in Table 7.3. The act sequences were not normally distributed according to Shapiro-Wilk (this study: W=0.872, df=32, p<0.01; Canter et al.: W=0.768, df=32, p<0.001).

The standardised residuals demonstrated a significant relationship between this study, and those in Canter et al. (1980) (r_s=0.377, N=32, p<0.05). The correlation coefficient shows a medium effect size according to the standard values provided by Cohen (1988). A significant finding for sequence of acts was also seen in the previous application of this approach (Section 7.2), again indicating repeatability in the approach.

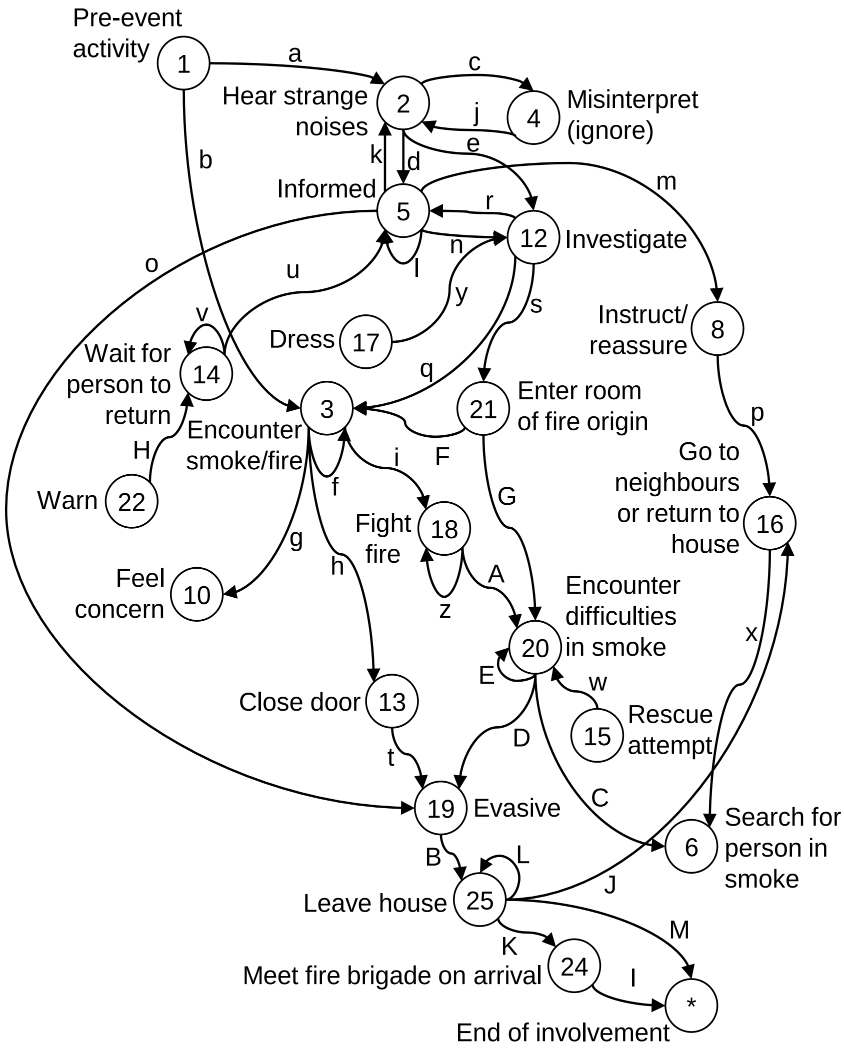

Figure 7.4. Transitions investigated for Study 3b

Table 7.3. Standardised residuals for the transitions, shown in descending order for Study 3b.

Transition	Study 3b	Canter et al. (1980)
z	5.66	3.58
K	5.61	4.47
q	5.59	8.88
u	4.56	2.68
c	3.64	6.26
F	3.30	2.24
J	3.28	4.47
A	2.13	1.79
I	1.97	2.46
y	1.88	1.79
B	1.75	14.30
s	1.52	3.13
g	1.38	1.34
t	1.10	0.45
h	0.95	0.89

Transition	Study 3b	Canter et al. (1980)
n	0.89	0.97
i	0.79	0.00
m	0.36	1.34
e	0.06	5.81
E	-0.06	3.58
p	-0.10	0.45
v	-0.11	0.89
r	-0.11	0.55
D	-0.16	3.90
G	-0.16	1.34
j	-0.18	4.02
k	-0.22	0.89
H	-0.35	0.89
d	-0.39	0.97
o	-0.45	1.72
L	-1.23	4.47
f	-1.23	2.12
a	See footnote*	9.84
I**		9.39
M**		2.24
w		1.79
b		1.34
C		0.89
x		0.45

* For blank cells the standardised residuals could not be calculated due to one act in the sequence not occurring in this study i.e. expected frequency was incalculable, as either numerator or denominator equalled 0. These transitions were removed from the analysis.

** "End of involvement" was not listed in the taxonomy of acts, and therefore strength of association values could not be calculated for acts leading to this one

The findings were investigated further to identify any opportunities to improve the approach through the next stages of its development. A scatter plot in Figure 7.5 shows the standardised residuals from this study plotted against those from Canter et al. (1980). The dashed line indicates a theoretical perfect correlation. A visual inspection indicates that there were under-representations of the transitions B and L in this study. These related to: "evasive" to "leave house" and "leave house" to "leave house". Possible reasons for these under-representations are made in the discussion.

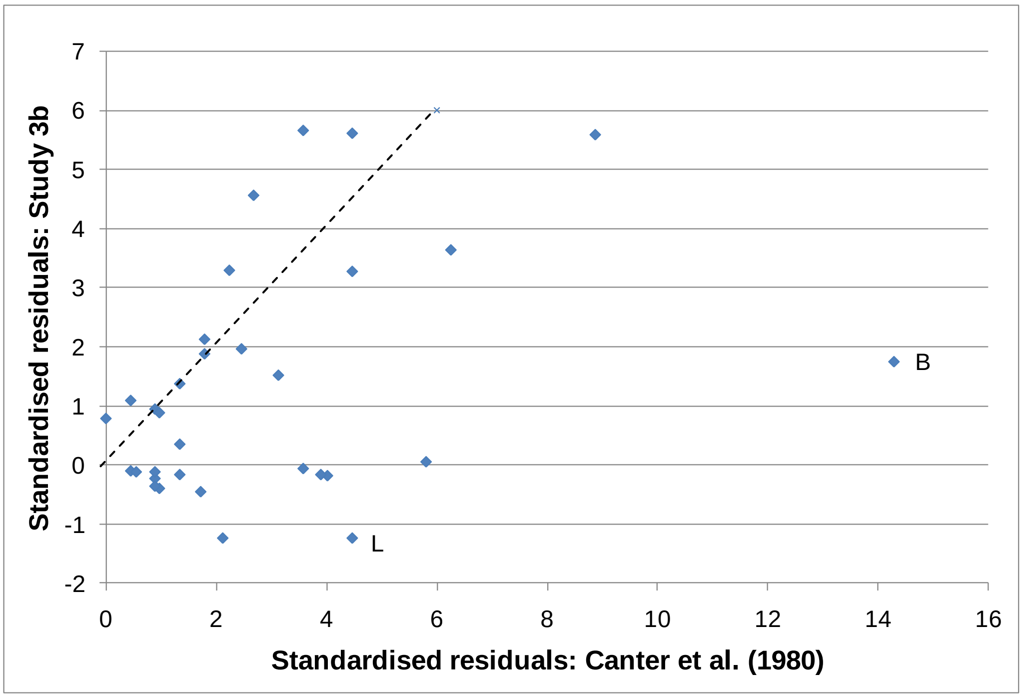

Figure 7.5. Strength of association values for the transitions from Study 3b and Canter et al. (1980). A theoretical perfect correlation is indicted (dashed line)

7.3.4 Discussion

The results demonstrated that in this scenario, the approach used yielded a significant relationship between the frequencies and sequences of acts participants predicted they would take in a domestic fire with those from a study of interviews with survivors of real domestic fires. These significant relationships were also identified in the initial application of this approach (Section 7.2). The correlation between predicted and reported acts for frequencies and sequences, for both predictive studies, indicated that (for the conditions in these studies) the approach demonstrated replicability. It also provided some indication (medium effect sizes) of the behaviours demonstrated in a real fire, although this was based on one reference study and limited to Spearman's rank correlations. However, the results demonstrated that the approach can be used to predict several actions which people have also taken in emergency situations. The resources required were low (a laptop computer, projector and whiteboard, and 20 participants) and it did not require any physical or virtual creation of an environment, only a brief written description of the scenario of interest. With further work, for example an investigation into generalisability to a range of different scenarios, it could be used to predict behaviour in novel situations as determined by the scenario presented to the participants.

While this study demonstrated significant relationships for the sequence of actions, Figure 7.5 revealed that some acts were under-represented in this approach, and that enhancements may have further improved the validity. "Evasive" to "leave house" (B) might have been infrequently observed due to differences in interpretation of participants' responses to Canter et al. (1980). For example, participants in this study stated where they would go to e.g. "I would get out of the house". This was never interpreted as "I would put distance between me and the fire" and then "I would get out of the house", although this sequence was identified more frequently in the Canter et al. (1980) study. In the next stage of the development of this approach (Section 7.4), the interviewing probed for more detail on participants' predicted locations.

Another sequence which showed apparent under-representation was L "leave house" to "leave house". This group actually included the sub-category of "enter area of minimal risk". Therefore, a change to the methodology would result in a participant who predicted that they would move out of their house and then move to another safe place increasing the occurrence of this transition. This was also realised through greater attention to reporting locations in subsequent studies.

Table 7.3 revealed that the sequences related to "end of involvement" (I & M), were not calculable, as these were not listed in the taxonomy of acts in Canter et al. (1980), and therefore were not recorded in the predictive approach. Similarly, "pre-event activity" (a) was not recorded in participants' reports; their sequences began with the first response action. In the further application of the approach (Section 7.4) "pre-event activity" and "end of involvement" were recorded as acts, which allowed computation of these sequences.

It was noted that significant relationships were found for the sequence and frequency of acts despite a stricter analysis procedure than that used in the initial study of this approach (Section 7.2). In this development study, only acts relating to changes in technology since 1980 were removed, whereas in the initial predictive study several acts were removed which were not anticipated due to the experimental protocol or scenario. This change was made to reflect more accurately the anticipated end-use of the approach. A human factors professional would not know which acts were unexpected due to the experimental protocol, and therefore this analysis was more representative of how the approach would be implemented to predict behaviours.

7.3.5 Conclusions

This study was conducted to investigate developments on an approach for predicting human behaviour in an emergency situation, in which participants were asked what actions they would take if they experienced a domestic fire in their house at night. The frequency and sequence of predicted acts were correlated with those from a study of human behaviour in real fires. A significant relationship was found for both, as they were in an initial study of the predictive approach, indicating reliability. This was despite a stricter analysis procedure, which excluded far fewer acts.

Recommendations have been made to improve the methodology, most importantly ask participants for more detail about their locations. These recommendations were implemented during the next development phase of this approach (Study 3c; Section 7.4).

The talk-through approach continued to show promise as a low resource approach, which (with development) could be used as part of the human factors professional's toolkit for predicting behaviour in novel situations. Notably, the approach did not require a physical or virtual mock-up, and did not put participants in any danger.

7.4 Further investigation (Study 3c)

7.4.1 Introduction

This study presents a further application of the predictive approach developed in the studies reported in Sections 7.2 and 7.3. The main purpose was to investigate the generalisability of the approach, applying it to a hotel fire instead of the domestic fire scenario used in the previous two applications. This application of the predictive approach also aimed to incorporate the recommendations arising from the previous studies, for example greater attention was given to reporting locations, and "pre-event activity" and "end of involvement" were included as specific acts (see Section 7.3.4).

7.4.2 Method

Participants

Adverts were placed around the University of Nottingham, and were circulated by email. The adverts explained that research was being conducted into human behaviour in emergencies. People who had been involved in a fire (or had a close friend or relative who had been involved), or suffering from any mental ill health were excluded from the study to avoid causing vulnerable participants any distress.

20 participants (12 male, 8 female; mean age=28.85, SD=8.01, range=18-50) met the application criteria and were allocated a 30 minute appointment each. Each participant was asked to sign a consent form which emphasised that they could withdraw at any point if they felt distressed. It was explained to participants that the testing was confidential.

Materials

Participants were given an information pack showing images from the Losium Hotel, Austria. The hotel was chosen due to the availability of detailed information, including floor plans. The information pack contained views of: the front of the hotel; the back of the hotel; a courtyard at the rear, a hotel room and plans of the ground and first floors (Figure 7.6 shows the first floor plan).

The experiment was conducted in a small meeting room, with a laptop computer and an overhead projector.

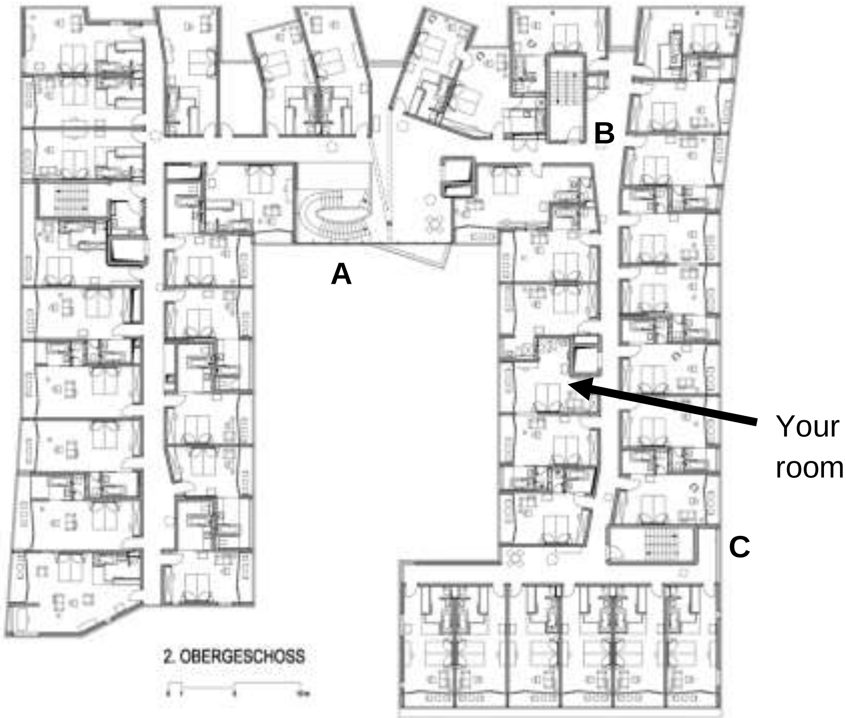

Figure 7.6. First floor plan of the hotel used in the scenario (Saieh, 2008). Only "Your room" was labelled on the plan used in the experiment. Points of interest include A: the main stairwell; B: the stairwell from which participants were told smoke was emitting, and C: the closest emergency exit to their room.

Procedure

Participants were given an information pack showing images of the hotel. The reception area was highlighted, as was the location of the room participants were told they were staying in (first floor). Participants were given a few moments to familiarise themselves with the images and floor plans and the opportunity to ask any questions.

The following narrative was read to the participants:

You are travelling alone and asleep in your bed in the middle of the night. You are awoken by footsteps and voices from outside your room. Please list and describe what actions you would take in this situation. Please report your predicted actions in detail. I may probe you for more detail if necessary. Please do not try to guess the correct answer, or what you think you should do – only the actions you are most likely to take. There is no right or wrong answer.

Figure 7.7. Narrative read to participants

If the participants' predicted acts led them to an interaction with another guest, they were told that the guests had limited proficiency with English. However, if the participant said they would leave their room to investigate, they were told that someone would turn to them and say “fire”. If the participant approached stairwell B in Figure 7.6 they were told that they could see smoke seeping through from under the door. This was included to investigate behaviour in smoke. However, they were told that they could still see through the smoke to the corridor on the other side.

Participants’ predicted acts were recorded on a laptop computer and displayed on a projection screen. They were told to stop their predictions either when they had exited the hotel or if they reached a state which would remain unchanged until the fire brigade arrived. Once complete, each statement was reviewed and participants were asked where each of the hypothetical acts would have taken place. This information was recorded.

After the experiment, the participants’ predicted acts were coded against the taxonomy for multiple occupancy fires reported in Canter et al. (1980). As in the previous applications of this predictive approach, every effort was made to map the reported acts onto the original taxonomy, rather than create new act categories.

In line with the recommendations arising from Section 7.3, greater attention was paid to locations. If the participants’ act sequences led to a change in location which was not reported as an act, this change was recorded as a new act at an appropriate place in the sequence. For example, had the participant stated they would “open door to investigate”, followed by “get my passport”, an intermediate act of “return to room” would be inferred and introduced to their act sequence. Also recommended in Section 7.3 was the inclusion of “pre-event actions” and “end of involvement”. These acts, and any instructions given to the participant (for example “you are awoken by footsteps and voices outside your room”), were included in the act sequence and coded.

The raw data from the first 10 participants was given to an independent researcher, who was given instructions according to the method reported above, and asked to repeat the coding to check for consistency. The inter-rater reliability (Kappa=0.708, N=554, $p<0.001$) was in the “substantial strength of agreement” category from Landis and Koch (1977).

Standardised residuals were calculated according to the method described in Section 7.2.2, although in this instance all predicted acts were included in the transition matrix: no acts were excluded from the analysis.

7.4.3 Results

The taxonomy of acts and frequencies with which they were predicted are shown in Table 7.4, as a percentage of the total number of comparable acts for both the predictive study (N=351) and the reference study (Canter et al., 1980, N=1703). As before, acts with values less than 1% for both studies are not shown in the table.

Table 7.4. Frequency of acts as a percentage of the total number of comparable acts per study. Acts are shown in descending order of frequency for Study 3c.

		Frequency (%)	
Code	**Action Category**	**Study 3c**	**Canter et al.**
6a	Seek information and investigate	17	10
24a	Leave immediate area	11	4
23a	Enter area of minimum risk	9	5
2a	Perception of stimulus (ambiguous)*	7	7
9a	Dress/gather valuables	6	5
1a	Pre-event actions	6	6
18a	Receive information (verbal)**	6	3
2b	Note behaviour of others (ambiguous)*	4	1
10a	Evasive	4	5
12a	Securing environment	4	3
3a	Perception of stimulus (unambiguous)**	3	5
5b	Disregard/ignore prior stimulus	3	1
16b	Overcome hindrance	3	0
18b	Note behaviour of others (unambiguous)	3	3
14a	Give instructions	2	1
14b	Receive instructions*	2	2
7c	Raise the alarm	1	1
16c	Experience negative feelings	1	0
7a	Disseminate warnings/information	1	3
15e	Offer help	1	0
5a	Incorrect interpretation	1	3
11a	Coping (self-related)	1	4
19a	Alter plan (self-initiated)	1	0
12b	Check security of others	1	0
15a	Give assistance	1	1
15d	Seek assistance	1	2
17a	Note persistence of stimulus**	1	1
4a	Correct interpretation	0	1
6b	Approach fire area	0	1
3b	Note worsening of immediate situation	0	3
3c	Note fire development	0	3
4b	Arrive at conclusion	0	1
15b	Receive assistance	0	3
15c	Note arrival of assistance	0	2
16a	Experience movement/breathing difficulties	0	3
18c	Note people who need to be rescued	0	1
20a	Duty related	0	2
27a	End of involvement	(see text above)	-

* These acts included both prompts by the experimenter and participant reported acts (Act 2a: 18 prompts by experimenter/7 from participants; Act 2b: 8 prompts by experimenter/6 from participants; Act 14b: 3 prompts by experimenter/2 from participants).

** These acts consisted entirely of prompts by the experimenter at appropriate stages in the act sequences

Note that 27a "End of involvement" was included to support analysis of the sequence data, but was not included in the taxonomy of acts in the reference study. Therefore, this category was excluded from the frequency analysis. Category 15e "Offer help" was reported in the predicted study, but could not be mapped to any individual act in the reference study. However, it was grouped with the other "assistance" acts.

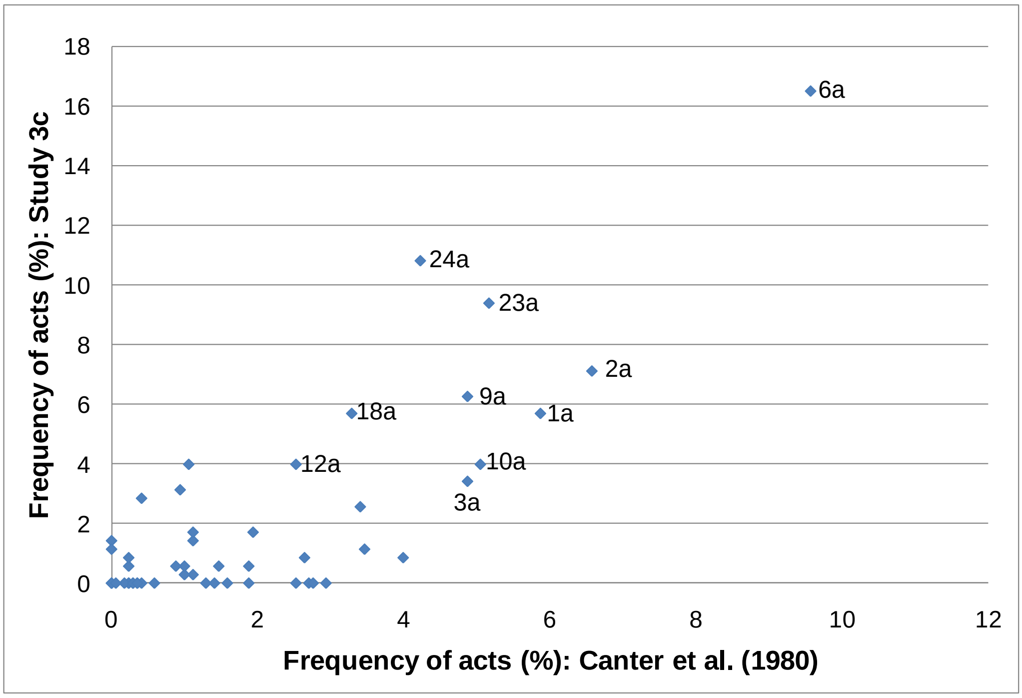

Figure 7.8. Frequency of acts for Study 3c and the reference study, shown as a percentage of the total number of acts per study. Labels (shown for higher frequency acts) refer to the codes in Table 7.4.

The Shapiro-Wilk test was run to test for normality. The results showed that the data for the both the reference study and the experimental data were not normal (Shapiro-Wilk=0.843, df=51, p<0.001 and Shapiro-Wilk=0.655, df=51, p<0.001). Therefore, the data were treated as non-parametric, as in the previous analysis. The frequency of predicted acts demonstrated a significant correlation and large effect size (based on Cohen, 1988) with the results reported in the reference study (r_s=0.572, N=51, <0.001). A scatter plot of the frequencies, again shown as a percentage of the total number of acts per study, is shown in Figure 7.8.

The decomposition diagram demonstrating transitions between the act groups in Canter et al. (1980) was shown in Section 4.4 (Figure 4.3) in the Saariselkä fire investigation. The standardised residuals for each of the labelled arrows from this study and Canter et al. (1980) are shown in Table 7.5.

Table 7.5. Standardised residuals for the transitions, shown in descending order for Study 3c.

Transition	Study 3c	Canter et al. (1980)
A	12.33	2.23
a	11.93	24.70
y*	10.86	1.78
j	10.46	2.96
z	9.31	6.70
o	6.80	2.68
m	5.94	4.02
s	5.94	3.20
c	5.17	14.75
w	4.94	4.92
G	4.26	6.37
t	3.68	8.05
d	3.30	5.20
M	3.00	25.49
O	2.34	4.92
v	2.23	4.02
q	1.93	1.81
x	1.93	4.02
i	0.61	5.36
K	0.39	6.37
l	0.36	2.68
u	0.03	1.78
e	-0.18	4.92
n	-0.23	3.09
k	-0.25	1.73
B	-0.26	4.85
r	-0.33	1.96
D	-0.37	4.47
C	-0.50	1.57
E	-0.52	1.34
F	-0.52	1.78
g	-0.55	0.84
N	-0.57	3.57
J	-0.81	2.60
h	-0.87	2.91
H	-0.87	11.77
I	-1.00	3.13
b	-1.29	0.34
f	-1.42	1.38
p	-1.42	11.21
L	**	6.70

*Note: in Canter *et al.*, transition y was shown to move from “evasive” to “encounter smoke with difficulties”. However, in the example for domestic fires the transition was the other way round. This pattern was logically be more likely, and therefore is assumed to be an error in the hotel fire scenario, and was treated in the direction for domestic fires.

**Note: this transition did not occur for the predicted scenario therefore the standardised residual could not be calculated.

Again, the data were tested with Shapiro-Wilk and proven not to demonstrate normality in the reference study (W=0.667, df=40, p<0.001) and the experimental data (W=0.817, df=40, p<0.001). Testing for non-parametric correlations, the standardised residuals demonstrated a significant correlation between this study and those reported in Canter et al. (1980) (r_s=0.344, N=40, p<0.05). The correlation coefficient indicated a medium effect size (Cohen, 1988).

A scatter plot showing the standardised residuals for this study against the reference study is shown in Figure 7.9. Despite the significant correlation reported above, variation can be seen in the results.

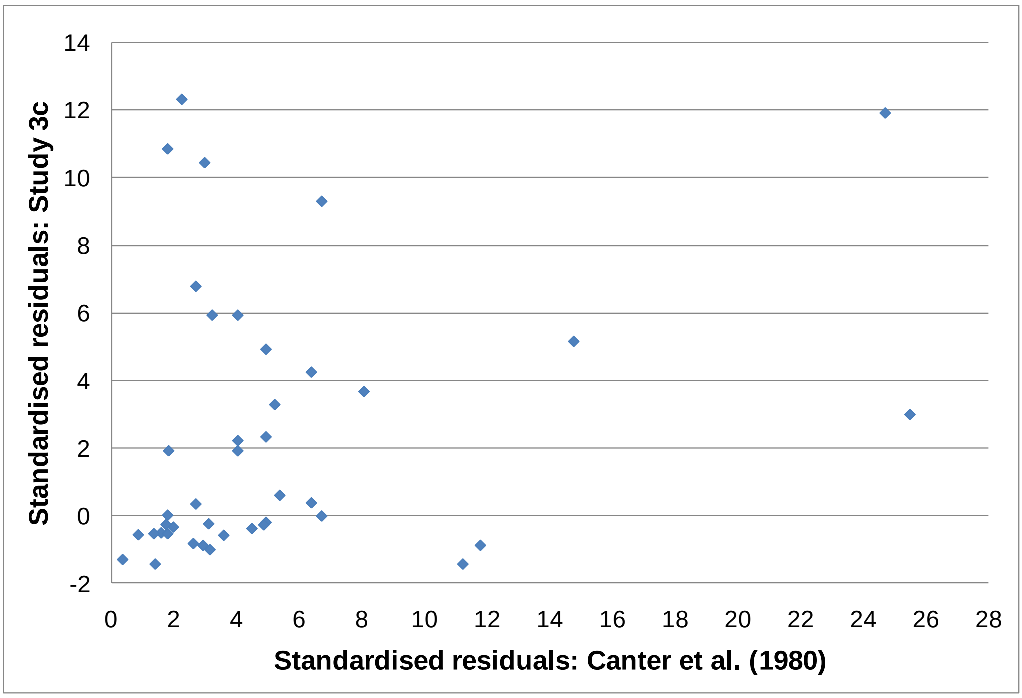

Figure 7.9. Standardised residuals for the transitions from Study 3c and Canter et al. (1980)

7.4.4 Discussion

This study (3c) has built on previous work which aimed to analyse a new approach for predicting human behaviour in emergency situations. As with the previous studies (3a and 3b) (Sections 7.2 & 7.3), the predicted behaviours from this study demonstrated significant relationships (and large/medium effect sizes) with a reference study of human behaviour in real fires (frequency of acts: r_s=0.572, N=51, p<0.001; sequence of acts: r_s=0.344, N=40, p<0.05). The use of a different scenario demonstrated that in this instance the approach was generalisable from domestic to hotel fires.

However, the limitations of the results must be highlighted. Although the results were significant, the data were non-parametric and tested using the Spearman's rank correlation. Thus, while the significance showed a greater than chance association between the variables, it was not possible to make a quantifiable prediction of the likelihood of an act, or act sequence, through this approach: testing was limited to the rank order of the variables. The variation seen in the scatter plot of sequences of acts (Figure 7.9) emphasises that the outcome of the predictive approach does not accurately represent the outcome of real events. That said, behavioural predictions will rarely be completely accurate due to the variability and complexity of human behaviour. This approach provided some indication of the predicted behaviours, accepting that they will not be a depiction of exactly what will happen.

A problem encountered in each application of this predictive approach was the limitations of the reference study. While Canter et al. (1980) was one of the most detailed studies of human behaviour in fire available in the scientific literature, it did not describe the specific circumstances of the fires. This was possibly because the data were collected from several incidents and therefore each situation was somewhat different. Thus, it was impossible to create a detailed scenario to present to the participants, without the possibility of introducing confounding variables. Unfortunately, access was not gained to incident reports which are likely to contain specific information which could be used to create a more detailed scenario and enable finer analysis of the predicted behaviours. These limitations were partly addressed in the standardised comparison in Chapter 8 through the use of a more detailed and controlled reference scenario.

Another difficulty was found through the use of the original taxonomy. This was selected to facilitate comparison between the predicted acts and those reported in the original study. However, the mapping between the predicted behaviours and those in the taxonomy was not always identical; on several occasions they had to be coded according to the closest behaviour in the taxonomy. This issue is also re-visited in Chapter 8, as the taxonomy was customised based on the witnessed behaviours.

7.4.5 Conclusions

This study has presented a new approach for predicting the human response to emergency situations. The work built upon previous studies (Sections 7.2 & 7.3) and investigated the comparability of the predicted behaviours obtained from this approach with those reported from real fires. This study used a different scenario, a hotel fire instead of the previous domestic fire scenario, to investigate the generalisability of the approach. As in the previous studies, the results demonstrated a significant correlation between predicted and actual behaviours for both frequency (r_s=0.572, N=51, $p<0.001$) and sequence (r_s=0.344, N=40, $p<0.05$). However, although significant the correlation coefficients indicate that the approach could not be used to predict exactly what people will do in any given situation. This study only used one reference study; further work was required using a more detailed scenario to determine whether the results would give some useful indication of likely behaviours in an emergency situation.

7.5 Comparison against the criteria for judging the quality of an approach

This section summarises an initial analysis of this approach against the criteria for judging the quality and adequacy of a method, taken from Wilson (2005).

Validity

Considering the predictive validity, each implementation of the approach (the initial investigation, Section 7.2; the development study, Section 7.3; and the further investigation, Section 7.4) produced results which demonstrated statistically significant relationships between the predicted acts and those reported in a study of behaviour in real fires (Canter et al., 1980) for both the frequency and sequences of acts. Effect sizes were all medium or large, according the standard effect sizes given by Cohen (1988). It was recognised that these results did not conclusively validate the obtained behaviour. Problems included: statistical testing was limited to non-parametric correlations for the frequencies of acts and standardised residuals which analysed only rank orders for the values not the absolute magnitudes; the test was not truly “blind”, for example the experimenter had available the categories of behaviour by Canter et al. (1980) when coding the data; and the study by Canter et al. (1980) was subject to the problems of using survivors of emergencies, for example hindsight bias (Section 2.2). However, it was still an achievement to demonstrate significant relationships. The validity of the results from this approach was investigated further in Phase 2: standardised comparison of approaches (Chapter 8).

Reliability

The results from both the initial investigation (Section 7.2) and the development study (Section 7.3), and the further investigation (Section 7.4) demonstrated significant relationships with results from Canter et al. (1980). Thus, finding a similar pattern of results on three applications indicates replicability in the approach.

Resources

The resources were relatively minor: twenty participants (plus associated recruitment effort), each taking up to one hour. The facilities required included a laptop and an office with an overhead projector. It took approximately 2-3 days to define the coding strategy, although the re-use of this was high in subsequent applications. Similarly, the analysis procedure was time-consuming on the first application (particularly calculating standardised residuals), but this was reduced on the second and third studies due to re-use of the Microsoft Excel spreadsheet. Following the established methodology, the total experiment could be conducted in less than two weeks.

Sensitivity

The approach identified likely behaviours and provided an indication of their frequencies and sequences. However, it provided no details of the reason for the predicted actions. This was partially achieved in the standardised comparison (Chapter 8) through the use of a supplementary questionnaire to understand participants' motivations. Similarly, in this study no indication was provided of evacuation times, although this was also addressed in Chapter 8 by asking participants to estimate their hypothetical pre-evacuation and evacuation times.

Ethics

As participants were reassured that they could withdraw at any time, and recruitment excluded those who have been involved with fires or are suffering from an mental illness, ethics concerns were relatively minor. Only one participant in all the studies (sixty participants in total) gave any indication of distress during the experiment, and they were still able to complete the session. There was minimal risk of physical harm, and no deception was required.

Generalisability

The approach has been applied to both domestic (Sections 7.2 & 7.3) and hotel fires (Section 7.4). However, further scenarios need testing before it can be proven to be generalisable to a wider range of emergencies.

7.6 Chapter summary

This chapter has presented a new approach for predicting human behaviour in emergencies. It was based on participants predicting what actions they would take in an emergency. The chapter reported three distinct studies, the first of which was an initial investigation into the approach. Thereafter, it was applied again, incorporating recommendations from the initial investigation, but also to investigate the reliability of the approach. The third study looked at generalisability by applying the approach to a different scenario.

The predictive approach was found to reveal hypothetical behaviours which had significant relationships and medium or large effect sizes (for the frequencies and sequences of acts) with a reference study of behaviour in actual fires (Canter et al., 1980). It also indicated reliability and generalisability through achieving significant relationships with the reference study on each application. The resources were low. The study was limited to identifying possible acts and their frequencies and sequences; it did not reveal the reasons why actions were taken. The approach was applied again in Phase 2: standardised comparison of methods, in the next chapter.

Phase 2: Standardised comparison

Phase 2 presents further research into the approaches which demonstrated the greatest potential success for predicting behaviour in emergencies during Phase 1. The selection of approaches evaluated within Phase 2 is shown in Figure II below.

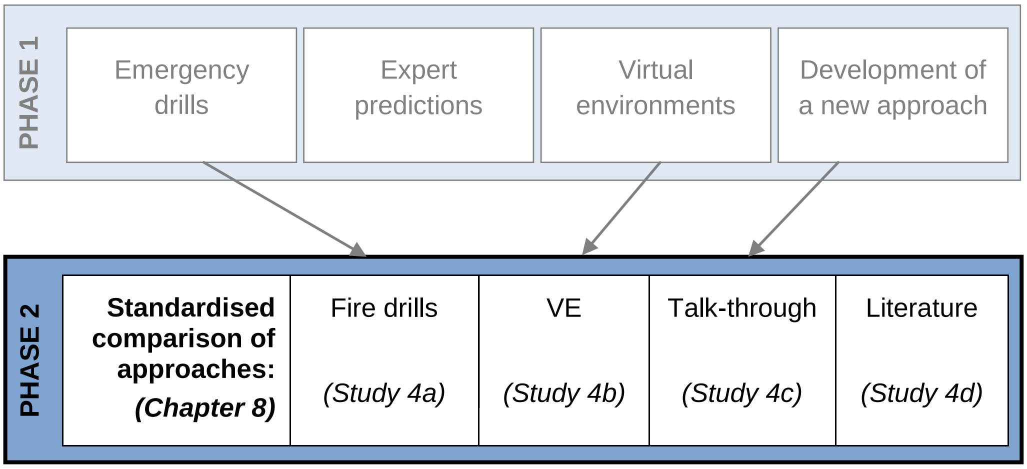

Figure II. Selection of approaches evaluated within Phase 2, showing their progression from Phase 1.

8. Standardised comparison of approaches for predicting human behaviour in emergency situations (Study 4)

8.1 Chapter overview

This chapter presents a comparison of approaches for behavioural prediction applied to the same (standardised) scenario – evacuation from the Psychology building in the University of Nottingham. The applied approaches included:

- An actual evacuation from the real building (referred to as *fire drill*)
- Evacuation using a virtual environment (VE)
- The talk-through approach (as described in Chapter 7)

Each approach was applied to the standardised scenario, which enabled comparison of the behaviours identified, in particular investigation of the concurrent validity of the behaviours through a process of triangulation. The results were also compared to behaviours in emergencies that have been predicted or reported in the scientific literature. The standardised comparison facilitated a controlled evaluation of the other criteria for judging the quality of an approach, for example the resources required, sensitivity, and ethics considerations.

8.2 Introduction

The research work conducted in Phase 1 analysed a variety of approaches for predicting human behaviour in emergencies. This included investigations through which conclusions were drawn about the utility of the approaches. However, a standardised comparison, in which the approaches were applied to the same scenario, was used in Phase 2 to allow for a more critical comparison of their performance against the criteria for judging their quality.

Three of the approaches in Phase 1 gave promising indications for their use in predicting behaviour, and demonstrated potential as practical tools for human factors professionals. These were:

1. **Fire drills.** This approach was introduced in Chapter 4. While this term is often used to describe practice building evacuations, in this study it was used as a short-hand reference to an experimental study of behaviours exhibited during a building evacuation. These approaches have the following common characteristics: real people in a real building, and a false emergency indicated only through an alarm.

2. **Virtual environment (VE).** The advantages and practicality of this approach were demonstrated in Chapter 6. The approach (in this instance) involved creating a VE (environment constructed using computer software and displayed on a typical computer screen) in which participants used the keyboard and touchpad to control their avatar (digital representation of a human). Thus, the VE study involved real people controlling avatars in a collaborative virtual environment.

3. **Talk-through approach**. This technique was introduced and developed in Chapter 7. It involved describing a scenario to participants, then asking them what actions they would take in response to that scenario. The talk-through approach involved real people describing their hypothetical behaviour.

The scenario under investigation was an evacuation from the Psychology building in the University of Nottingham. This building was chosen as the layout was sufficiently complicated to offer a choice of exit routes from the start of the evacuation. Furthermore, the Head of School and School Manager were supportive of the fire drill study, as the results contributed to the school emergency evacuation preparations. In the scenario, the indication of an emergency was limited to sounding the fire alarm, as ethical considerations prevented the presentation of any stimuli with the potential to cause greater distress, such as incorporating smoke. This criterion was particularly important for the fire drill study in the real building as this method presented the additional risk of physical injury. However, the same indication of an emergency was used throughout to ensure comparability of the results.

8.3 Method

The general methodology for this study was to compare the behaviours obtained from each approach, but also to analyse them with reference to the scientific literature. Comparing the behaviours obtained from each approach to each other approach provided an indication of their validity through a process of triangulation (Wilson, 2005). Howitt and Cramer (2011) describe triangulation as a combination of concurrent validity (applying the methods concurrently) and predictive validity (the ability of the each method to predict the results derived from each other method). The results were also compared to those derived from scientific literature to provide further evidence for their predictive validity. Applying the approaches to the same scenario also enabled a detailed comparison of their performance against other criteria for judging their quality, namely sensitivity, ethics and resources. To judge the reliability, and in particular replicability, reference was made to previous applications of each approach to determine consistency in the results (Drury, 2005). An overview of the assessments made as part of the standardised comparison is shown in Table 8.1.

Table 8.1. Assessments made in the standardised comparison of approaches

Evaluation criteria	Approach: Fire drill	VE evacuation	Talk-through approach	Literature
Validity	Correlation of frequency/sequence of acts from all approaches; comparisons of pre-evacuation time; time taken to evacuate and route choice. Consideration was also given to influence of various factors on evacuation.			Grounding of the results using scientific literature
	Correlation between recorded and self-reported acts and evacuation times.	Correlation between recorded and self-reported acts and evacuation times.		
Reliability	Consideration of the results against previous applications of this approach (Chapter 4).	Consideration of the results against previous applications of this approach (Chapter 6).	Consideration of the results against previous applications of this approach (Chapter 7).	
Sensitivity	Comparison of the suitability of the results for human factors applications.			
Resources	Comparison of the resources (time, cost) to apply each approach.			
Ethics	Comparison of the ethics considerations for each approach.			

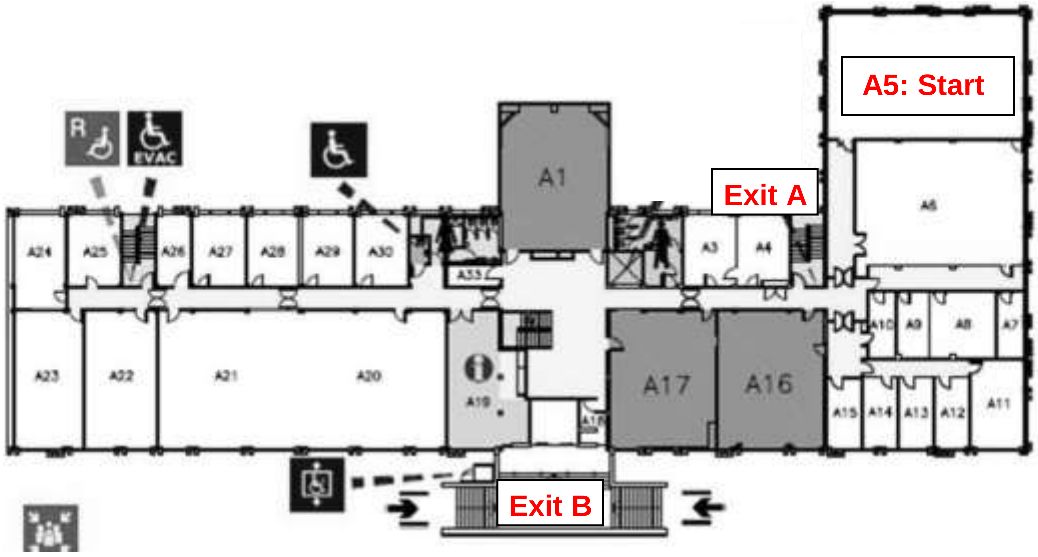

Figure 8.1. Plan view of A-floor, Psychology building. Labels indicate the start point in the evacuation scenario (computer room A5), the closest exit (A) and the main entrance/exit for the building (B).

In each experimental condition, the evacuation scenario began with participants in the A5 computer room, which is located on the A-floor of the Psychology building (Figure 8.1). Participants were physically in the A5 computer room for the fire drill; for the VE study their avatars were in the computer room but participants were located elsewhere; in the talk-through study the scenario asked them to imagine they were in A5 even though the study was conducted in another building. There were two obvious exit routes from the A5 computer room, as shown in Figure 8.1. To leave from Exit A, the closest exit, occupants had to take stairs from the corridor to the lower floor then exit to a car park at the rear of the building. Exit B is the main door to the building, which all participants will have used at some point to gain access to the building prior to the experiment. Participants in the fire drill arrived at the front door of the building at the start of the experiment, before being shown to A5.

Participants were not forewarned that there was going to be a fire alarm in any condition. A between-subjects design was used to prevent participants anticipating the event. In each of the experiments they were given a task with which to engage themselves prior to the fire alarm starting. An overview of the characteristics of each study is presented in Table 8.2; the details are described in Sections 8.3.1-8.3.3.

Table 8.2. Overview of the characteristics of each study

	Fire drill	VE	Talk-through
Participants (N)	22	19	22
Group size	22	Seven groups of 2-4 participants	Individual, although told to imagine there were ~20 people in A5
Representation of building	Real	Virtual (Second Life)	Paper plan view
Movement of participants	Real	Virtual	Hypothetical
Pre-alarm activities	Personality questionnaires	Reading news articles	Hypothetical personality questionnaire
Presentation of scenario	Real	Virtual	Verbal description
Communication between participants	Real	Voice communications through VE	Hypothetical communications
Capture of participants behaviour	Video footage & post-trial questionnaire	Screen recordings & post-trial questionnaire	Hypothetical acts transcribed by experimenter
Experimenter actions after alarm	Remain out of sight; return to A5 after five minutes to give evacuation order	Remain outside building, hidden if possible. Return to A5 after five minutes to give evacuation order	Ask participants what actions they think they would take.
Physical location of trial	Psychology building	Conference room (not in Psychology)	Laboratory (not in Psychology)

8.3.1 Method for the fire drill evacuation

Participants

Recruitment adverts were emailed to the undergraduate psychology students asking them to take part in a navigation study into route choice in buildings. Posters were also put up in the Psychology building in the weeks building up to the study. Thus, recruitment was targeted at participants with some building familiarity. This strategy was taken to reflect the typical building occupants: most of the time, the majority of the occupants will have some knowledge of the building. A total of 22 participants were recruited (5 male, 17 female; mean age=20.2, SD=2.1; range=18-27).

Materials/apparatus

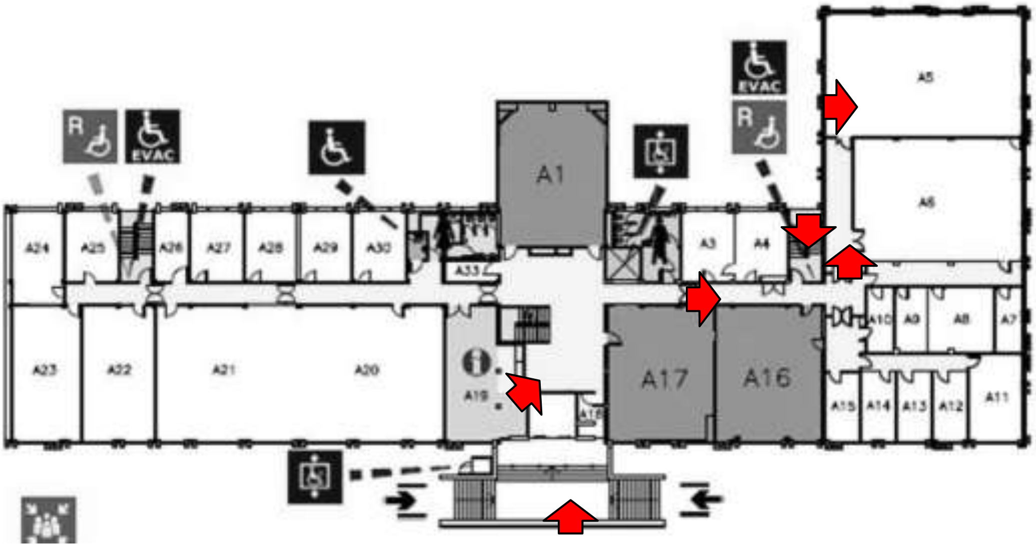

Figure 8.2. Location and direction of cameras on A-floor

The location and direction of cameras used to capture participants' evacuation from the building are shown in Figure 8.2. An additional two cameras were positioned to record the basement exits. The corridor cameras were mounted on door frames looking down on the participants.

In addition to the cameras, a set of post-trial questionnaires were used in the study (Appendix II). These asked for the following data:

- A description of what actions they thought they took
- The estimated time to leave A5, and the building
- Ratings for various influences on their choice of exit, including: familiarity with route; behaviour of other participants; instruction from authority figure; emergency exit signs; and distance.

Procedure

Participants were invited to arrive at the main entrance to the Psychology building on a Sunday morning, a time chosen to minimise disruption to other work being conducted in the building. They were led to the A5 computer room (Figure 8.1) where they were issued with an identification number on a printed A4 sheet (Figure 8.3). Participants were told that they would be captured on video while completing the navigation tasks, and were told that that these numbers would be used for identification.

Figure 8.3. Participants' identification labels

The (fictitious) navigation task was then explained in further detail to the students. They were told that they would be given tasks to complete, for example, “find Professor Smith’s office”. Participants were asked to read and sign a consent form for a navigation study and complete several personality questionnaires. These were merely used to give the participants commitment to a prior activity, and were not used in the analysis. During this time, the cameras were started, and left on until after the study. The experimenter left the room, explaining that they would return shortly to issue the participants with the navigation tasks.

The experimenter immediately left the A5 computer room and gave a signal to start the building fire alarm. They then went up the stairs in the reception area where they could monitor the evacuation unnoticed.

Three other researchers, fully aware of the experiment procedure, were located around the building – two at the basement exits and one outside the main front exit. The role of these researchers was to gather the evacuees after they had left the building, thus preventing them from wandering off. The researchers were also available to help during the trial in case of any unanticipated problems.

After five minutes had elapsed the experimenter returned to the A5 computer room and told all remaining participants that there was a fire alarm and that they should evacuate the building. This time limit was to avoid causing any further distress to the participants if they were kept waiting indefinitely. It was also to minimise disruption to other staff using the building.

Once all participants had evacuated, the fire alarm was turned off. Participants were taken back to the computer room (A5) and asked to complete the post-trial questionnaires (Appendix II).

The whole experiment was conducted with procedures which received approval from the University of Nottingham Faculty of Engineering Ethics Committee. In particular, measures were taken to ensure anonymity of the participants and to obtain consent to analyse the video footage. It was deemed important to de-brief participants on the purpose of the trial. It was explained to participants that they may have behaved differently had they known about the evacuation, and hence the necessity for not pre-warning them of the fire alarm. Furthermore, the Head of School, School Manager and Safety Officer were involved in the planning of the study, and the Head of School was nearby on the day of the evacuation in case of any problems.

8.3.2 Method for evacuation in the virtual environment

Participants

Recruitment adverts were placed across the campus, and emails were sent asking for undergraduate students to take part in a study which would involve completing various tasks in a virtual environment. Participants were told not to apply if they were susceptible to any of the symptoms associated with simulator sickness. They were also told not to apply if they suffered from any mental ill-health, or knew someone who had been involved in a fire. This was to protect the participants in case of any distress caused by the virtual environment evacuation scenario. Participants were also required to have some familiarity with the Psychology building (either have attended lectures or used the computer rooms) to ensure that they had an appreciation of the layout and scale of the building. This criteria was particularly important as the practicality of equipment set up and room availability meant that the experiment was located in a different building to Psychology i.e. the computers were set up in a conference room which was outside the area represented in the VE.

19 participants were recruited (8 male, 11 female; mean age=21.2; SD=5.0; range=18-41). The participants were invited in groups as limitations in bandwidth, physical space and computer availability prevented all participants from being evaluated at the same time. Seven groups were run, ranging in size from two to four participants in each. Recruiting the participants in groups still allowed for social interaction and the influence of other people on behaviour, important determinants on the outcome of an evacuation (Aguirre et al., 1998; Pan et al., 2006). Virtual evacuation studies have rarely been conducted as collaborative exercises, despite recognition of the necessity for doing so (Mol et al., 2008; Ren et al., 2008; Smith and Trenholme, 2009).

Materials/apparatus

This study investigated evacuation from the A5 computer room using the Second Life collaborative virtual environment (Linden research Inc, 2010), run on standard laptop computers. The author of this thesis extended and developed the virtual environment started by Yogesh (2009) described in Section 6.4 of this thesis. It was extended to include the A5 computer room and surrounding lecture theatres and offices. Development work also included the lower floor evacuation routes and cosmetic enhancements to Yogesh's (2009) original environment. The VE was built using a scaled map of the plan view of the Psychology building (similar to that shown in Figure 8.4) and photographs from the real building. The development work focussed on the main evacuation routes to support exits A and B (Figure 8.1). Offices and lecture theatres which were not main evacuation routes were left empty of furniture, although participants were generally able to enter these areas. The building area covered by Yogesh's (2009) original VE, and the area developed for the standardised comparison, are indicated in Figure 8.4. An image of Yogesh's (2009) VE is shown in Figure 8.5; images taken from the modified environment used for the standardised comparison are shown in Figure 8.6 - Figure 8.9.

Figure 8.4. Plan view of the area developed in Second Life. The area in red was originally developed by Yogesh (2009). For the standardised comparison the building area was extended to that shown in green; the lower floor evacuation routes were also developed and enhancements were made to the original work.

Figure 8.5. Ariel view of the final VE developed and used by Yogesh (2009), oriented to facilitate comparison with Figure 8.4

Figure 8.6. VE developed for the standardised comparison (Study 4): exterior view

Figure 8.7. VE developed for the standardised comparison (Study 4): view along corridor

Figure 8.8. VE developed for the standardised comparison (Study 4): A5 computer room

Figure 8.9. VE developed for the standardised comparison (Study 4): lower floor evacuation route

The virtual environment benefited from the voice communications feature in the Second Life environment. This enabled users to talk and hear other participants, within range, in the virtual environment using microphone headsets. The sound was presented in stereo and was sensitive to distance: the volume of the voice communications decreased as participants moved further away from each other.

A recording from the actual fire alarm in the Psychology building was originally loaded into the virtual environment to be used for the fire alarm. However, a pilot study revealed that when played back in the Second Life environment, the sound became distorted and was unrecognisable as a fire alarm. Participants comments included "I wondered if the computer/headphones were damaged", and "didn't realise it was the fire alarm". Therefore, the pilot study data were removed from the results and the sound was replaced by a sound file containing a metal clanger type fire alarm. This lost none of the quality when uploaded into Second Life, and informal discussions with participants revealed that it was recognisable as a fire alarm.

In addition to the VE, the materials used included a consent form and a questionnaire (Appendix III), which was similar to the one issued in the fire drill study. This asked for:

- A description of what actions they thought they took
- An estimation of the time to leave A5, and the building
- Ratings for various influences on their choice of exit.

Procedure

Upon arrival, participants were asked to sign the consent form explaining that their actions and conversations would be recorded in response to nominal and emergency situations. It was emphasised that they could withdraw from the study at any point should they feel distressed, or feel any adverse effects of using the virtual environment. They were asked to complete the checklist of simulator sickness symptoms (Kennedy et al., 1993). In addition, the experimenter provided an introductory briefing, explaining that participants would be asked to complete various tasks in the VE, and that their navigation routes and reaction to stimuli in the virtual environments would be recorded.

The participants were given various familiarisation tasks to complete in the virtual environment to practise controlling their avatars. This was done in groups, so several avatars were being controlled within the VE at the same time. The tasks involved using all the controls and took the participants through different areas of the virtual Psychology building. They were led by the experimenter's avatar within the VE. Participants were familiarised with the different speeds their avatars could move, namely: walk (the default speed, equivalent to 2.16m/s), run (equivalent to 3.15m/s) and slow walk (equivalent to 0.56m/s). The default speed was between the comfortable and maximum mean walking speeds for healthy adults in their twenties (Table 8.3).

Table 8.3. Mean walking speeds for healthy adults (Bohannon, 1997)

	Mean walking speeds (m/s)	
	Comfortable	Maximum
Men	1.39	2.53
Women	1.41	2.47

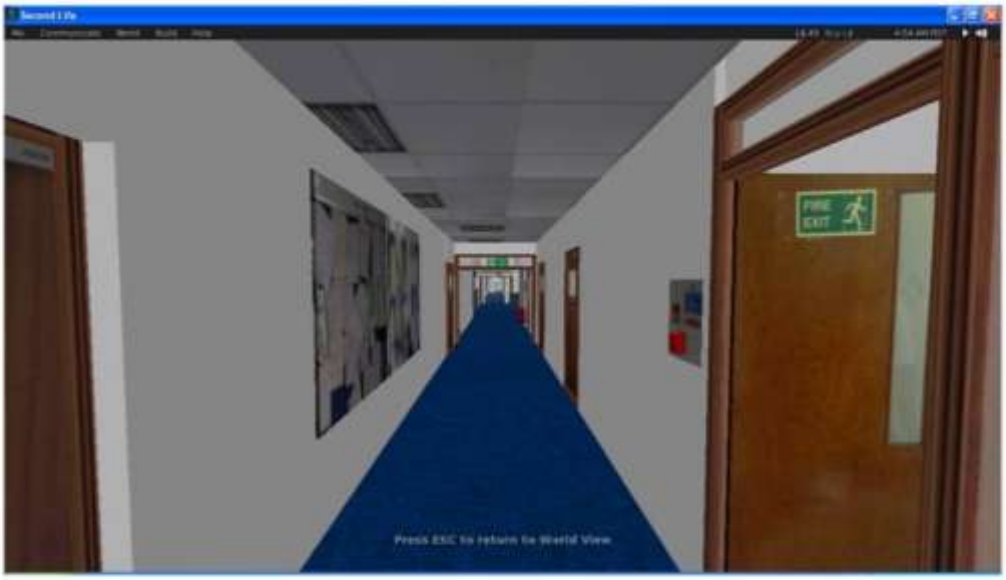

Figure 8.10. Example of "eye-view"

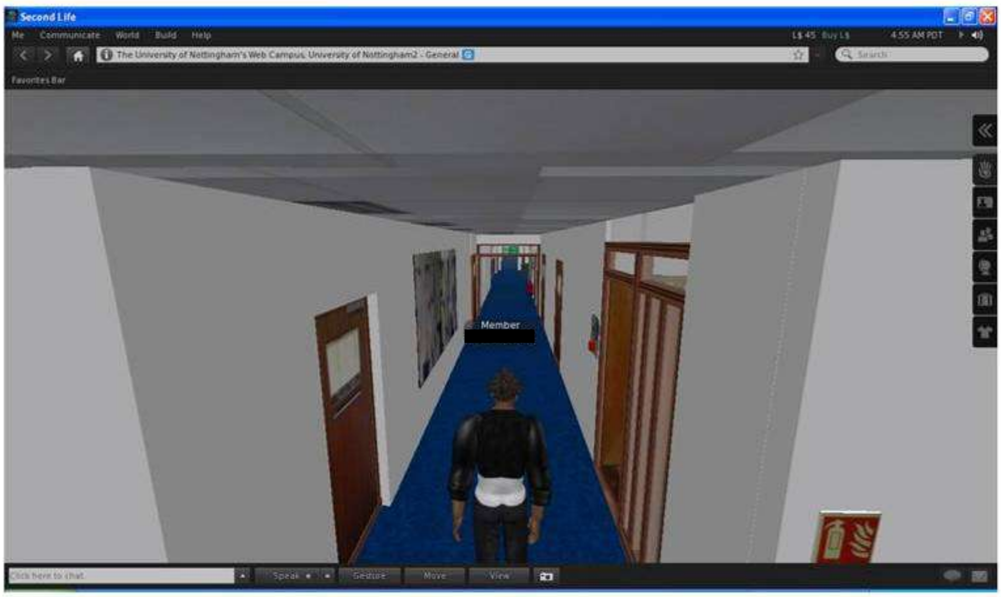

Figure 8.11. Example of "world-view" in which avatar is visible

Participants were asked to communicate through the headsets and remain in the virtual environment for the duration of the experiment, unless they felt distressed or experienced any other negative emotions or adverse effects. They were also asked to use the eye-view (Figure 8.10), which enabled them to control their avatar as if seeing through their eyes. However, in some cases a bug in the Second Life system meant that this feature did not work and participants had to navigate within the VE with a view of their avatar (Figure 8.11).

Within the VE, participants were asked to follow the experimenter's avatar to the A5 computer room (Figure 8.1 and Figure 8.8). They were told to read news articles which were printed on posters on the walls of the computer rooms (Figure 8.12) as they would need the information for a subsequent task.

The experimenter told the participants (through the VE) that he (his avatar) would return with detailed instructions before leaving the computer room to trigger the alarm. He commanded his avatar to wait outside the building such that he could see the participants if they evacuated. After five minutes, the experimenter returned his avatar to the computer room (A5) to tell any remaining participants to evacuate. The participants' avatar movements and behaviours were recorded using screen capture technology, as were the participants' voice communications.

Figure 8.12. Participant reading posters in the A5 computer room

After all participants had controlled their avatars out of the building, they were told to remove their headsets. They were asked to complete the post-trial questionnaire (Appendix III).

8.3.3 Method for the talk-through approach

Participants

Recruitment followed a similar approach to that taken for the VE study (8.3.2), namely recruitment posters and sending emails to undergraduate students. Participants were again told not to apply if they suffered from mental ill-health or if they knew someone who had been involved in a fire. This was a requirement of the ethics committee from previous applications of the approach. Participants were again required to have some familiarity with the Psychology building. A total of 22 participants were recruited for the study (9 male, 13 female; mean age=20.4, SD=1.62; range=18-25).

Materials

This experiment used a questionnaire (Appendix IV) which contained a plan view of the psychology building (Figure 8.13). The questionnaire also contained space for the experimenter to record the hypothetical actions, and locations where they were predicted to occur. Rating scales were included with which the participants rated their expected influences of: familiarity with route, behaviour of other participants, instruction from authority figure, emergency exit signs and distance on their choice of evacuation. The questionnaire also asked participants to estimate the time it would take them to evacuate A5, and to evacuate the building.

Figure 8.13. Plan view shown to participants in the talk-through approach

Procedure

The talk-through approach followed the methodology presented in Chapter 7, in which participants were asked to predict the actions they would take in an emergency situation. In this instance, the scenario was the same as the previous two approaches: evacuation from the A5 computer room of the Psychology building.

Participants were first asked to sign a consent form which summarised the background to the research and the purpose of the current study. It told them to withdraw from the study at any point if they felt distressed or experienced any other unwanted emotions.

The experimenter explained to participants that a scenario was going to be described to them, and they would be asked to state, in order, each action they thought they would take in this situation. They were then presented with the map shown in Figure 8.13. The experimenter highlighted certain features to help orient the participant. These included the main entrance, the foyer, the lecture theatre A1, the corridors, stairwell exit (A) and the A5 computer room.

Participants were asked to imagine that they had come to use the A5 computer room on a weekend morning. They were told that they had not yet logged on, but were completing a paper-based personality questionnaire as part of a job application. They were told that around 20 other people were using the computer room at the same time. Participants were then told that they heard an alarm. They were asked what actions they would take, and where they would conduct them. Participants' hypothetical acts were transcribed by the experimenter.

Once participants' predicted actions led them to evacuate the building, the experimenter read the act list back to them to check for accuracy. Locations were also confirmed.

With this approach the experimenter was unable to judge a period of five minutes after which they would tell the participant to evacuate, as implemented in the other two methods. Therefore, the experimenter only explicitly told participants to evacuate should their predicted sequences lead to a static state, which would not lead to them evacuating without this direction.

Finally, participants were asked to complete the ratings for hypothetical influences on exit choice, and were asked to estimate the time it would have taken them to leave A5 and the building.

8.4 Results

The structure of the results section is shown in Table 8.4 below. The results from the approaches are presented in categories to enable easier comparison. Key findings are highlighted in boxed text within the results sections to emphasise important data.

Table 8.4. Structure of the results section

Result	Section	Summary
Overview	8.4.1	Summary descriptions of the outcome of each approach.
Frequency of acts	8.4.2	Comparison of the frequencies of acts generated through each approach, and analysis against the scientific literature. Comparison of the frequencies of acts obtained from the video footage from the fire drill and from the screen captures from the VE to acts reported in the post-trial questionnaires.
Sequence of acts	8.4.3	Comparison of the sequences of acts. Evaluation against sequence data published in the literature where possible. Comparison of the sequences of acts obtained from the video footage/screen captures to those reported in the post trial questionnaire.
Time to evacuate	8.4.4	Analysis of pre-movement and total evacuation times. Reference to figures published in the scientific literature. Analysis of the objective measures (video footage/screen recordings) against self-reports of evacuation time from the post-trial questionnaires.
Perception of danger	8.4.5	Analysis of the subjective ratings for danger.
Exit choice	8.4.6	Comparison and analysis of route choice (with reference to the scientific literature). Also analysis of influences on route choice.

8.4.1 Overview of the outcome from each approach

This section provides a summary description of the outcome from each approach. Detailed results are provided in the following sections.

Fire drill: after the alarm went off, the majority of the participants remained seated. One participant left the room to investigate, walking back to the main entrance. This participant returned to the computer room. Two different participants left the building via the main entrance, followed shortly afterwards by the participant who had originally investigated and one other participant. These four stayed outside the building, while the other participants remained seated in A5 until the experimenter told them to leave after five minutes. The participants all left via the route they had used to enter the building: the main entrance (B).

Virtual environment[a]: The results of the virtual environment were more variable than those in the fire drill. Nine participants evacuated before being explicitly instructed to by the experimenter. Five participants re-entered the building after their initial evacuation. Five participants left the building through the corridor stairwell exit (A); nine left through the main exit (B); one participant went to the foyer, then into the basement before exiting to the rear of the building. In general, the virtual environment demonstrated a more varied response than that witnessed in the fire drill.

Talk-through approach: The talk-through approach also elicited behavioural predictions which demonstrated greater variability than those of the fire drill. This was particularly evident during the early stages of the (hypothetical) emergency. Of the 22 participants, 19 predicted they would exit via the main entrance (B); 3 would exit by the closer corridor stairwell (A). Two participants predicted a return to the computer room having left to gain information.

8.4.2 Frequency of acts

The frequencies of acts (shown as a percentage of the total number of acts for each study) are shown in Table 8.5. The acts from the fire drill were taken from the video footage; the VE acts were from the screen grabs of the avatars in the VE; and the talk-through acts were participants' hypothetical behaviours in response to the described scenario. The act frequencies are shown in descending order of combined act frequency across the three studies. They are shown as percentages of the total number of acts per study to enable cross-method comparison.

[a] Data related to four participants were removed from the analysis. This was because their evacuation choices may have been influenced by a discrepancy in the virtual environment. This issue is expanded upon in the discussion.

Table 8.5. Frequency of acts emerging from the standardised comparison of approaches, shown in descending order of combined act frequency across the three studies

		Frequency (% of total no. of acts in each study)		
Code	**Action category**	**Fire drill**	**VE**	**Talk-through**
10a	Evasive	27.45	19.61	16.19
6a	Seek information and investigate	9.02[a]	11.44	6.48
5b	Disregard/ ignore stimulus/ stay seated/ continue prior activity	9.02	7.84	8.91
24a	Leave immediate area	9.41	6.21	9.72
23a	Enter area of minimum risk	9.02	5.88	8.91
1a	Pre-event actions	8.63	4.90	8.91
30a	End of involvement	8.63	4.90	8.91
29a	Follow others/move with others/copy others	8.24	4.90	1.21
2b	Note ambiguous behaviour of others	0	7.84	5.26
14b	Receive instructions	7.06	3.59	0.40[b]
18b	Note unambiguous behaviour of others	0	2.94	5.26[c]
34a	Problem with VE	0	7.52	0
9a	Dress/gather valuables	0	0	6.48
12a	Securing environment/return to A5	1.18	2.29	1.21
27a	Discuss/debate/ask	1.18	0.98	2.43
35a	Get lost	0	3.92	0
17a	Note persistence of stimulus	0	0	2.43[c]
38a	Re-enter building	0.39	1.31	0
4b	Arrive at conclusion	0	0	1.62
39a	Overshoot	0	1.31	0
16c	Experience negative feelings	0	0	1.21
18a	Receive information (verbal)	0	0	1.21[b]
32a	Look for/at fire signs	0	0	1.21
33a	Wait (not in A5 computer room)	0.39	0.33	0.40
14a	Give instructions	0	0.65	0.40
19a	Alter plan (self-initiated)	0.39	0.65	0
26a	Encounter colleague/superior	0	0.65	0
2a	Perception of stimulus (ambiguous)	0	0	0.40
7a	Disseminate warnings/information	0	0	0.40
22a	Note nothing unusual/stay calm	0	0	0.40
15a	Give assistance	0	0.33	0
	Total number of acts	**N=255**	**N=306**	**N=247**

[a] Note that for the fire drill study, the "seek information" act most commonly referred to looking at an alarm board located in the foyer on their way past. In Canter et al. (1980) this act was reported at an earlier stage of the evacuation sequence. It was an unusual feature of the fire drill scenario that the alarm information was available at this late stage in the evacuation.

[b] For the talk through study, these acts were prompts by the experimenter at an appropriate stage in the act sequences

[c] For the talk through study, these acts were comprised of both prompts by the experimenter and unprompted reports by participants (Act 17a: 3 prompts by experimenter/3 from participant; Act 18b: 10 prompts by experimenter/3 from participant).

The acts were coded in accordance with the approach presented in Section 7.4. The taxonomy from multiple occupancy and hospital fires in Canter et al. (1980) was used as a basis for the coding, being the closest available behavioural taxonomy. Acts 27a-

39a were not in this reference taxonomy, but were demonstrated during one or more of the studies in Phase 2 of this research, and therefore additional act categories were created. Any acts from the Canter et al. (1980) taxonomy which were not witnessed in any of the Phase 2 studies were removed from the analysis. The fire drill and VE were populated with acts coded from the video footage and the screen capture respectively; the data from the participants' reports of the events in the post-trial questionnaire are presented later in this section.

The Shapiro-Wilk test was run to check for normality. The results showed that the data for each condition were not normal (Table 8.6). Therefore, the concurrent validity was investigated using non parametric correlations using Spearman's rho. Significant correlations were found between each condition as shown in Table 8.7. The correlation coefficients indicate medium (VE and talk-through) or large (fire drill and VE/fire drill and talk-through) effect sizes, according to the values from Cohen (1988).

Table 8.6. Shapiro-Wilk test for normality: frequencies of acts

	Shapiro-Wilk	df	p
Fire drill	0.593	31	<0.001
VE	0.751	31	<0.001
Talk-through	0.769	31	<0.001

Table 8.7. Spearman's rho correlation coefficient for frequency of acts

	r_s	N	p
Fire drill and VE	0.689	31	<0.001
Fire drill and talk-through	0.575	31	<0.01
VE and talk-through	0.394	31	<0.05

In addition to the frequencies of acts captured on video (fire drill) and screen recordings (VE) the actions participants reported completing in the post-trial questionnaires for these studies were coded separately against the taxonomy shown in Table 8.5. Initially, the frequency of act data for the fire drill, obtained from the video footage and the post-trial questionnaires were investigated, to determine the concurrent validity of these methods. Any act in the taxonomy which was not identified in either the video or the questionnaire data was removed. Then, the data were tested for normality (Table 8.8). The data were not normal and therefore a Spearman's rho was conducted. A high correlation was found between the frequency of acts captured from the video footage and the post-trial questionnaire (Table 8.9).

Table 8.8. Shapiro-Wilk test for normality: fire drill frequencies of acts taken from the video footage and those from the post-trial questionnaires

	Shapiro-Wilk	df	p
Fire drill (video)	0.611	29	<0.001
Fire drill (self-report)	0.819	29	<0.001

Table 8.9. Spearman's rho correlation for the fire drill frequencies of acts from the video footage and those from the post-trial questionnaires

Fire drill: frequency of acts	r_s	N	p
Video of actions and self-reported actions	0.527	29	<0.01

The process above was repeated to investigate the relationship between the frequencies of acts from the screen capture of the VE studies to those reported in the post-trial questionnaire. Shapiro-Wilk tests demonstrated non-normality in the data (Table 8.10), and therefore a comparison of the acts was made using Spearman’s rho. A significant finding, and large effect size, was found (Table 8.11).

Table 8.10. Shapiro-Wilk test for normality: VE frequencies of acts taken from the screen grabs and those self-reported in the post-trial questionnaires

	Shapiro-Wilk	df	p
VE (screen grabs)	0.769	29	<0.01
VE (self-report)	0.859	29	<0.001

Table 8.11. Spearman's rho correlation between the VE frequency of acts taken from the screen recordings and those self-reported in the post-trial questionnaires

VE: frequency of acts	r_s	N	p
Screen recordings and self-reported acts	0.658	29	<0.001

While the correlations in the preceding section demonstrated significant relationships for the act frequencies obtained from each of the approaches, it was also pertinent to review the data against the published literature on human behaviour in fire (Table 8.12). This was to gain an understanding of the predictive validity of the act frequencies. The Canter et al. (1980) study provided no sufficiently similar scenario to conduct a quantitative analysis; therefore the analysis was based on a qualitative review of the observed acts against the literature. The frequencies shown for the fire drill and VE were based on the video footage and screen capture data respectively, although reference is made to the data from the post-trial questionnaires. Act frequencies observed in the standardised comparison which were not supported by the literature are highlighted. Acts with very low frequencies (<5% in each study) were omitted from the table, unless their occurrence (or absence) contributed to an understanding of the predictive validity of the behaviours.

Table 8.12. Review of observed acts against the literature. Colour coding indicates concerns with frequencies of the reported acts (yellow) or act frequencies which are not supported by the literature (red).

Action category	Frequency (% of acts in each study) Fire drill	VE	Talk-through	Literature
Evasive	27.45	19.61	16.19	Act refers to a movement through the building towards the exit, reported in almost all building evacuations (e.g. Olsson and Regan, 2001; Gwynne et al., 2003).
Seek information and investigate	9.02	11.44	6.48	Witnessed in domestic, multiple occupancy and hospital fires in Canter et al. (1980); was also expected for the Psychology evacuation. Act has also been reported in high rise building evacuations (Proulx and Reid, 2006; McConnell et al., 2009).
Disregard/ignore stimulus/stay seated/continue prior activity	9.02	7.84	8.91	Several sources indicate a tendency to disregard alarms or to continue activities prior to leaving (Pauls and Jones, 1980a; Proulx, 1995; Purser and Bensilum, 2001; Gwynne et al., 2003; Gershon et al., 2007; Proulx, 2007; Galea et al., 2009; Xudong et al., 2009).
Leave immediate area	9.41	6.21	9.72	Act refers to leaving the initial location. Reported in several studies (Olsson and Regan, 2001; Gwynne et al., 2003; Gershon et al., 2007).
Enter area of minimum risk	9.02	5.88	8.91	Referred to leaving the building. Reported in several studies (Canter et al., 1980; Olsson and Regan 2001; Gwynne et al., 2003).
Pre-event actions	8.63	4.90	8.91	*Included to support generation of act sequences*
End of involvement	8.63	4.90	8.91	*Included to support generation of act sequences*
Follow others/move with others/copy others	8.24	4.90	1.21	Groups and social factors have been recognised as influential on the outcome of an evacuation (Aguirre et al., 1998; Drury et al., 2006; Pan et al., 2006; Gershon et al., 2007; Kuligowski, 2009)
Note ambiguous behaviour of others	**0.00**	7.84	5.26	See "*Follow others/move with others/copy others*". In the fire drill, the video analysis alone did not identify this level of granularity. However, it was reported in participants' accounts of their actions in the post-trial questionnaire.

Action category	Frequency (% of acts in each study) Fire drill	VE	Talk-through	Literature
Receive instructions	7.06	3.59	0.40	This act was a feature of the experimental design, although it has been noted in several evacuation studies in the literature (Edelman et al., 1980; Kanno et al., 2006; Gershon et al., 2007; Proulx, 2007)
Note unambiguous behaviour of others	**0.00**	2.94	5.26	See "*Follow others/move with others/copy others*". This act was only recorded in the post-trial questionnaire for the fire drill.
Problem with VE	0.00	**7.52**	0.00	Mantovani et al. (2001) reported problems with their VE. However, an important finding was that this act would not occur in the scenario for which the prediction is being made i.e. a real building evacuation.
Dress/gather valuables	**0.00**	**0.00**	6.48	Noted in several evacuation studies (Canter et al., 1980; Proulx, 1995; Proulx, 2001; Purser and Bensilum, 2001; Proulx and Reid, 2006). Was not identified through video analysis of fire drill, although was reported in the post-trial questionnaire. Was not possible in this VE study.
Securing environment/ return to A5	1.18	2.29	1.21	Some evidence in the literature supported the occurrence of this act (Canter et al., 1980; Edelman et al., 1980; Pan et al., 2006).
Discuss/debate/ask	1.18	0.98	2.43	Noted in WTC studies (Proulx and Reid, 2006; Gershon et al., 2007; Galea et al., 2009).
Get lost	0.00	**3.92**	0.00	Some studies have cited the benefits of signage design (or negative effects of poor design)(Pan et al., 2006; Gershon et al., 2007; Kobes et al., 2009). Unfamiliarity with building has also been cited as a hindrance to evacuation (Gershon et al., 2007; Gershon, 2009). However, no evidence was given in the literature for being lost.
Overshoot	0.00	**1.31**	0.00	Difficulties in controlling avatars have been reported (Mantovani et al., 2001; Gamberini et al., 2003; Mol et al., 2008). However, this specific action was not reported for any real-life evacuation scenarios.

Action category	Frequency (% of acts in each study) Fire drill	VE	Talk-through	Literature
Look for/at fire signs	0.00	0.00	1.21	Some studies have cited the benefits of signage design in evacuation, or negative effects of poor design, indicating their use (Pan et al., 2006; Gershon et al., 2007; Kobes et al., 2009).
Perception of stimulus (ambiguous)	0.00	0.00	0.40	Proulx (2007) reported failure to recognise the alarm as a reason for not evacuating. Not determined through the video analysis of the fire drill, but was reported in the post-trial questionnaire. Not reported at all for the VE study. This may have been because the alarm noise used in this study was more recognisable than in the real building.
Note nothing unusual/stay calm/ no danger	0.00	0.00	0.40	Rational behaviour generally dominates in emergencies (Pauls and Jones, 1980b; Sime, 1995; Proulx, 2001). This behaviour was not possible to obtain from the video analysis or VE screen capture, but was captured in the perception of danger ratings in the post-trial questionnaires.

Frequency of acts: key findings

- Medium or large effect sizes (Cohen, 1988) were obtained for correlations of the frequencies of acts between each pair of approaches
- Large correlation coefficients (Cohen, 1988) were seen between the frequencies of acts recorded (from the video footage in the fire drill and screen recordings in the VE) and those reported in the post-trial questionnaires.
- The obtained acts generally showed comparability to those reported in the scientific literature, except for a small number of acts specific to VEs

8.4.3 Sequence of acts

As in previous investigations (Section 4.4 and Chapter 7) standard residuals were used to investigate transitions between acts. These indicate the extent to which occurrence of an act increases the likelihood of the proceeding one. The standard residuals were calculated using transition matrices, which tallied all transitions between the act categories shown in Table 8.5. Act categories were omitted if no appropriate actions were recorded during the particular study. For example, act 14a “give instructions” was not recorded during the fire drill; therefore this act was not included in the fire drill transition matrix.

The standard residuals were calculated for each transition matrix. Tests revealed that they were not normally distributed (Table 8.13).

Table 8.13. Shapiro-Wilk test for normality of standardised residuals

	Shapiro-Wilk	df	p
Fire drill	0.553	169	<0.001
VE	0.610	441	<0.001
Talk-through	0.574	529	<0.001

For each pair of approaches, the standard residuals for all transitions which occurred in both conditions were tested for correlations to investigate their concurrent validity. The results of the correlation test are shown in Table 8.14. Each one was significant, and demonstrated medium or large effect sizes, according to the classification given by Cohen (1988).

Table 8.14. Spearman's rho correlation coefficient for comparable sequence data

	r_s	N	p
Fire drill and VE	0.429	169	<0.001
Fire drill and talk-through	0.534	121	<0.001
VE and talk-through	0.492	196	<0.001

Thereafter, decomposition diagrams were generated for the behavioural sequences generated through each approach (Figure 8.14-Figure 8.16). The diagrams show the transitions between acts for which the standardised residuals were greater than two. These can be considered as the transitions which occurred most often relative to their expected occurrence; they indicate the extent to which a transition is more likely than expected by chance alone (Bakeman and Gottman, 1986). The act categories are shown within the nodes. The standardised residuals are shown on the arrows leading between the nodes, which point to subsequent acts. To clarify the decomposition diagrams, negative standardised residuals were omitted from the diagrams. These represented transitions which were unlikely, and so contributed less to an understanding of the sequences of behaviour than those which were more likely than expected. Also, transitions are only shown for acts which accounted for 2% or more of the total frequency of acts per study. This avoided cluttering the diagram with low frequency acts. The threshold was decided upon as the lowest value which enabled complete sequences (from pre-event actions to end of involvement) to be generated for each of the approaches. A high-level analysis of the decomposition diagrams is shown in Figure 8.17, which focuses on the consistency of the behavioural sequences.

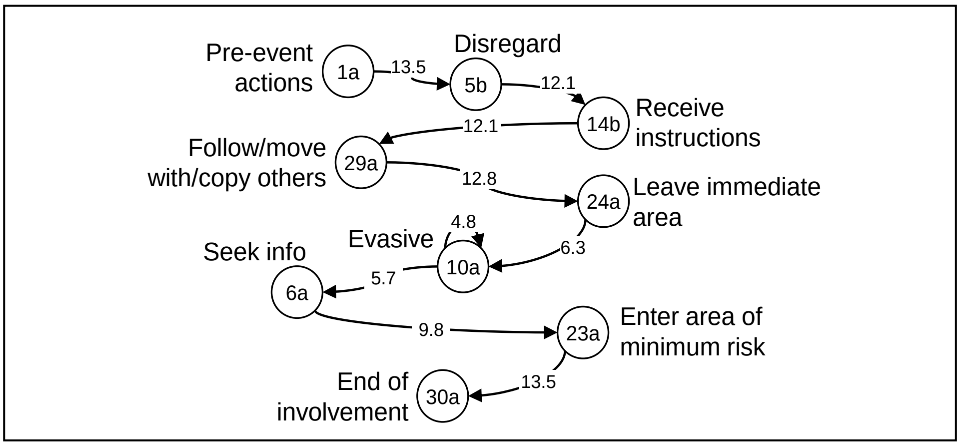

Figure 8.14. Decomposition diagram: fire drill

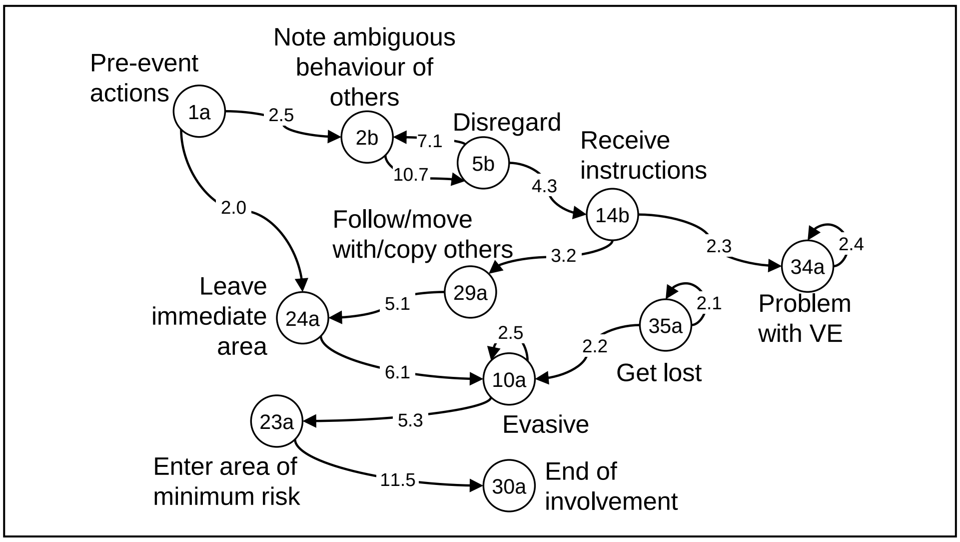

Figure 8.15. Decomposition diagram: VE

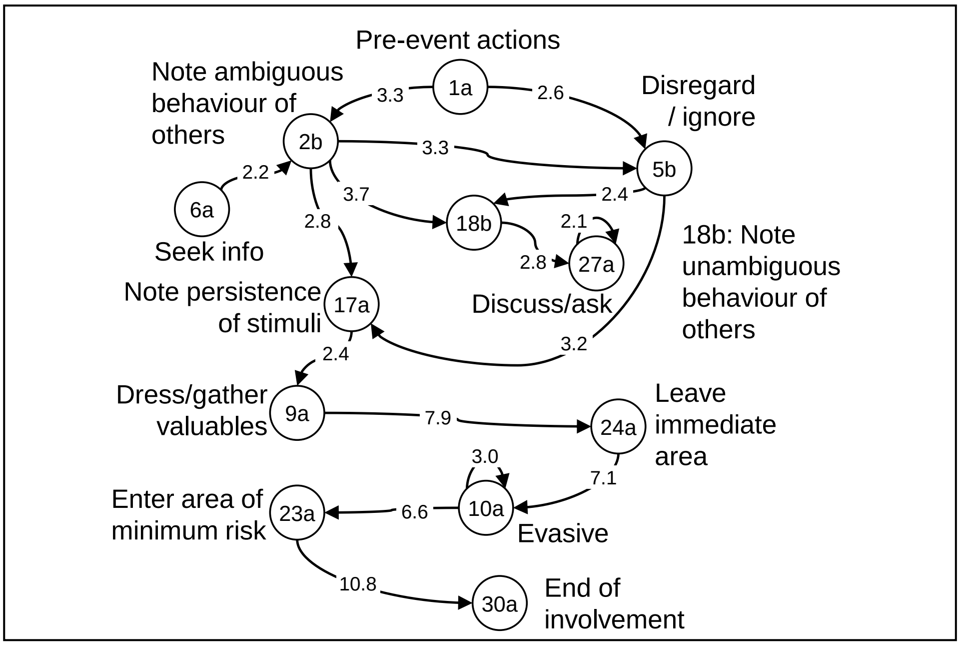

Figure 8.16. Decomposition diagram: talk-through

In general, the VE and talk-through diagrams (Figure 8.15 and Figure 8.16) demonstrate a more varied sequence of behaviours than in the fire drill (Figure 8.14). This can also be seen in Figure 8.17, which shows fewer recurrent loops and divergent sequences for the fire drill. The fire drill was also had a more consistent sequence of transitions, with only one start point, and one end point.

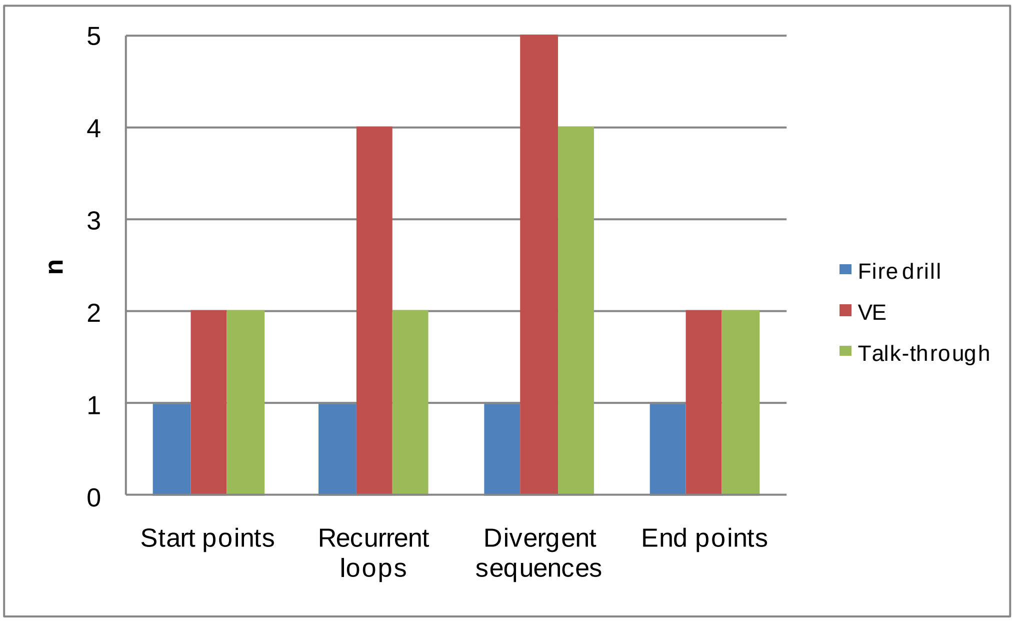

Figure 8.17. Analysis of the decomposition diagrams

This consistency in the results of the fire drill was a consequence of the common behavioural pattern which many people demonstrated in the drill: staying in the computer room for the full five minutes, then receiving instructions to leave and evacuating as a group. As reported in Table 8.12, group behaviour such as this has been previously reported in the scientific literature (Aguirre et al., 1998; Drury et al., 2006; Pan et al., 2006; Gershon et al., 2007; Kuligowski, 2009).

The VE sequence (Figure 8.15) demonstrates a more complex behavioural pattern from the first instance, with participants having similar likelihood for “note/seek behaviour of others” or “leave immediate area”. There is also a return loop of actions between “note/seek behaviour of others” and “disregard” as people in A5 continually checked what other occupants were doing. This diagram has sequences associated with “problem with VE” and “get lost”.

The talk-through decomposition diagram (Figure 8.16) demonstrates a variety of likely sequences through the early stages of the hypothetical emergency. These involved noting the behaviour of others, disregarding the cues, discussion and seeking information. However, after the persistence of the stimuli (alarm), participants predicted dressing/gathering belongings, then leaving the computer room and then walking out of the building.

As with the frequencies of acts, the concurrent validity of the sequences of acts obtained through analysis of the post-trial questionnaires were investigated against those obtained from the video footage (fire drill) and screen capture (VE). For the fire drill, act sequences which were demonstrated in both the video footage and the questionnaire data were analysed. The standardised residuals were tested using Shapiro-Wilk and found not to be normally distributed (Table 8.15). A Spearman's rho test demonstrated a medium correlation between the sequences of acts from the video footage and from participants' accounts of the event in the post-trial questionnaires (Table 8.16).

Table 8.15. Shapiro-Wilk test for normality: fire drill sequences of acts taken from the video footage and self-reported in the post-trial questionnaires

	Shapiro-Wilk	df	p
Fire drill (video)	0.568	100	<0.001
Fire drill (self-report)	0.659	100	<0.001

Table 8.16. Spearman's rho correlation between the fire drill sequences of acts taken from the video footage and those self-reported in the post-trial questionnaire

Fire drill: sequences of acts	r_s	N	p
Video of actions and self-reported acts	0.429	100	<0.001

The process described above was repeated for the sequences of acts from the VE screen grab data and the VE post-trial questionnaire accounts of the event. The data were found not to be normal (Table 8.17). Spearman's rho correlation demonstrated a medium effect size between the screen grab and questionnaire data (Table 8.18).

Table 8.17. Shapiro-Wilk test for normality: VE sequences of acts taken from the screen recordings and self-reported in the post-trial questionnaires

	Shapiro-Wilk	N	p
VE (screen grabs)	0.639	289	<0.001
VE (self-report)	0.625	289	<0.001

Table 8.18. Spearman's rho correlation between the VE sequence of acts taken from the screen recordings and those self-reported in the post-trial questionnaires

VE: sequence of acts	r_s	N	p
Acts from screen grabs and self-reported acts	0.449	324	<0.001

As for the frequencies of acts, the sequences of acts were also compared to the scientific literature to investigate the predictive validity. Unfortunately the scientific literature provided little detailed data on sequential aspects of evacuation. The few relevant findings from the literature are presented below. Evidence of these behaviours from the data described above is provided.

Table 8.19. Comparison of findings from the scientific literature to the sequence data from the standardised comparison

Behavioural sequences info from literature	Fire drill	VE	Talk-through
Occupants have often conducted non-evacuation activities before evacuating, including gathering belongings, investigating, or warning (Proulx, 1995, 2001; Purser and Bensilum, 2001; Proulx and Reid, 2006; Galea et al., 2009)	The majority of participants disregarded the alarm, and continued prior activities. However, the detail of the acts they undertook were only captured using the post-trial questionnaire; the detail was not captured by camera.	Participants were particularly likely to note/seek the behaviour and disregard the alarm, continuing pre-event actions prior to evacuating. Other pre-evacuation activities were limited; there was no opportunity to dress or gather valuables in this study.	The predictions gave some evidence of conducting (hypothetical) non-evacuation activities prior to evacuating. These included completing tasks and gathering belongings.
People have demonstrated commitment to prior activities after hearing an alarm (Purser and Bensilum, 2001)	Participants did continue with the personality questionnaire issued to them prior to the alarm. Again, this had to be captured in the post-trial questionnaire – the video did not gather this level of detail.	Several participants continued to read the posters before evacuating.	Several participants predicted that they would continue with prior activities.
People may follow others who they see evacuating (Tubbs and Meacham, 2009)	Some evidence. Of the four participants who evacuated prior to instruction, both evacuated in pairs. However, the remainder of the participants stayed seated until instructed to leave. They then all evacuated together.	Some participants copied others who were evacuating. Generally, participants evacuated together once being told to leave the building.	Some participants predicted that they would follow/move with others, but this was not reflected in Figure 8.16.
If a group is split up, a member may return to the building to re-form the group before exiting as a whole (Pan et al., 2006).	The first participant who left the building returned before evacuating with someone else.	Some participants re-entered the building; this may have been to re-form the group.	There was no evidence of a (hypothetical) return to the building. However, there were some indicators of group/social behaviour in Figure 8.16.

Behavioural sequences info from literature	**Fire drill**	**VE**	**Talk-through**
After initial cues, occupants follow an investigate or misinterpret sequence (Canter et al., 1980).	Misinterpret (disregard) is notable in Figure 8.14. The investigate sequence is not shown, but was evident by the participant who tried to find info before returning to A5.	Disregard is also evident in Figure 8.15. The seek (investigate) sequence is not apparent.	Investigation (seek info) and misinterpret (disregard) sequences are evident in Figure 8.16. There was no notable sequence leading to “seek info”, however.
After receiving initial cues, behavioural sequences progress through the generic stages of *Interpret*, *Prepare*, and *Act* (Canter et al., 1980).	The interpret stage is evident through the disregard behaviour in Figure 8.14. The prepare act was largely negated by being told to evacuate, and so participants skipped this stage moving directly to *Act* (evacuate).	Figure 8.15 shows some evidence of the *Interpret* stage at the anticipated sequence, through noting the behaviours of others and the disregard loop. There is little evidence for *prepare*. After receiving instruction, participants evacuated, as evidence of the *Act* phase.	Several acts (e.g. disregard/ discuss) in Figure 8.16 give evidence for *Interpret. Prepare* is evidenced through Dress/gather valuables. All subsequence acts are associated with the *Act* phase (evacuation).

Sequences of acts: key findings

- Medium or large correlation coefficients were seen for the sequences of acts obtained from each approach
- The VE and talk-through demonstrated a more varied sequence of behaviours than in the fire drill
- Medium effect sizes were seen for the sequences of acts recorded (from the video footage in the fire drill; from the screen recording in the VE) and those reported by participants in the post trial questionnaires
- Some comparability was demonstrated with the sequences of acts seen in the standardised comparison and those reported in the literature; the available data in the literature was limited.

8.4.4 Time to evacuate

Pre-evacuation times

A comparison was also made of the times taken to evacuate both from the computer room and from the building. The time taken to leave the computer room can be approximated to the pre-evacuation time. The mean times, and time for the first and last persons to leave the computer room, are shown in Figure 8.18. These were taken from the video footage for the fire drill, the screen recordings for the VE, and for the estimated time to leave for the talk-through. Only estimated times were available from the talk-through approach, but these were used in the analysis to determine the accuracy of the participants' predictions. For all approaches, the times were recorded each time a participant left the computer room, if they returned and left again. This occurred in the fire drill (n=2), VE (n=3) and talk-through (n=3). The first time to leave was analysed as this was the beginning of the evacuation process, which was the figure of most interest. The full descriptive statistics are shown in Table 8.20.

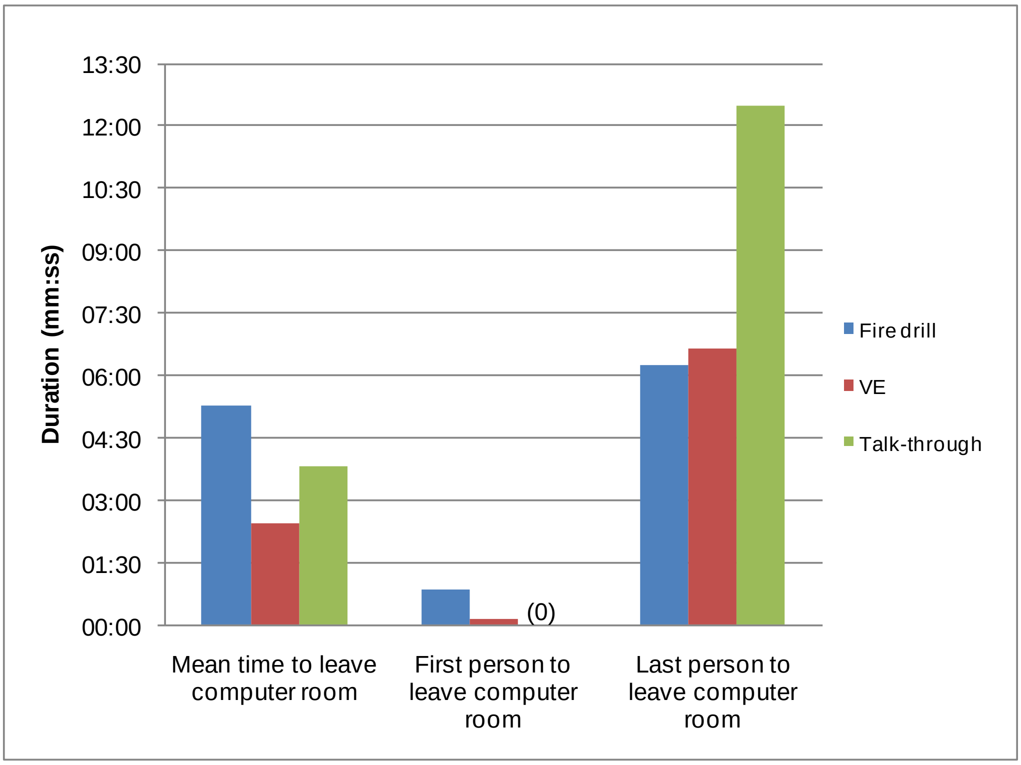

Figure 8.18. Time taken for participants to leave the computer room (for the first time, if more than once)[a] . Fire drill values were taken from the video footage, VE values from the screen recordings, and the talk-through values were participants' estimated times.

[a] Note that participants were not told to leave the computer room after five minutes in the talk-through approach, therefore these results should be treated with caution

Table 8.20. Descriptive statistics for time to leave computer room (mm:ss) after alarm started

	mean	SD	Min	Max
Fire drill	05:16	01:30	00:50	06:14
VE	02:26	02:41	00:08	06:39
Talk-through	03:48	03:11	00:00	12:30

Shapiro-Wilk tests revealed that the times for leaving the computer room were not normal (Table 8.21). Therefore, a Kruskal-Wallis test was run to investigate differences in the time to leave the computer room. A significant difference was found in the times to leave the computer room (x^2=17.505, df=2, p<0.001). Stepwise comparisons using Mann-Whitney U test revealed significant differences in each comparison, except VE and talk-through (Table 8.22). All post hoc tests in this study were corrected using Bonferroni. This strategy was adopted to minimise type I errors, but the increased risk of type II errors must be acknowledged (Perneger, 1998).

Table 8.21. Tests of normality for time to leave the computer room

	Shapiro-Wilk	df	p
Fire drill	0.582	22	<0.001
VE	0.724	15	<0.001
Talk-through	0.875	22	<0.05

Table 8.22. Mann-Whitney U comparisons

	Mann-Whitney U	p (0.017)
Fire drill – VE	42.00	<0.017
Fire drill – Talk-through	109.00	<0.017
VE – Talk-through	124.50	NS

Because participants could not be told to evacuate after five minutes in the talk-through approach, as was possible in the fire drill and the virtual environment, a comparison was made of the number of participants evacuating prior to being told to do so (values were taken from the video footage for the fire drill, screen recordings for the VE and estimated time to leave for the talk-through). Instruction was given to evacuate after approximately five minutes, therefore the percentage of participants leaving in five minutes or less is shown in Figure 8.19.

A chi-square test was run and showed significance (X^2=21.687; df=2; p<0.05). Therefore, the predictive approach did have an effect on the number of people leaving before five minutes.

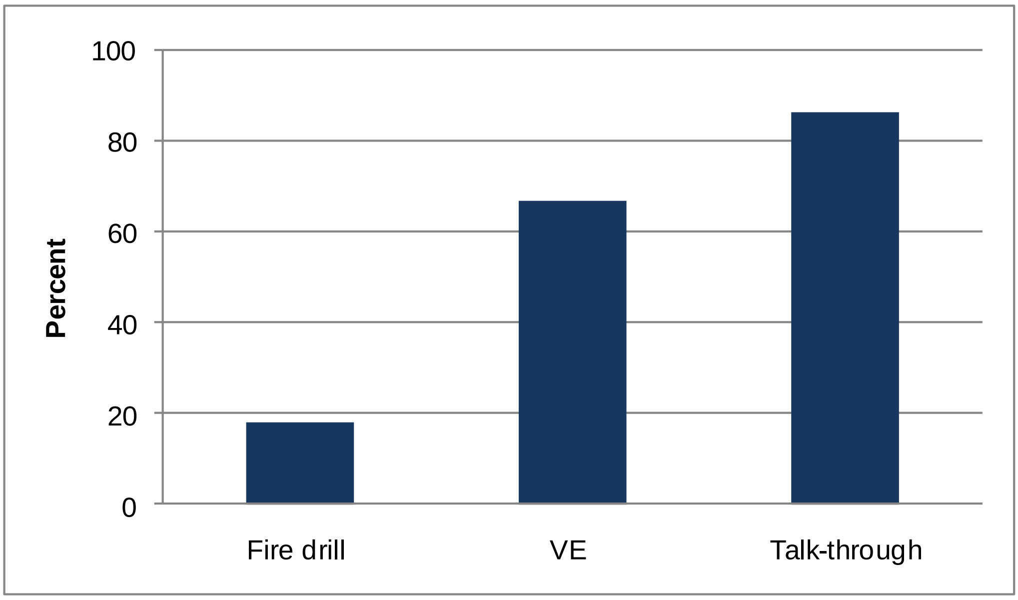

Figure 8.19. Percentage of volunteers who left the computer room in five minutes or less.

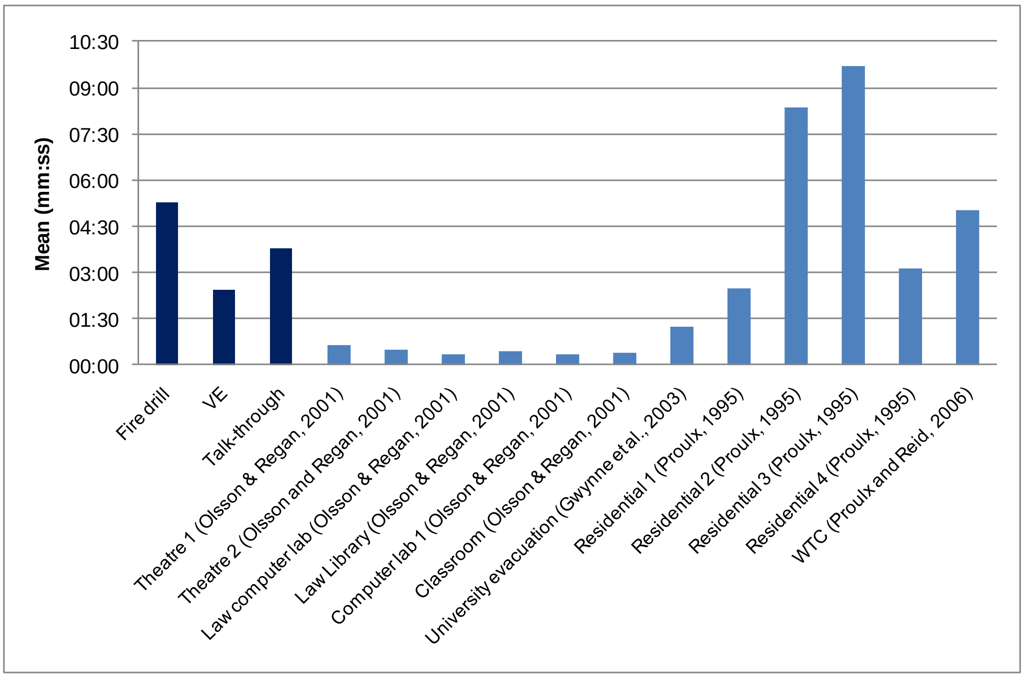

Figure 8.20. Mean pre-evacuation times from Study 4 and the literature. Fire drill values were taken from the video footage, VE values from the screen recordings, and the talk-through values were participants' estimated times.

The mean times taken to leave the computer room are shown on a bar chart with results reported in the scientific literature in Figure 8.20 to investigate the predictive validity of the approaches. While the study by Olsson and Regan (2001) shows very low mean pre-evacuation times, the ranges found in other studies (Proulx, 1995; Proulx and Reid, 2006) suggested that those in the standardised comparison were representative. Note that for the fire drill and VE studies in the standardised comparison participants were told to evacuate after five minutes, and therefore the pre-evacuation times may have been longer if participants had not been told to leave. This would likely have led to longer pre-evacuation times in these studies.

In addition to the times taken to evacuate from the video footage in the fire drill and screen recordings in the VE, participants were also asked to estimate the time it took them to leave the computer room in the post-trial questionnaires. The mean values for the self-report measures of time taken to leave the computer room are are shown in Figure 8.21 against those taken from the video footage and screen recordings; the descriptive statistics are shown in Table 8.23 (the measures from the recordings were also shown previously in Table 8.20: they are repeated here to enable easier comparison).

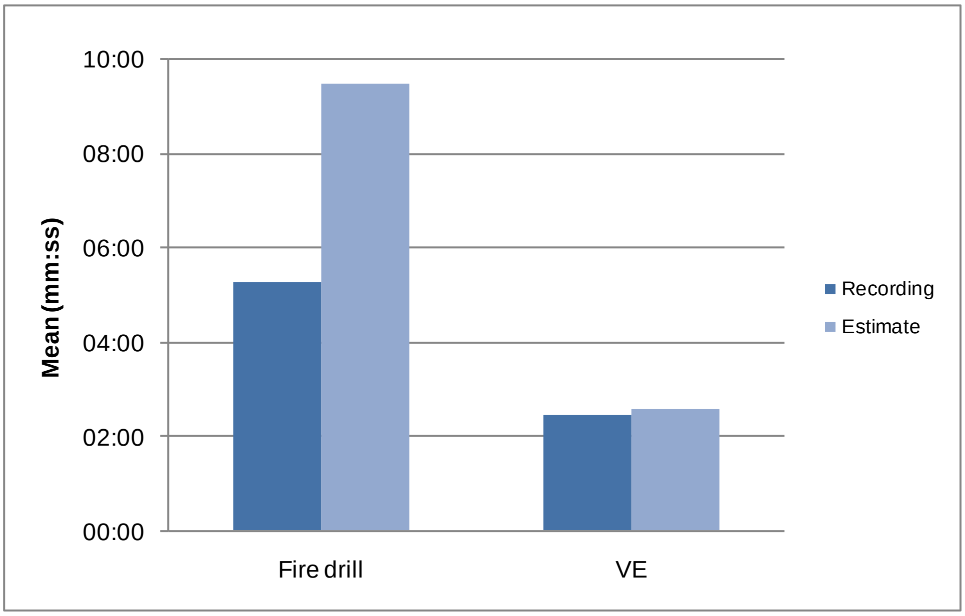

Figure 8.21. Mean times taken to leave A5, showing values taken from the video footage (fire drill) and screen recordings (VE) against those taken from the participants' estimates in the post-trial questionnaire

Table 8.23. Descriptive statistics for time to leave computer room (mm:ss) after alarm started: values taken from video footage recordings (fire drill) and screen recordings (VE) and from participants' estimates for both

	mean	SD	Min	Max	N
Fire drill (from video footage)	05:16	01:30	00:50	06:14	22
Fire drill (estimates)	09:30	03:49	03:00	15:00	18
VE (from screen recordings)	02:26	02:41	00:08	06:39	15
VE (estimates)	02:36	03:29	00:02	10:00	14

As the video footage and screen recordings were previously found not to demonstrate normality (Table 8.21), Spearman's correlation coefficient was used to investigate the relationship between participants' self-reported times and the objective measures (Table 8.24).

Table 8.24. Spearman's correlation for subjective and objective measures for time taken to leave the computer room for the fire drill and VE

	r_s	N	p
Fire drill: estimates/video footage	0.268	18	NS
VE: estimates/screen recordings	0.962	14	<0.001

There was no significant relationship between the self-reported times and those captured on video for the fire drill, indicating that participants were not able to accurately report the time it took to evacuate. The VE screen recordings and self-reports of time taken to leave the computer room however demonstrated a large effect size, and significant relationship. Thus, a stronger association was seen between the objective and subjective measures in the VE. The results should be treated with caution however, as the question did not explicitly ask participants to state the time taken to leave the computer room *after the alarm started* and therefore some possibility existed for participants to misinterpret the question.

Time taken to leave building

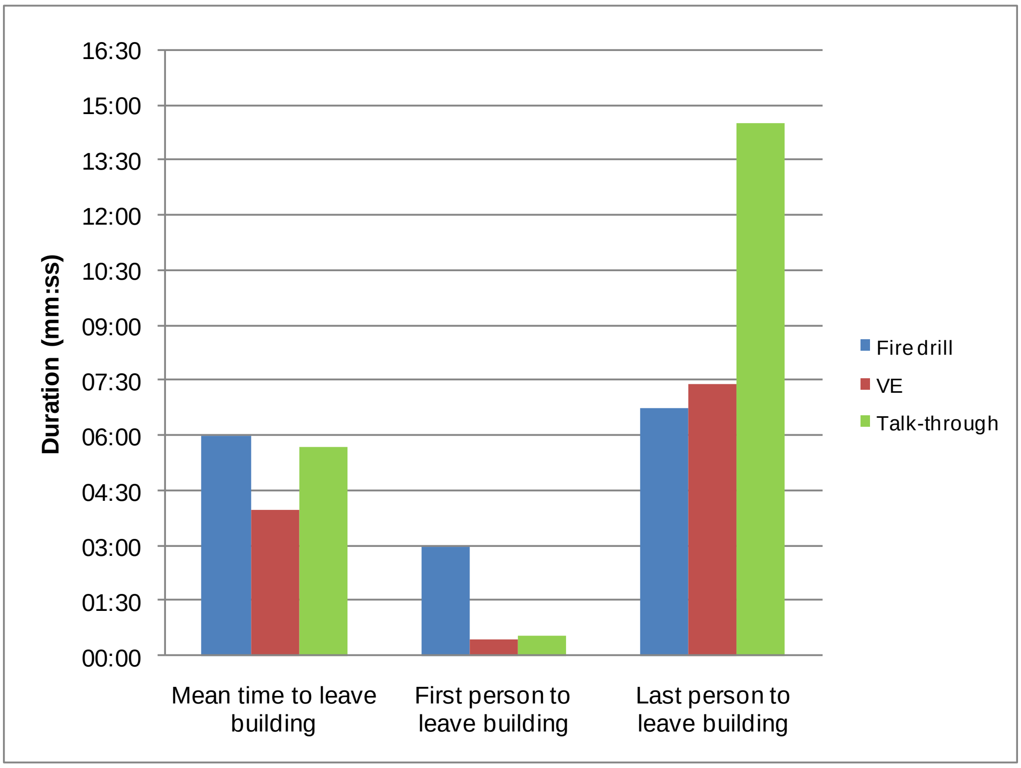

Figure 8.22. Times taken to leave the building[a]. Fire drill values were taken from the video footage, VE values from the screen recordings, and the talk-through values were participants' estimated times.

[a] Note that participants were not told to leave the computer room after five minutes in the talk through approach, and that this is likely to have affected the times taken to leave the building

The times taken to evacuate the building are shown in Figure 8.22, with values taken from the video footage for the fire drill, the screen recordings for the VE, and a hypothetical prediction of time taken to leave the building for the talk-through approach. In contrast to the time taken to leave the A5 computer room, the *final* time to evacuate the building was used if the participant left the building more than once (i.e. exited, returned and left again). This was because the final evacuation was considered the end of the evacuation process, and therefore the most important figure. This sequence of actions happened in the fire drill (n=1) and in the virtual environment (n=3[a]). The final time to evacuate the building has been commonly reported in the literature, presented as either mean (based on number of evacuees) or maximum (last person to leave) values (Proulx, 1995; Shields and Boyce, 2000; Olsson and Regan, 2001; Averill et al., 2009; Xudong et al., 2009). The descriptive statistics for the time taken to leave the building are shown in Table 8.25.

Table 8.25. Descriptive statistics for time to leave the building (mm:ss) after alarm started

	mean	SD	Min	Max
Fire drill	06:00	01:07	02:57	06:45
VE	03:59	02:26	00:27	07:23
Talk-through	05:40	03:22	00:33	14:30

Shapiro-Wilk was run for the time taken to evacuate the building. The fire drill data were found not to be normally distributed (Table 8.26). Therefore, the Kruskal-Wallis test was run. This was found to be significant (x^2=9.808, df=2, p<0.01). Mann-Whitney U comparisons (with Bonferroni correction) were made of each pair of conditions (Table 8.27). The only significant difference was between the fire drill and the virtual environment.

Table 8.26. Normality tests for time to leave building

	Shapiro-Wilk	N	p
Fire drill	0.590	22	<0.001
VE	0.911	14[b]	NS
Talk-through	0.936	22	NS

Table 8.27. Mann-Whitney U comparisons

	Mann-Whitney U	p (0.017)
Fire drill – VE	64.50	<0.017
Fire drill – Talk-through	154.00	NS
VE – Talk-through	110.00	NS

[a] A further participant in the VE study left the building then re-entered, but their computer crashed, effectively ending the trial with the participant inside the building.

[b] Note one participant's Second Life froze and therefore they could not exit the building in the VE

Evacuation time may always be influenced by the size and layout of a building, due in part to the time it takes to travel the distance from the point of origin to the exit. However, comparing the results from the Psychology building evacuation to others reported in the literature gave a crude estimate of the predictive validity of the approaches. This comparison can be seen in Figure 8.23; the data generated in the Psychology evacuation studies fell within the range of those found from a comparable investigation.

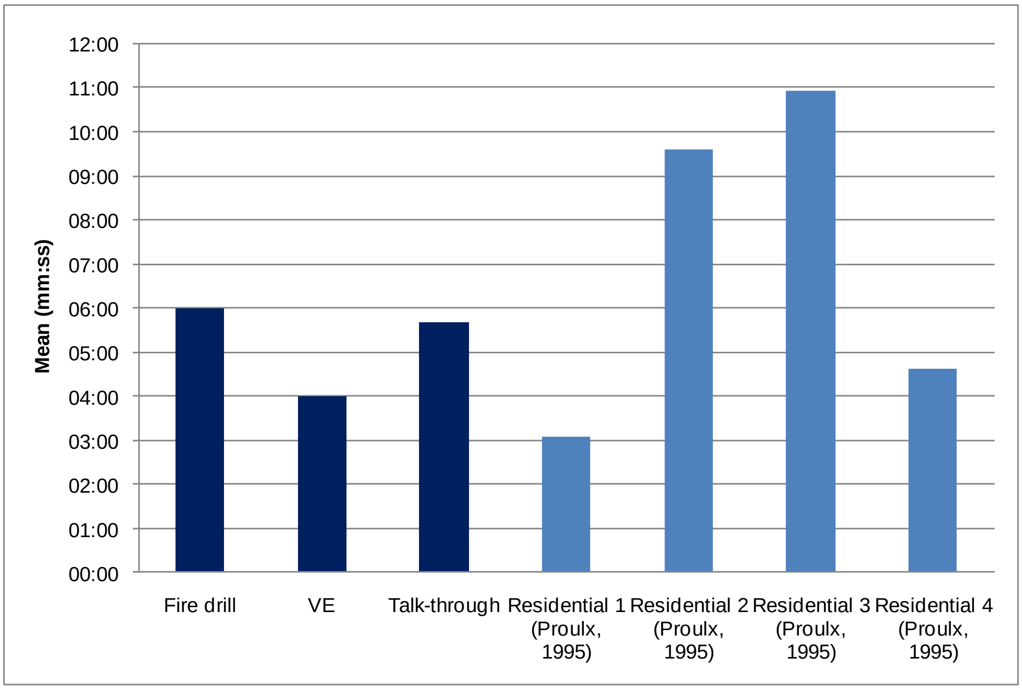

Figure 8.23. Mean building evacuation times. Fire drill values were taken from the video footage, VE values from the screen recordings, and the talk-through values were participants' estimated times.

However, as discussed earlier in this section, the total building evacuation time (i.e. longest time it took any one person to evacuate) is also a commonly used measure. The results are presented in Figure 8.24, with other values from previous scientific literature. It can be seen that the value for the talk-through approach (based on estimates of time taken to leave the building) produced the longest evacuation time.

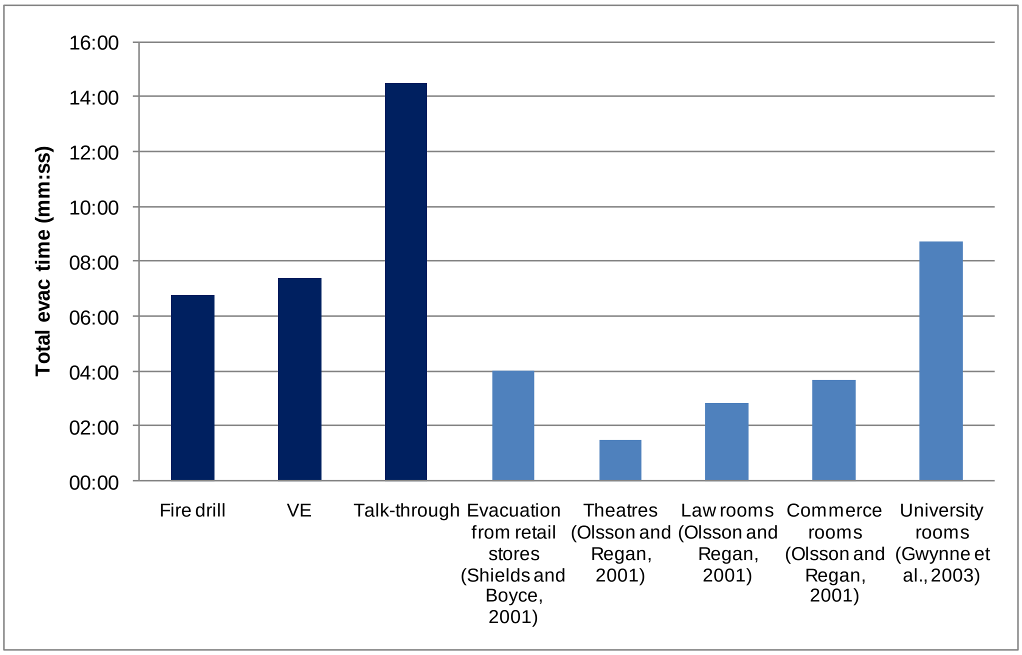

Figure 8.24. Total building evacuation time (maximum value for time taken to leave the building). Fire drill values were taken from the video footage, VE values from the screen recordings, and the talk-through values were participants' estimated times.

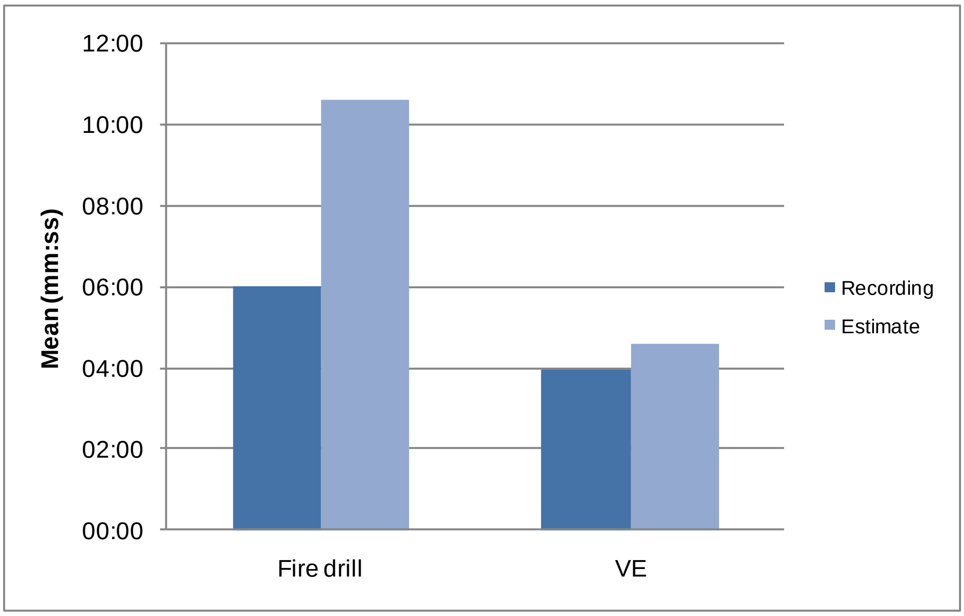

Figure 8.25. Mean times taken to leave building, showing values taken from the video footage (fire drill) and screen recordings (VE) shown against those taken from the participants' estimates for both studies

As part of the post-trial questionnaires in the fire drill and VE studies, participants were asked to judge "how long do you think it took to leave the building?" The mean values for the self-report measures of time taken to leave the building are shown in Figure 8.25, against those taken from the video footage and screen recordings. The descriptive statistics are shown in Table 8.28 (the objective measures were also shown previously in Table 8.20).

Table 8.28. Descriptive statistics for time to leave building (mm:ss) after alarm started: values taken from video footage (fire drill) and screen recordings (VE) and participants' estimates from the post-trial questionnaires

	mean	SD	Min	Max	N
Fire drill (from video footage)	06:00	01:07	02:57	06:45	22
Fire drill (estimate)	10:38	04:17	03:30	20:00	18
VE (from screen recordings)	03:59	02:26	00:27	07:23	14
VE (estimate)	04:35	03:31	01:00	11:00	11

The video footage from the fire drill was previously found not to demonstrate normality (Table 8.26). The VE self-reported time taken to leave the building also did not in this instance show normality according to Shapiro-Wilk ($W=0.832$, $df=11$, $p<0.05$). Therefore, Spearman's rho correlations were used to investigate the relationship between participants' self-reported times to the objective measures (Table 8.29).

Table 8.29. Spearman's rho correlation between subjective and objective measures for time taken to leave the building for the fire drill and VE

	r_s	N	p
Fire drill: self-report/video footage	0.237	18	NS
VE: self-report/screen recordings	0.859	11	<0.01

As for pre-evacuation time, the results demonstrated no significant relationship between the self-reported times and those captured on video for the fire drill. The VE once again demonstrated a large effect size, and significant relationship between the self-report measures and the screen recordings. As before, the results should be treated with caution. Participants were asked "how long did it take you to leave the building?", and may have interpreted this as the time from when the alarm started, or the time from leaving the computer room. Discretion was used in interpretation of participants' responses, and any unclear responses were omitted.

In the VE study, participants were also asked to rate their gaming experience. The results were correlated against time taken to leave A5, time taken to leave the building, and travel time (time from leaving A5 to leaving the building). No significant correlations were found (Table 8.30).

Table 8.30. Correlation between gaming experience and evacuation time

	r_s	N	p
Gaming experience and time taken to leave A5	-0.226	15	NS
Gaming experience and time taken to leave building	-0.431	14	NS
Gaming experience and travel time	-0.426	15	NS

Time taken to evacuate: key findings

- Significant differences for pre-evacuation times were seen between the fire drill and VE, and between the fire drill and talk-through
- The number of participants who evacuated before five minutes differed between the approaches
- For building evacuation times, a significant difference was found only between the fire drill and VE study.
- Mean pre-evacuation and building evacuation times were within the range shown in other studies in the scientific literature. However, the talk-through approach produced a long maximum building evacuation time, based on participants' estimates.
- High correlations coefficients (large effect sizes) were seen between screen recordings of the VE and self-reported evacuation times; no significant relationship was found between the video footage and self-reported evacuation times in the fire drill
- No significant correlation was demonstrated between gaming experience and any of the evacuation time measures.

8.4.5 Perception of danger

In the post-trial questionnaire for each approach, participants were asked to rate their perception of danger for different stages of the evacuation, including:

- on hearing the alarm,
- on deciding to leave the computer room, and
- when exiting the building.

These were post-trial reports for the fire drill and VE, and *expected* perception of danger for the talk-through. The median ratings on a scale from 1 (no danger) to 5 (great danger) are shown in Figure 8.26.

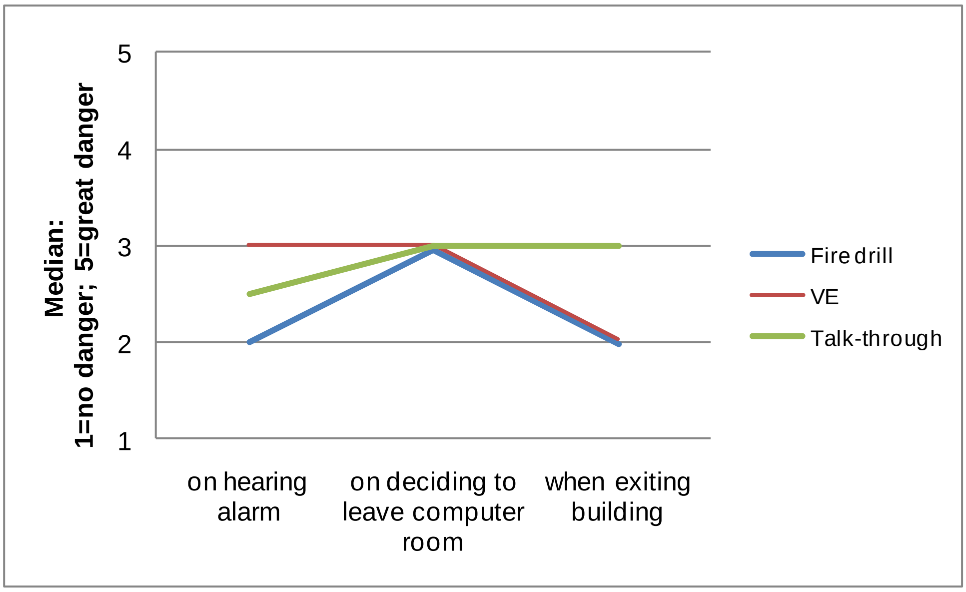

Figure 8.26. Perception of danger

The concurrent validity was investigated by testing for differences in the perceived level of danger between the predictive approaches for each question. There were no significant differences (Table 8.31). These results indicated that a similar perception of danger was experienced in all three conditions.

Table 8.31. Kruskal-Wallis investigation into differences in perception of danger between the predictive approaches

Rated perception of danger	X^2	df	p
On hearing alarm	3.045	2	NS
On deciding to leave computer room	4.309	2	NS
When leaving building	1.112	2	NS

The perception of danger when leaving the computer room was correlated against the time taken to leave the computer room. This was to investigate whether a higher perception of danger led to a shorter time to leave the room. The results can be seen in Table 8.32. Only the talk-through approach demonstrated a significant correlation (and medium effect size based on Cohen, 1988). This may have been because participants in the talk-through approach were not told to evacuate after five minutes in their hypothetical act sequences, and therefore the distribution of the data was not restricted by this cut-off. However, the VE demonstrated a medium effect size, albeit non-significant.

Table 8.32. Correlation between perceived danger when deciding to leave computer room and time taken to leave computer room

Approach	r_s	N	p
Fire drill	0.198	21	NS
VE	-0.347	15	NS
Talk-through	-0.469	22	<0.05

In the scientific literature there has been some evidence to suggest that a higher threat of danger has been associated with a quicker response (Edelman et al., 1980; Galea et al., 2009; McConnell et al., 2009), which supported the findings from the talk-through approach. A more general report by Pauls and Jones (1980a) was that occupants have often treated the sense of threat lightly. This provides evidence for the predictive validity of the approaches tested, given the low-medium perception of danger ratings made by participants in each trial.

Perception of danger: key findings

- In all conditions, perceived danger was rated medium to low; no significant differences existed between the ratings obtained for the different approaches
- The talk-through study showed a negative correlation and medium effect size between the hypothetical perception of danger and estimated time taken to leave the computer room. This association has been reported in studies of real emergencies.

8.4.6 Exit choice

The three exits used by the participants in each of the studies are shown in Figure 8.27. Exit A is accessed from a corridor stairwell and leads to a car park at the rear of the building. Exit B is the main entrance/exit door to the building. Exit C is accessed from stairs in the foyer to an exit route from the basement to the same car park at the rear of the building, but through a different external door. Images from the areas of the building leading to these exit routes can be seen in Figure 8.28-Figure 8.30.

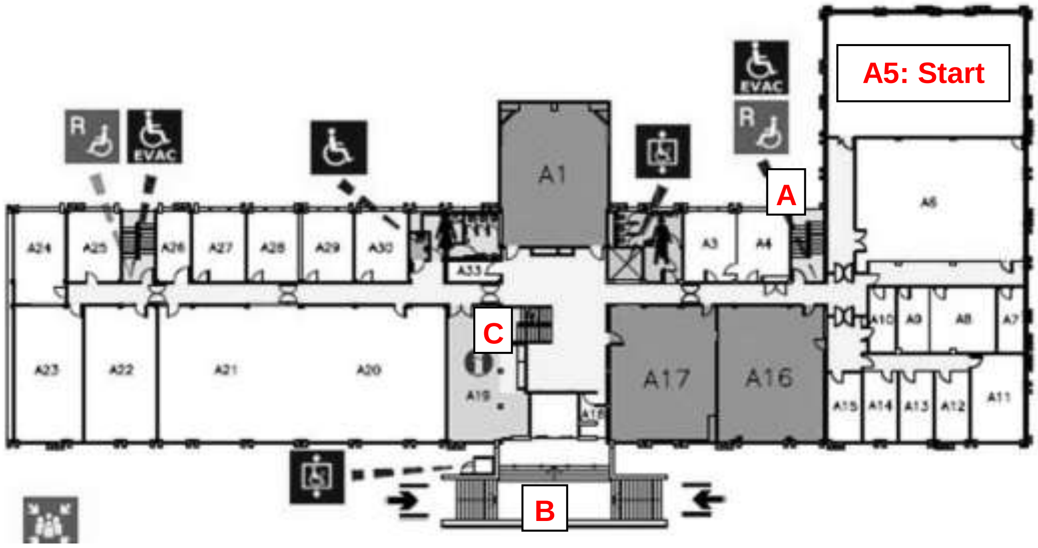

Figure 8.27. Plan view of A-floor

Figure 8.28. Exit A: corridor stairwell

Figure 8.29. Exit B: main entrance/exit

Figure 8.30. Exit C: stairs leading to basement and exit to car park

The exits used by the occupants can be seen in Figure 8.31, shown as percentages for each approach. The fire drill was based on the video footage, the VE based on the screen recordings and the talk-through values were participants' anticipated exit routes. Figure 8.31 shows a preference for the front/main entrance/exit, although over 30% of participants in the VE study took the corridor exit (A).

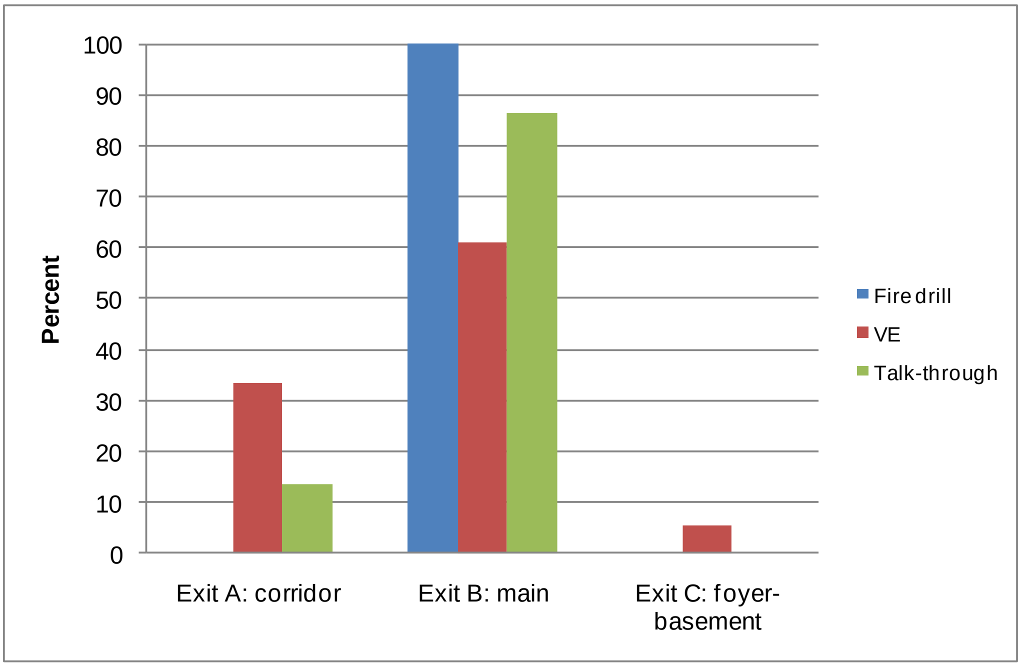

Figure 8.31. Exit choice (as a percentage of exit usage in each condition)

A chi-square test showed significance (X^2=12.199, df=4, $p<0.05$), which indicated an association between the predictive approach and route choice[a].

The exit choices demonstrated were generally consistent with the findings reported in the scientific literature, indicating predictive validity. These can be summarised as a tendency to use the most familiar entry/exit route (the main entrance), although some use is reported of the nearest exits (Canter et al., 1980; Edelman et al., 1980; Mawson, 2005; Xudong et al., 2009). Thus, the fire drill and the talk-through approach revealed exit choices consistent with these findings.

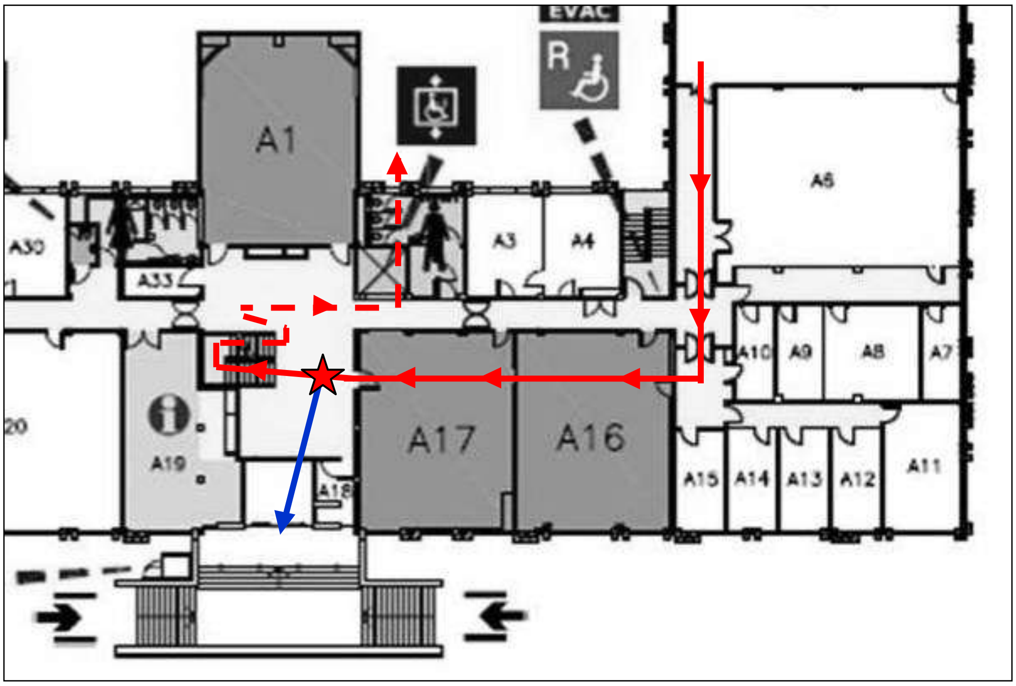

Figure 8.32. Exit route taken by one participant in the VE study (shown in red). The dashed line indicates the distance travelled on lower floor. Blue line shows a much shorter and more straightforward exit path.

The exit choices seen in the VE were generally acceptable, apart from one participant who used Exit C (foyer stairwell to basement). The exit path taken by this avatar is shown by the red line in Figure 8.32; the dashed line indicates the distance travelled in the basement. From the position indicated by the star, the main entrance was a simple path approximately five metres to the left. However, the participant directed their avatar down the foyer stairs, and looked round the basement before finding a fire exit. While a further avatar was in the vicinity of the stairwell at the time this decision was made, the participant did not follow them directly, and therefore their decision to extend considerably the time and distance to leave the building was only understandable if the main entrance/exit was unnoticed. This was likely; the restricted

[a] Six of the nine cells had an expected frequency smaller than five, and therefore the results should be treated with caution.

field of view in the VE meant that the main entrance was not visible from the position indicated by the star unless the participant turned their avatar to the left. In the real building, peripheral vision, plus other cues such as noise and light would make the main entrance much more obvious from this location. This raised concerns about navigation in the virtual environment, a point which will be revisited in the discussion.

Participants were also asked to rate in the post-trial questionnaire how influential several different factors were on their choice of exit, from 1 (not at all influential) to 5 (very influential). For the fire drill and VE, these were based on their experiences during the trial; the talk-through asked for hypothetical influences on exit choice. The results are shown in Figure 8.33.

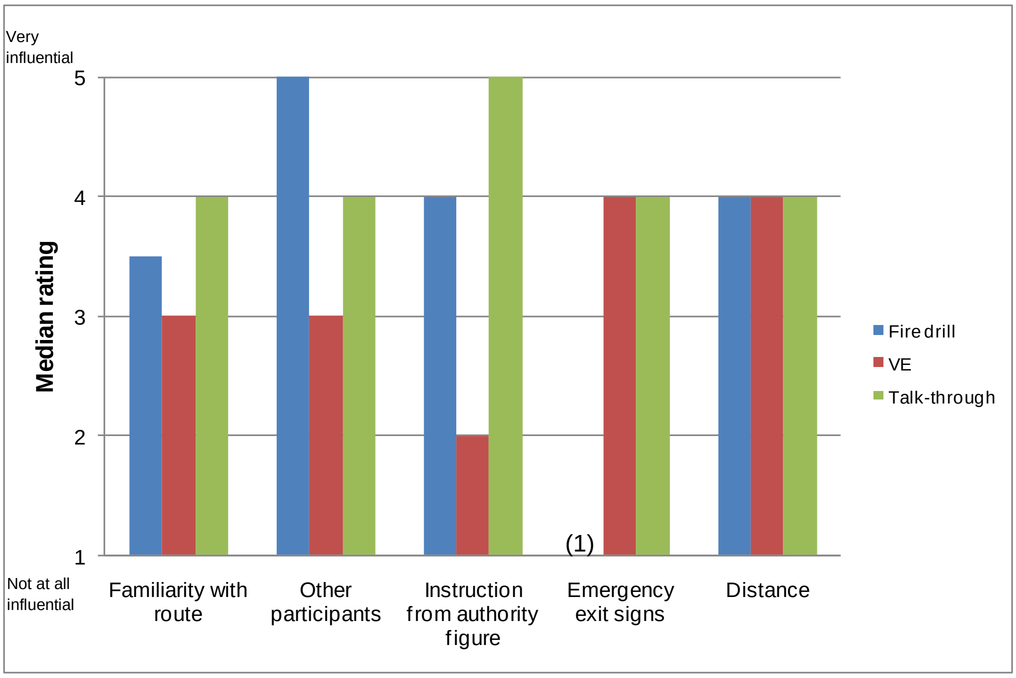

Figure 8.33. Median rating for influence on choice of exit

Kruskal-Wallis tests were run for each of the influencing factors on exit choice to investigate differences between the predictive approaches. The results are shown in Table 8.33. For each significant finding, stepwise Mann-Whitney U tests, with Bonferroni correction, were run to identify the significant differences in each condition, as shown in Table 8.34 - Table 8.36.

Table 8.33. Kruskal-Wallis investigation into the differences in influences on route choice between the different conditions

Influencing factor	X^2	df	p
Familiarity with route	6.190	2	<0.05
Other participants	4.082	2	NS
Instruction from authority figure	13.888	2	<0.01
Emergency exit signs	19.768	2	<0.001
Distance (i.e. shortest route)	2.392	2	NS

Table 8.34. Mann-Whitney U comparisons: Familiarity with route

	Mann-Whitney U	p (0.017)
Fire drill – VE	133.50	NS
Fire drill – Talk-through	176.50	NS
VE – Talk-through	89.5	<0.017

Familiarity with route was a significantly greater influence on choice of exit in the talk-through approach than with the VE. It was difficult to explain these findings from the results. However, it was anticipated that in the VE factors other than familiarity, such as signage, were more influential on the choice of exit. As described earlier in this section, the scientific literature suggests that familiarity has a notable effect on exit route.

Table 8.35. Mann-Whitney U comparisons: Instruction from authority figure

	Mann-Whitney U	p (0.017)
Fire drill – VE	134.00	NS
Fire drill – Talk-through	128.00	<0.017
VE – Talk-through	63.50	<0.017

Instruction from authority figure was significantly greater in the talk-through condition than in the VE or the fire drill. This may have been due to inherent differences in the applications of the approaches. In the talk-through approach participants were asked "How influential would direction by an authority figure be on your choice of exit?" However in the fire drill and VE, participants were asked "How influential was direction by an authority figure on your choice of exit?" Thus, in the latter, participants who had no contact with the experimenter likely rated this lower. In fact, the experimenter did not actually instruct the participants on which exit to take, only to evacuate. It was likely however that the instruction to leave by the experimenter affected the ratings for this particular question.

Evidence exists in the scientific literature to suggest that instruction from an authority figure is influential on the evacuation (Gershon et al., 2007). Thus the results for the talk-through, and to a lesser extent the fire drill, seem more appropriate than those seen for the VE.

Table 8.36. Mann-Whitney U comparisons: Emergency exit signs

	Mann-Whitney U	p (0.017)
Fire drill – VE	87.00	<0.017
Fire drill – Talk-through	49.00	<0.017
VE – Talk-through	163.50	NS

The VE and talk-through approaches yielded a significantly greater influence of signs on choice of exit than in the fire drill. In the VE signage may have helped navigation which can be harder due to reduced field of view (Neale, 1997). However, it was interesting to note that the talk-through approach yielded a significantly greater (predicted) influence of exit signage than was reported by the participants in the fire drill.

Participants were also asked to rate their familiarity with the Psychology building on a 5 point scale, ranging from not at all familiar to very familiar, as differences may have influenced their choice of exit. The median ratings for each condition are shown in Figure 8.34. A Kruskal Wallis test revealed no significant difference between the conditions (X^2=0.764; df=2; p=NS).

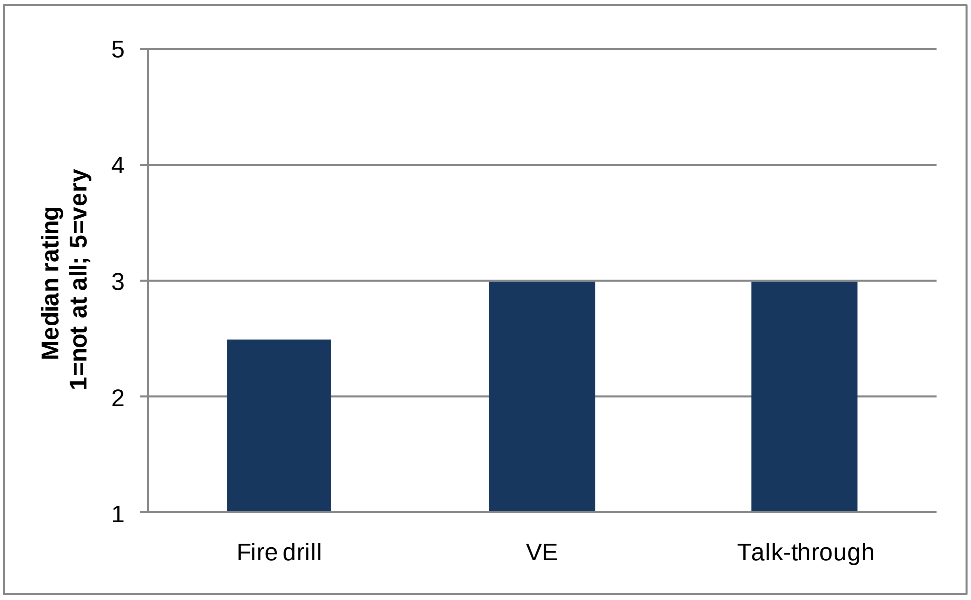

Figure 8.34. Familiarity with Psychology building

Exit choice: key findings

- Participants in each of the studies demonstrated an overall preference for the main exit; however a high proportion of participants in the VE used the nearest exit.
- Exit choice was generally compatible with the scientific literature, although concerns were raised about navigation in the VE due to an unusual exit choice.
- Emergency exit signs were seen to exhibit a greater influence on exit choice in the VE and talk-through than in the fire drill.

8.5 Discussion

The discussion is based upon consideration of the results and application of the approaches against the criteria for judging the quality of an approach, namely validity, reliability, sensitivity, resources, and ethics. Unless stated, analysis is conducted at the *approach* level, i.e. the combination of a *setting* and *method* to produce a *measure*, as defined in Section 1.5, although comments are made on the settings, methods and measures individually in several instances.

8.5.1 Validity

The validity is discussed in relation to each aspect of the experiment reported in the results.

Frequency of acts

The frequency of acts obtained from the fire drill, VE and talk-through approaches demonstrated medium or large effect sizes when correlated with one another. Furthermore, the fire drill and VE showed significant relationships and large effect sizes between the frequencies of behaviours captured on video/screen recordings and from post-trial questionnaires asking participants to report the actions they had taken.

The comparability across the approaches gave an indication of their concurrent validity for frequencies of acts although the strength of this comparability is limited to the effect sizes indicated which were based on rank correlations. The validity of the approaches for obtaining frequency of act data was supported through an investigation of predictive validity, by qualitative comparison with the scientific literature. In the fire drill several of the act categories were only captured in the post-trial questionnaire as they were too detailed or cognitive to be captured by video. These included “note behaviour of others (ambiguous)”, “note behaviour of others (unambiguous)”, “dress/gather valuables”, “perception of stimulus (ambiguous)”, and “note nothing unusual/stay calm”. This limitation of the video analysis method could be easily rectified in a fire drill type study such as this one through the addition of a post-trial questionnaire. However, it may not be resolvable in situations in which CCTV analysis is conducted of unplanned emergencies. It may also not be possible in public buildings, if the evacuees leave after the event.

Some problems were identified with the act frequencies obtained from the VE approach. In particular, several participants encountered problems with the VE, which included avatars freezing, loss of control of the direction of gaze, or difficulties in getting through the doors. Several participants overshot when navigating in the environment. Getting lost was a relatively common occurrence, which may in part have been due to the restricted field of view in the desktop VE (Neale, 1997). These acts led to reduce the predictive validity of the VE for obtaining representative acts.

However, the act frequencies were significantly correlated with those from the other approaches (*concurrent validity*, described above). A further problem with the VE setting was that participants had no ability to dress or gather valuables. This may be possible with a more developed VE, but is unlikely to be as natural as in the real world.

Finally, the frequency of predicted acts obtained from the talk-through approach demonstrated no particular issues either in comparison to the other approaches or to the scientific literature.

Note that while some concurrent validity is implied through a significant correlation, there can still be largely different frequencies for any particular act. Therefore, the results must be used with caution.

Sequences of acts

As for frequencies of acts, the approaches demonstrated significant relationships for the sequences of acts when correlated against each other, with medium or large effect sizes. The associations indicated concurrent validity, to the levels implied by the correlation coefficients/effect sizes.

For the fire drill approach, the decomposition diagram shows a relatively consistent response by participants in which they generally ignored the alarm, received the instruction to evacuate, and then left the building with other people. The decomposition diagrams for the VE and the talk-through approaches demonstrate a more variable response. The differences may have in part been due to differences in group sizes, with the VE study being conducted in groups of 2-4 participants, and the talk-through being conducted individually. The fire drill was studied as one large group. That said, the fire drill group was not entirely cohesive with some participants evacuating before being told to do so.

No approach stood out with particularly high agreement or disagreement with the available scientific literature on sequences of action. It should be noted that, as for frequency of acts, the data from the self-report questionnaires for the fire drill and VE were necessary to support a full comparison of the results with the literature.

Time to evacuate

Considering first the time to leave the computer room, participants in the fire drill overall took significantly longer than the other approaches, which demonstrated no significant difference between them. When investigating the number of participants who evacuated before being told to do so, again the VE and talk-through demonstrated much higher numbers than in the fire drill. The mean pre-evacuation times in the literature were inconclusive due the spread of data, which suggested that while the concurrent validity of the approaches used in the standardised comparison was low, the times obtained were within reasonable tolerances. Thus, despite the differences found between the conditions, the mean times were still within the range found in the literature implying predictive validity. Note that participants in the fire drill and VE were told to evacuate after five minutes. Without this intervention, the mean pre-evacuation time for the fire drill in particular may have approached the upper boundary of acceptability. However, given Proulx's (2007) statement that participants will rarely respond to an alarm without additional cues such as sight or smell or smoke, the fire drill times were deemed acceptable, despite the differences to those from the VE and talk-through, and similar fire drill studies reported in the literature (Olsson and Regan, 2001; Gwynne et al., 2003).

A further interesting finding was that participants were unable to accurately recall the time taken to leave the computer room in the fire drill study, another facet of concurrent validity with particular relevance to approaches which use post-event surveys. This supports the uncertainty by Proulx and Reid (2006) about this method of capturing quantitative data. Thus, concerns were identified with the validity of the participant reports of time to leave the computer room.

Considering the time taken to leave the building, a comparison of all three approaches demonstrated a significant difference only between the data produced in the fire drill and VE approaches. The mean time to evacuate the building was less with the VE. The mean times from each approach were within reasonable limits when compared to values published in the literature. However, the talk-through approach revealed a notably long maximum building evacuation time, based on participants' estimations. As was the case for pre-evacuation times, concerns were raised about the concurrent validity of participants' reports of the time taken to evacuate the building in the fire drill study when compared to the data obtained from the video footage.

No significant correlation was found between gaming experience and time to leave A5, time to leave the building or travel time between A5 and the exit, although in each case the correlation coefficient was negative, indicating that more experienced gamers took less time to evacuate. Previous studies have demonstrated an association between user-estimates of gaming experience and performance in desktop VE navigation tasks (Smith and Du'Mont, 2009). Expert gamers have also been seen to take less time to evacuate from a building in a VE evacuation task than non-gamers (Smith and Trenholme, 2009). While the negative correlation coefficients found in the VE study (Table 8.30) indicated the possibility of an influence of gaming experience, the lack of a significant finding suggested other factors accounted for the variation in evacuation times.

Perception of danger

Participants in each condition reported a mid to low level of danger. While there was no actual danger or presence of smoke or fire in any of the conditions, people in actual emergencies have also reported a perceived low level of risk (Pauls and Jones, 1980a). This finding indicates predictive validity in the approaches used in the standardised comparison for this measure. No significant differences were identified between the perceptions of danger in each condition, at each of the three stages of the evacuation: on hearing the alarm, on deciding to leave the computer room, and on exiting the building. Thus, concurrent validity between the approaches was indicated. Only the talk-through approach demonstrated a significant correlation and medium effect size between the anticipated perception of danger when deciding to leave the computer room and the time taken to leave the computer room. This indicated an association between a higher perception of danger, and a shorter time to leave the computer room. There has been some evidence in the literature to support this (Galea et al., 2009). The absence of a correlation in the other approaches may have been affected by the instruction to the participants to leave the building after five minutes. However, this instruction was a necessary part of the method for the fire drill and VE, and therefore the talk-through approach justifiably demonstrated greater predictive validity for the effects of perception of danger on evacuation time.

Exit choice

All participants evacuated through the main door in the fire drill; most did in the VE and most predicted they would do so in talk-through approach. The corridor exit was the second most popular choice in the VE and talk-through studies; in the VE one person exited through the basement exit which was accessed through the stairwell in the foyer. Apart from this last participant, the findings were proven to be consistent with those in the scientific literature, indicating predictive validity in the approaches.

It is pertinent to mention the circumstances which led to the removal of four participants from the analysis of the VE data. Adding textures to an object in Second Life can sometimes place the image on all sides of an object. Thus, when creating fire signs, the words “Fire exit” were placed on the back and sides of the cuboid. In one instance in the VE, this resulted in a fire sign inaccurately leading participants into the corridor on the far side of the foyer (Figure 8.35). In reality, the back of this sign is blank (Figure 8.36). During the VE study four of the participants entered this corridor. Another walked towards it, then returned to the main entrance; their data were retained and this incident was considered an overshoot. The four participants who entered this corridor would only have had to travel a short distance to their left to exit via the main entrance (Figure 8.27). However, they chose not to exit, instead continuing straight on.

Figure 8.35. VE anomaly

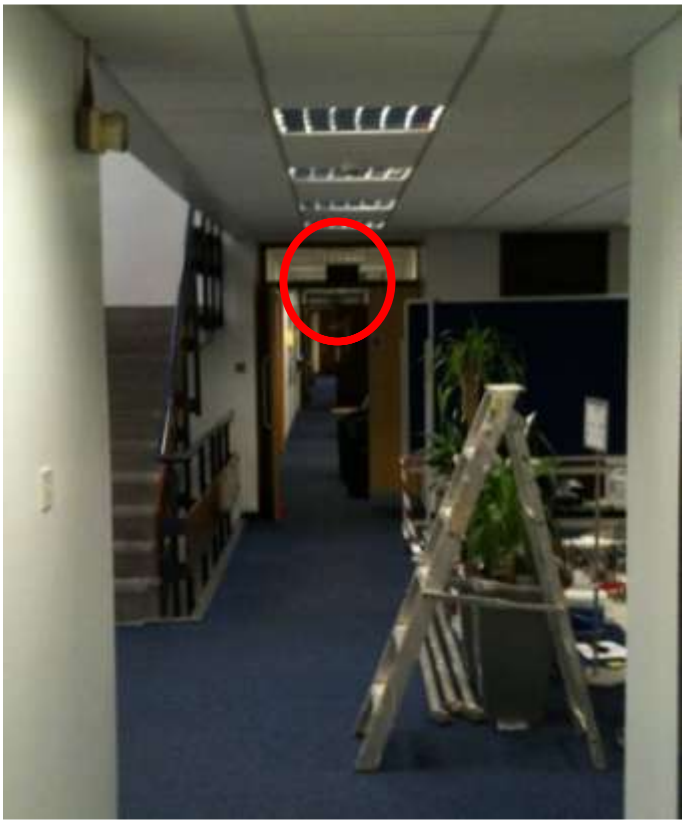

Figure 8.36. Photo of real building

It is unclear whether participants in the VE continued straight on due to the inaccurate signage or simply because they failed to notice the main entrance on their left. However, no-one in the fire drill made this mistake. In the real building, the main entrance is very obvious, and an evacuee would be unlikely to miss it. It was easily done in the VE – the main entrance only become notable upon deliberately turning the avatar's head to the left. Thus, the concurrent validity of the VE for exit choice was compromised, as participants were less aware of their surroundings than participants were in the real environment. This may have been in part due to the restricted field of view in the VE (Neale, 1997).

Finally, considering the influences on exit choice, participants in the talk-through approach reported a higher predicted influence of familiarity than in the VE. The importance of familiarity has also been reported in the literature. Anticipated instruction by authority figure also received higher ratings in the talk-through approach, but this may have been because the actual instruction by authority figures in the VE and fire drill was limited to being told to evacuate. The VE and talk-through approach reported a greater influence of signage design on the exit choice than in the fire drill. In the latter, some of the participants reported noticing a fire exit sign but walking straight past it. While some evidence was found in the scientific literature for the influence of signage design, it appeared to be rated too highly in the VE and the talk-through approaches.

Summary of validity

A summary of the discussions of validity is shown in Table 8.37.

Table 8.37. Summary of discussion of validity

Criteria	Fire drill	VE	Talk-through
Frequency of acts	Required a post-trial questionnaire, but consistent with other approaches and the literature, indicating concurrent and predictive validity.	Some issues caused by problems with the VE; some people got lost. Otherwise comparable to other approaches (concurrent validity).	No problems identified, thus indicating concurrent and predictive validity.
Sequence of acts	Consistent with other approaches (concurrent validity).	Consistent with other approaches (concurrent validity).	Consistent with other approaches (concurrent validity).
Time to evacuate	Time to leave computer room and time to evacuate building were within a range of times found in the scientific literature, indicating predictive validity.	Time to leave computer room and time to evacuate building were within a range of times found in the scientific literature, indicating predictive validity.	Mean estimated time to leave computer room and time to evacuate building were within a range of times found in the scientific literature, indicating predictive validity. Maximum building evacuation time was greater than previous reports in the literature.

Criteria	Fire drill	VE	Talk-through
Perception of danger	General level consistent with other approaches and the literature, indicating concurrent and predictive validity.	General level consistent with other approaches and the literature, indicating concurrent and predictive validity.	General level consistent with other approaches and the literature, indicating concurrent and predictive validity. Demonstrated a (negative) correlation between perceived danger and time to leave computer room, which provides further evidence for the predictive validity.
Exit choice	All participants evacuated via main door, demonstrating predictive validity.	Some concerns raised about exit choices. Influence of exit signage rated higher than expected.	Most participants predicted evacuating by main door; some would use nearest exit. Acceptable predictive validity. Anticipated influence of exit signage rated higher than expected.

8.5.2 Reliability

This section discusses the reliability of the three approaches. To do this, reference was made to previous applications of each method, to identify whether a similar pattern of findings was obtained for each application. Thus, the focus of the investigation was replicability of the approaches.

Fire drill

The fire drill was compared to the study of the hotel fire evacuation in Saariselkä (Section 4.4). This was a detailed source of similar data, albeit from a different scenario. A comparison of the results serves as an indicator of the replicability of the approach (Table 8.38).

Table 8.38 demonstrates good comparability between each of the criteria for the Saariselkä fire drill, and the fire drill evacuation from the School of Psychology. The one exception is a different pattern in the median ratings for perception of danger. In particular, evacuees from the hotel fire clearly felt no danger upon exiting the building; a slightly higher perception of danger was noted at this stage of the Psychology building evacuation.

Table 8.38. Investigation of the reliability of fire drills as a predictive approach

Criteria	Hotel fire evacuation (Section 4.4)	Psychology fire drill
Frequency of acts	Significant correlation and medium effect size with Canter et al. (1980) (r_s=0.414, N=52, p<0.01)	Significant correlations with results from VE (r_s=0.689, N=31, p<0.001) and talk-through approach (r_s=0.575, N=31, p<0.01)
Sequence of acts	Significant correlation with Canter et al. (1980) (r_s=0.459, N=26, p<0.05)	Significant correlations with results from VE (r_s=0.429, N=169, p<0.001) and talk-through approach (r_s=0.534, N=121, p<0.001)
Time to leave room occupied at time of alarm	Mean=3:54, range=2:30 – 5:00. (Times were estimated by participants post-evacuation). Comparable with scientific literature.	Mean=5:16, range=00:50 – 06:14. (Times were captured from video footage). Comparable with scientific literature.
Time to leave building	Mean=4:24, range=3 – 5:30. Within the boundaries presented in other scientific investigations.	Mean=6:00, range=02:57 – 06:45. Acceptable when compared to scientific literature.
Perception of danger	Median ratings decline from 3 when alarm was first heard, to 2 when deciding to leave room to 1 when exiting the building.	The comparable median ratings were 2, 3, 2.
Exit choice	All participants exited via main entrance.	All participants exited via main entrance.

These findings indicate that the data obtained from fire drills are replicable, although obviously this comparison is limited to the two instances shown above.

Virtual environments

To investigate the replicability of VEs, comparison was made to the study conducted by Yogesh (2009) (Section 6.4). While this does not offer as detailed an assessment as was made for the fire drills, some conclusions can be drawn.

In Yogesh's (2009) study, participants evacuated immediately; they were told to leave the building and timed from their initial movements. Thus, the time can essentially be considered movement time. Yogesh (2009) found that evacuation in the VE was significantly longer than in the real environment, and also more variable. A similar test was performed to compare the Psychology building VE travel time to the fire drill travel time (time after leaving A5 to exiting the building). The travel times were found not to be normally distributed (Table 8.39), therefore the Mann-Whitney comparison was used. This showed no significant differences between the groups (U=127.5; N=22,15; p=NS). However, as in Yogesh (2009), the VE showed much greater variability than in the real world (Table 8.40). Thus, the VE approach demonstrated a longer mean time to evacuate than in the real world in both studies (although not significantly different in the standardised comparison), and greater dispersion in the data. Yogesh (2009) also reported problems with the VE such as participants getting lost and stuck at doors; problems which were found in the Psychology VE evacuation.

Table 8.39. Normality tests for time taken to leave building (travel time)

Time to leave building	Shapiro-Wilk	n	p
Fire drill	0.879	22	<0.05
VE	0.676	15	<0.001

Table 8.40. Comparison of travel times from Yogesh (2009) to the standardised comparison

	Mean (seconds)	**SD (seconds)**
Yogesh (2009) real evacuation	21.79	5.28
Yogesh (2009) VE	43.03	17.91
Psychology evacuation: fire drill	34.64	4.17
Psychology evacuation: VE	74.27	79.47*

* Range=19-280, suggesting a large positive skew

Concerning exit choice, Yogesh (2009) found no significant difference between the real and virtual environments, and 85% of participants left via the main entrance, with the remainder leaving by a closer stairwell exit. In the VE Psychology evacuation, 61% left by the main entrance, with 33% leaving via a closer corridor stairwell and 6% leaving through the foyer stairwell. Treating the exit choice as “main”, “stairwell” and “other”, a chi-square provided no evidence for a difference in exit choice between Yogesh (2009) and the Psychology evacuation (x^2=2.227; df=2; p=NS).

These results indicate that the VE approach has demonstrated a similar pattern of results (a longer and more variable travel time than in the real world[a]) and similar issues (e.g. getting lost, stuck) on repeat applications. The exit choices by the participants also appeared to be similar on both applications. Unfortunately the original study (Yogesh, 2009) did not provide an opportunity to investigate the frequency or sequence of acts or perception of danger.

[a] While this discussion has focussed on replicability, the more variable travel times seen in the VE also have implications for the concurrent validity of this measure.

Talk-through

The development of the talk-through approach (Chapter 7) already provided evidence of the replicability of this approach through consistent correlations with the frequency and sequence of acts from Canter et al. (1980)(Studies 3a-3c). The standardised comparison (Study 4) demonstrated consistent correlations with the frequency of acts from the fire drill and VE. These results are summarised in Table 8.41.

Table 8.41. Results from each application of the talk-through approach

	Initial investigation (Study 3a)	Development study (Study 3b)	Further investigation (Study 3c)	Standardised comparison	
				Fire drill	VE
Frequencies of predicted acts	r_s=0.694, N=23, $p<0.001$	r_s=0.323, N=55, $p<0.05$	r_s=0.572, N=0.572, $p<0.001$	r_s=0.575, N=31, $p<0.01$	r_s=0.394, N=31, $p<0.05$
Sequences of predicted acts	r_s=0.457, N=21, $p<0.05$	r_s=0.377, N=32, $p<0.05$	r_s=0.344, N=40, $p<0.05$	r_s=0.534, N=121, $p<0.001$	r_s=0.492 N=196, $p<0.001$

Thus, while this comparison is limited to only the frequency and sequence of hypothetical acts, the talk-through approach has consistently demonstrated medium to high correlation coefficients with the reference studies in each application. This finding implies replicability in the approach.

8.5.3 Sensitivity

In this section, discussion will be made of the appropriateness of the outcome from each approach for human factors predictions of behaviour in emergencies. To do this, it was necessary to first consider the possible applications of the predictive approaches, introduced in Chapter 1. In a human factors context, these are likely to include informing the design of buildings and systems, the development of emergency response procedures and use during the development or execution of training.

Firstly, all the approaches were able to give an indication of what acts people will do during a fire drill evacuation. This is likely to be essential information for a human factors (HF) professional during design, training and procedure development. For example, one of the origins of this research was Emergency Response Coordinators asking human factors researchers at the University of Nottingham what people would do in an emergency situation, albeit for a different scenario than a fire drill.

The approaches in the standardised comparison not only indicated what acts may be conducted, but also provided sequential information regarding the order with which they may be conducted. Thus, a more detailed and elaborate picture of the predicted behaviours was possible. The importance of understanding behaviours in sequence was explicitly reported by Canter et al. (1980). As mentioned in Section 8.4.2, to maximise the utility of the frequency and sequence of act data from a fire drill or VE approach it was necessary to supplement the video/screen capture data with a post-trial questionnaire in which participants report their actions. This was particularly important for some of the cognitive acts, such as “noting the behaviour of others” or “perception of stimulus”.

Time taken to evacuate would also be a useful measure for an HF practitioner during the development of emergency response procedures. For example, with knowledge of the delay to evacuation demonstrated in the Psychology fire drill, recommendations could be made that an authority figure sweeps the building as soon as possible, instructing occupants to leave. However, the subjective measures for time taken to evacuate were not shown to be accurate for the fire drill evaluation. A further concern was the issue raised by Gwynne et al. (1999) that each approach provided only one instance (or a very small number of instances) of evacuation time, and that over several runs a distribution of times would be expected. As a hypothetical example, the time measured for evacuation during the fire drill study might have been from a particularly slow group; on a repeat application the participants may have left the building much faster.

The ability of the approaches to obtain an indication of perceived danger could be used to understand the level of threat experienced by occupants’ during an emergency. This could in turn help predict their cognitive states and therefore their ability to conduct tasks during the emergency. This may be of particular use for someone with specific tasks they must complete during an emergency, such as a Fire Officer, as it could be used to indicate their capability to perform these tasks.

Predictions of exit choice would be useful during building design, for example to ensure that the exits are sufficiently large to accommodate the anticipated number of people. Similarly, when developing emergency response procedures consideration should be given to the anticipated flow of evacuees. Staff training would benefit from this prediction, to avoid for example any misconceptions that the majority of evacuees will leave through the nearest exit. The subjective ratings for influence on exit choice would support this work, to give some understanding of why occupants made various decisions during their evacuation. Thus, this information would support a more effective design, which is sympathetic to the different strengths of influence on exit choice.

One disadvantage of the fire drill approach is that it requires a physical representation of the building of interest; thus investigating design solutions would necessarily be at a later stage in the development process than with the other approaches. For example, investigating different designs in the VE would simply be a matter of altering the VE; with the talk-through approach different designs could be investigated with a change to the scenario. The level of detail which can be investigated with the talk-through approach may be limited however, due to its use of fairly basic building plans and short verbal scenarios, rather than the rich environment used in VEs. This concern was supported by the higher than expected ratings for influence of signage design in the talk-through approach. While the VE also received high ratings for the influence of signage design, this was attributed to navigation difficulties caused by a restricted field of view (Section 8.5.1) rather than a lack of sensitivity to detail in the environment.

A further disadvantage of the practical application of the talk-through approach is that it has limited ability to control for the prediction of social interaction between the participant and other people who may be involved in the scenario. For example, if during their predictions the participant states that they would “ask the person next to them what to do” the experimenter needs to be prepared with a response. These interactions are in fact an important part of emergency scenarios (Pauls and Jones, 1980b; Sime, 1995; Aguirre et al., 1998; Mawson, 2005; Drury et al., 2006; Pan et al., 2006; Gershon et al., 2007).

To summarise the discussion on sensitivity, each of the approaches studied would be suitable for providing information which has several uses in human factors applications. The fire drill approach is however dependent upon the existence of a physical representation of the building for which the prediction is being made. The talk-through approach requires thorough preparation to address social interactions. This discussion has assumed that the methods used for the fire drill and VE video/screen capture data are supplemented with post-trial questionnaires, which record participants’ perceptions of what happened, as well as subjective ratings for influence on exit choice.

8.5.4 Ethics

Each of the approaches used had ethics considerations, particularly due to the nature of the scenario under investigation. The fact that a fire evacuation investigation was being conducted meant that participants had to be screened for any mental illness, or for having close friends or family members who have been involved with a fire to minimise the risk of distress. This was relaxed slightly for the fire drill, as these are run annually without any screening of participants. However, if the study had deviated from a simple evacuation (for example by artificially blocking exits or by generating smoke) considerably more effort would have been required in the precautions to ensure that participants did not experience any distress or risk of physical injury. With the scenario tested, the VE study and the talk-through approach led to slightly more concerns about the well-being of participants than the fire drill, simply because these types of events are not run annually like fire drill evacuations and therefore less was understood about how participants would react. Despite these considerations and precautions, no participants experienced any notable distress, or sustained any injury, during any of the trials.

The VE setting presented the risk of simulator sickness (Nichols and Patel, 2002). A small number of people are particularly susceptible to the symptoms of simulator sickness, and this necessitated further screening, and a briefing on the symptoms they should look out for, prior to the start of the experiment. Fortunately, no-one complained of any significant symptoms, other than annoyance at the continuous fire alarm noise in the headphones.

The participants in the fire drill and VE studies were deceived to a certain extent about the true purpose of the investigation, having been recruited believing they would be participating in a navigation experiment. This was to ensure that their behaviours and actions were not affected by pre-warning of the evacuation. There was no other way to obtain the results, and measures were implemented to mitigate any after-effects, for example screening participants for mental ill-health and providing a post-event de-brief. However, deception in research trials requires careful consideration and sound justification (Banyard and Flanagan, 2005).

In summary, no participants experienced distress or injury in any of the trials. However, if the scenario had been any different to a typical building evacuation, the fire drill approach would have required greater consideration to protect the physical, as well as mental, well-being of the participants. The VE presented the risk of simulator sickness, although no participants reported any adverse effects during this trial. The use of deception requires careful consideration, and strong justification, in any research.

8.5.5 Resources

The resources associated with each approach can be seen in Table 8.42. Overall, the talk-through approach had the lowest resources. This was expected, and the resources of this approach were commented on in Chapter 7. The approach involved no specialist equipment, and had very low development costs. The fire drill had the added expense of video cameras, and extra expense for participants as the study took a whole hour to conduct, whereas the talk-through approach took under half an hour. If conducted during office hours, the fire drill would also have the cost associated with disruption to normal working activities. Finally, an additional cost associated with the fire drill was the specialist staff required: senior academics were present in case of any unanticipated problems, and a porter was required to manage the alarms and locking/unlocking of the building. The talk-through and VE studies were conducted entirely by the experimenter themselves.

Table 8.42. Resources associated with each condition

Criteria	Fire drill	VE	Talk-through
Participant payment	***£10/participant*** trial lasted one hour	***£5/participant*** trial lasted 30 mins	***£5/participant*** trial lasted no longer than 30 mins
Training		***1-2 weeks*** to learn Second Life	
Development/ preparation	***Three working days.*** The development work involved a series of planning meeting, works requests, and correspondence to organise the drill. This was done over several weeks, but did not amount to a great deal of time.	***Three weeks.*** The Second Life development took about six weeks (part-time), which equates to approximately three weeks full time. This would have been extended by ~one to two weeks had the building not have been started already.	***Three working days.*** The development and preparation work was limited to refining the scenario, and administration work such as participant recruitment.
Equipment	***£650*** Six video cameras on suction pads. These were cheap cameras, approximately £100 each, plus memory cards, plus suction pads.	***~$12.50 + $40.00 US/month*** Land in Second Life costs money, the value of which fluctuates depending on “market conditions” (Linden Research Inc, 2010). As a crude guide, the area used for the VE study could be purchased for around $12.50, but with a monthly fee of $40. Also required five standard computers to run the study.	***Nominal*** The equipment involved only printed questionnaires.
Cost of conducting trial	***7 hours*** Three researchers for two hours. Two senior academic staff (in case of any problems). Two hours for a porter to unlock building, turn on/off alarms.	***7 hours*** One researcher for each group of participants (six groups plus pilot). Each study took around half an hour, with an additional half hour set up time.	***11 hours*** Each participant took around half an hour therefore 22 participants x 30mins =11 hours.
Time for data analysis	***~2 weeks*** 1 week for coding the video footage plus another week for coding participants’ reports	***~2 weeks*** 1 week for coding the screen recordings plus another week for coding participants’ reports	***~1 week*** for coding the participants’ reported acts

Criteria	Fire drill	VE	Talk-through
Cost to third parties	***Building occupants were evacuated (and therefore could not work) for around 20 mins***	***Occupies a large meeting room for the duration of the trial***	***Occupies a small meeting room for the duration of the trial***
Potential for re-use	***Data analysis procedures only*** i.e. re-use of the Excel spreadsheet.	***Virtual environment plus data analysis procedures*** The virtual environment could be re-used (with or without modification) for future studies.	***Data analysis procedures only*** i.e. re-use of the Excel spreadsheet.

The largest cost associated with the VE was the development time. This was still fairly rapid, at around three weeks (plus an additional 1-2 weeks for training, if required), but was significantly more than the other two methods. Second Life also has land costs in the VE. Other development kits, such as those provided with computer games, do not have the land cost (Smith and Trenholme, 2009), but there may be a fee associated with purchasing the game. They also have similar development costs (Smith and Trenholme, 2009). However, the VE had the highest potential value for re-use as the environment constructed could be used for subsequent research. The other methods were limited to re-use of the data analysis procedures (i.e. Excel spreadsheet and statistical analysis). The data analysis itself was greater with the fire drill and VE then the talk-through approach, as these involved coding and analysis of both objective (video footage/screen recordings) and subjective (participant reports) data. The talk-through approach only required coding of participants' self-reported actions.

In summary, the talk-through had required the lowest resources, followed by the fire drill, then the VE. This was specific to the scenario as tested; running a fire drill during office hours may have changed how the approaches ranked.

8.5.6 Study limitations

As for all the experimental work presented in this thesis, the results were only applicable to the scenario tested. Under different conditions, the findings may be considerably different. The results from this chapter can provide only an indication as to how the approaches may perform in different applications. Furthermore, the predictive validity of the approaches was determined by assessing their ability to predict the behaviours reported in other studies, not their ability to forecast future events.

8.6 Chapter summary

This chapter has presented a standardised comparison of three approaches to a similar scenario: a fire evacuation from a building on a university campus. The approaches included a fire drill evacuation in the real building, evacuation in a virtual environment, and analysis of participants' predictions of their evacuation actions using the talk-through approach presented in Chapter 7. Using the same scenario enabled an assessment of the approaches against the criteria for judging their quality, for this particular scenario. The assessment demonstrated that the concurrent and predictive validity of the approaches was generally acceptable under the conditions tested. Reliability was also demonstrated in the approaches, with high confidence for the fire drill method, and reliability of the frequency and sequence of acts obtained for the talk-through approach. The sensitivity of the approaches was appropriate, with all of them producing output which would be useful for an HF practitioner. The talk-through approach required the lowest resources, followed by the fire drill and finally the VE, mainly due to the effort required to build an appropriate VE for the scenario of interest.

9. General discussion

9.1 Chapter overview

The key findings from the research conducted for this thesis are discussed in this chapter, with reference to previous scientific work. The original contributions to knowledge are also presented. Recommendations are made for predicting human behaviour in emergencies, and the limitations of the research are highlighted.

9.2 Summary of research findings

The outcomes from the various studies and investigations from this research are summarised in Table 9.1. This also shows, through colour coding, analysis of the approaches against the criteria used for judging them. This analysis was a subjective assessment based on the approaches as implemented in the various studies and investigations described in the previous chapters. Reference is made to Table 9.1 throughout this chapter. Table 9.1 refers to findings from the studies and research conducted for this thesis; findings from previous research were also used to generate recommendations for predicting behaviour in emergencies in Section 9.4.

9.3 Discussion of research findings

This section initially provides an overview of the research findings. Thereafter, each of the approaches investigated is discussed individually and in comparison to previous research.

It can be seen from Table 9.1 that, overall, validity was good for the fire drill approach, and reasonable for VEs, the talk-through approach and use of literature. The use of experts, as implemented in Study 2 (Chapter 5) generally produced results with low validity for the criteria analysed. Table 9.1 shows that replicability was high for fire drills, the talk-through approach and VEs. Using literature presents a risk of variation in the predictions made by different researchers. Concerning sensitivity, VEs were considered a particularly suitable approach for HF professionals as designs can be generated and evaluated quickly. The talk-through approach and use of literature performed well against ethics and resources, although there were no major concerns with any of the methods against these criteria.

Table 9.1. Analysis of predictive approaches. Colour indicates performance against the criteria (green=good; yellow=minor issue; red=poor)

	Fire drills	VE	Talk-through	Literature[a]	Experts[b]
Validity					
Frequency of acts	Comparable with other studies	Comparable with other studies; some predicted acts specific to VEs	Predicted acts comparable with other studies	Mainly qualitative descriptions	No significant correlation with reference study.
Sequence of acts	Comparable with other studies	Comparable with other studies.	Comparable with other studies	Only a small number of qualitative descriptions	Not obtained with paired comparisons method.
Time to evacuate	Comparable with literature	Comparable with literature, but high variability in travel times	High (estimated) total evacuation time	High variability in data	Not obtained with paired comparisons method.
Perception of danger	General level comparable with a reference study	General level comparable with a reference study	Negative correlation between perception of danger and time to leave A5.	A small number of reports are available.	Not obtained with paired comparisons method.
Exit choice	Comparable with literature	Comparable with literature; some concerns about navigation in a VE	Comparable with literature; greater than expected influence (predicted) of signage.	Generally consistent reports.	Not obtained with paired comparisons method.
Reliability	High replicability of measures shown above.	Similar pattern of results on repeat applications	Medium to large effect sizes (freq & seq of acts) with reference studies on four applications	Different researchers may make different interpretations.	Not evaluated.
Sensitivity	Necessity for a physical building reduces usefulness for design investigations	Useful to HF practitioner, when used with post-trial questionnaire	Useful to HF practitioner. Limited use for investigating design detail or social aspects.	Useful to HF practitioner, if sufficiently detailed and relevant data is available.	Paired comparisons method does not generate behaviours or sequences of behaviours.
Ethics	Risk of physical injury limits scenario Deception required	Risk of simulator sickness Deception required	No distress or injury, but precautions were required.	No risk of distress or injury.	No issues experienced, but risk of experts remembering distressing event
Resources	Low/med – requires video cameras and causes disruption to workers, plus participant payment and analysis time.	Low/med – development time, participant payment and analysis. High re-use.	Low – only development of a descriptive scenario, participant payment and analysis.	Low – only time of researcher	Low – only one hour, plus development and analysis time, but requires access to experts.

[a] Findings based on the use of literature in the Standardised Comparison (Chapter 8) and predictions derived from the literature review (Appendix I/Chapter 2).
[b] Not included in standardised comparison, therefore assessment is based on Study 2 only

Fire drills

The fire drill studies conducted for this thesis investigated the utility of this approach for predicting human behaviour in emergencies. While previous studies have used similar approaches involving video footage and questionnaires to investigate human behaviour in fire drills or building evacuations (Proulx, 1995; Shields and Boyce, 2000; Gwynne et al., 2003; Xudong et al., 2009), these studies were conducted to understand the behaviours demonstrated, rather than analyse the approach taken. The fire drill component of the standardised comparison was one of the most detailed assessments of this approach. In particular this study included an investigation of the validity of the behaviours by comparing those witnessed with the behaviours predicted by other methods. Previous studies have reported the outcome of fire drill studies without giving consideration to their validity (Shields and Boyce, 2000; Olson and Regan, 2001; Gwynne et al., 2003; Xudong et al., 2009), or have investigated more specific aspects of behaviour such as merging on stairs (Boyce et al., 2009; Melly et al., 2009) or occupants' choice of lift or stair use (Heyes and Spearpoint, 2009).

The work conducted for this thesis demonstrated validity of the behaviour obtained with the fire drill approach. This conclusion was drawn from the investigations of predictive validity in the hotel fire evacuation (Study 1: Section 4.4) and concurrent and predictive validity investigations in the standardised comparison (Study 4: Chapter 8). These findings support the opinions of authors who have advocated their use for studying behaviour in emergencies (Pauls and Jones, 1980a; Proulx, 2001). Concerning the methods used within this approach, there was an association between the participants' self-reported actions and those captured on video (apart from pre-evacuation and evacuation times), which supports previous work that suggests evacuees' accounts of events are valid (Wood, 1980). Furthermore, analysis of the results from a study of a hotel fire evacuation in Saariselkä, Finland (Section 4.4) and the fire drill evacuation from a university building (Chapter 8) indicated replicability of the approach. The methods used revealed useful measures for human factors applications (Section 8.5.3), but this approach requires a physical building to be implemented; thus its use for investigating design alternatives is limited. Ethics considerations, including the risk of injury and distress were identified with this approach as they limited the scenario. These concerns have also been identified by other authors (Edelman et al., 1980; Muhdi et al., 2006). Finally, the approach requires low resources overall, but echoing concerns by Pauls and Jones (1980a), it may cause disruption to other building occupants.

Virtual environments

While studies have been conducted to investigate the use of virtual environments for predicting evacuation behaviour, none had previously investigated all the facets studied for this research. For example, no inferential statistical investigations on exit times from buildings have been presented (Mantovani et al., 2001; Gamberini et al., 2003; Mol et al., 2008; Smith and Trenholme, 2009). Also, the research for this thesis used a multi-player system with voice communications, which has not been used in previous fire drill studies. Finally, and as mentioned above, previous work has never reported on the various criteria for judging the quality of an approach in the detail reported in this thesis.

There were similarities in some of the results of the VE study conducted as part of the standardised comparison and those from prior investigations. For example, Smith and Trenholme (2009) showed a longer time to evacuate in the VE than in real life. The VE study in the standardised comparison also showed a longer evacuation time, although not significantly different. However, in the standardised comparison, variability in the travel times in the VE was greater than in the fire drill, as also reported by Yogesh (2009). Smith and Trenholme (2009) raised concerns about the face validity of the behaviour of participants in VEs, such as opening doors with smoke coming underneath them. Mantovani et al. (2001) reported problems in a VE study such as participants overshooting a desired turn, and colliding with objects. In the standardised comparison several participants re-entered the building, which may have been due to a lack of perceived danger in the VE. Concerns were also raised in the standardised comparison about participants' navigation in the VE, which seemed affected by a limited field of view. Certain guidelines have proven effective at reducing user disorientation, such as positioning objects on walls to indicate an area exists off-screen when partially obscured by nearer walls and features in the VE (Smith and Marsh, 2004). In many instances the complexity of the VE used in the standardised comparison had the same effect. For example Figure 8.35 shows that the staircase indicated an area existed to the left of the visible foyer area. However, these guidelines may be worth implementing more systematically in further work. Despite the concerns raised, overall the frequency and sequence of acts in the VE showed concurrent validity with the results of the fire drill, talk-through approach and literature analysis.

Using any VE poses a risk of simulator sickness (Nichols and Patel, 2002), and while no participants were affected in Study 4, a mild case was reported in the development of the emergency training simulator (Section 6.3). Precautions were therefore required to avoid any unpleasant feelings. Concerning the resources, Smith and Trenholme (2009) reported constructing their VE within three weeks – a similar time was taken to develop the VE for the standardised comparison. While this was greater than the resources used for the fire drill or talk-through, the use of virtual buildings is in fact one of the main advantages of this method. Design alternatives can be developed and compared quickly, in far less time than required to generate and evaluate physical alternatives in a real environment. Furthermore, the potential for re-use of the VE is high.

Talk-through approach

The talk-through approach was developed in recognition of the difficulty of ethically studying behaviour in emergency situations, a problem which has been cited by other authors (Edelman et al., 1980; Muir, 1996; Galea et al., 2007; Gershon et al., 2007; Kobes et al., 2009). Participant predictions were seen as an approach for obtaining the anticipated behaviours without any risk of physical injury. Thus, while the talk-through method has been used in other applications (Kirwan and Ainsworth, 1992), this was the first use for emergency situations, and in conjunction with sequence analysis (Bakeman and Gottman, 1986). This was one of the main contributions of this research – the development and assessment of a new approach for predicting human behaviour in emergencies.

The use of sequential analysis enabled investigation of the sequences of the predicted actions. Sequences of actions have rarely been reported in detail in studies of behaviour in emergencies, yet they have been recognised as critical to the outcome of an event (Canter et al., 1980). Despite previous concerns about the validity of participant predictions (Heyes and Spearpoint, 2009), the talk-through approach in the work conducted for this thesis demonstrated concurrent validity for the frequencies and sequences of acts with the results of the fire drill and VE in Study 4. It also demonstrated comparability of the behavioural predictions with a reference study of human behaviour in real fire (Canter et al., 1980). Furthermore, the technique demonstrated a negative correlation between perception of danger and time to leave the building, which was reported in other literature surrounding emergency scenarios (Galea et al., 2009). This finding was consistent with research in risk theory which demonstrates an influence of risk perception on decision making (Slovic, 2002; Burns, 2007). Some of the potential limitations of the talk-through approach are that it may not be sufficiently sensitive to investigate detail, such as signage design as indicated by the higher than expected influence of signage on participants' hypothetical actions. Furthermore, it was difficult to incorporate social interactions other than those described in the initial scenario. It also produced a high (estimated) total evacuation time when compared to values in the scientific literature (Figure 8.24).

The approach was found to require even less resources than the VE study, amounting only to participant payment and the researcher's time to develop the scenario and analyse the results. The talk-through approach proved reliable, demonstrating associations between the frequencies and sequences of (predicted) actions with a reference study on four separate occasions.

Literature

The use of literature for behavioural prediction can also be discussed, although this was mainly used for further evaluation of the predictive validity of the other approaches, rather than to generate predictions itself. Chapter 2 revealed that behavioural predictions based on the literature are faced with issues, for example, limited quantitative data regarding behaviour other than evacuation time (Proulx, 1993; Pan et al., 2006; Proulx, 2007; Kuligowski, 2009; Tubbs and Meacham, 2009). Despite being 30 years old, the reference study by Canter et al. (1980) was one of the most detailed, quantified studies of behaviour in emergencies available in the literature, hence its extended use for validation in this thesis. Using literature to validate the outcome of the other approaches revealed a lack of data on which to make the comparisons, a problem which has also been found in previous investigations (Mawson, 2005; Pan et al., 2006; Proulx et al., 2006). Chapter 2 also raised concerns about possible variations in interpreting the literature by different researchers.

Despite these concerns the approach yielded useful data, where sufficiently relevant and detailed data were found. As no participants were used there were no ethical considerations, and resources were limited to the time of the researcher.

Expert predictions

Finally, the experts study was the first investigation into the use of the paired comparisons technique for predicting the likelihood of certain behaviours in an emergency. Some previous predictions made using expert knowledge had relied on techniques similar to hierarchical task analysis which generally work well in procedural applications or for investigating specific aspects of an emergency (Groner, 2009; Zachary Au, 2009). However, they may not be capable of providing a more general overview of what actions people will take in situations which are not dominated by procedure, such as those typically experienced by building occupants in an emergency. Dombroski et al. (2006) revealed wide variability in experts' quantified predictions of compliance with an evacuation order, and also gave indications of contradictions with other predictions such as those made by Gershon et al. (2007). Thus, concerns were identified with the predictive validity of the expert predictions, concerns which were also witnessed in the paired comparisons study (Study 2; Chapter 5) in this thesis. Dombroski et al. (2006) also cited the need for a detailed reference scenario, again something found in Study 2, with experts reporting difficulty in making predictions based on the very brief description of the emergency situation provided. A further limitation was that the approach used in Study 2 did not in itself generate the behaviours of interest; these had to be identified through another method. As a result of these issues, this approach was not developed further.

9.4 Recommendations for predicting human behaviour in emergencies

To achieve Aim 4 “develop guidance for HF practitioners”, it was necessary to develop from the research recommendations for HF professionals involved with behavioural prediction in emergency situations. These recommendations were based on the research work conducted for this thesis, but also draw upon the previous scientific literature reviewed in Chapter 2. This section focuses on recommendations for applying the approaches; recommendations for future research are made in Section 10.4.

Three tables are presented. The first (Table 9.2) presents the recommended applications for each predictive approach, based on the purpose of the prediction. This was derived from consideration of all the criteria used to evaluate the approaches, but in particular *sensitivity*. Table 9.2 can guide users to the most suitable approach(es) for their investigation. The second table (Table 9.3) presents recommended methods and measures/outcomes associated with each approach. This was based upon consideration of the criteria *validity* and *reliability*, both from the research conducted for this thesis, and prior research. Table 9.2 and Table 9.3 can therefore be used to identify approaches which are suitable for a behavioural prediction, after which the analyst can select methods which can be applied to obtain the desired data. The final table (Table 9.4) will help the user consider certain constraints (for example resource or ethical considerations) which may affect the choice of approach. These tables are intended as a guide for approach/method selection, but it is recognised that it may be difficult to satisfy all criteria, as also reported by Wilson (2005).

As mentioned in Section 1.7, consideration should be made to the limits of any behavioural prediction. Users should be aware of the data and conditions from which the recommendations were derived and in particular the limits of the predictive validity of the approaches.

9.4.1 Recommended applications for various approaches

Recommendations for suitable approaches for predicting behaviour in a variety of applications are provided in Table 9.2. The intention is that a user would examine this table to first select their approach for making a behavioural prediction, and then select methods to obtain their desired measures/outcome using Table 9.3.

The recommendations in Table 9.2 include approaches for generating behavioural predictions during different stages in the development lifecycle. These have been categorised as *Early* (during the concept phase), *Mid* (as the design becomes refined and detailed CAD models become available) and *Late* (any time after the availability of physical properties representing the design intent).

Participant predictions, and in particular the talk-through approach is recommended during the early design phase, based on the low resources required for this approach, which is still able to give an indication of likely behaviours as seen in Chapters 7 and 8. The use of experts and literature may be suitable, although this will be limited to the specific methods and measures recommended in Table 9.3. It may be hard to develop a VE or use simulation at this stage, given the available data. Fire drills are not suitable during this phase, as they require a physical property. Investigating behaviour using reports by survivors is recommended as a rich source of data, but gaining access to survivors may be difficult.

Moving to the mid design phase, VEs are recommended as the information will now be available to generate the environment of interest. Conversely, the benefits of using participant predictions are likely to be surpassed by VEs or simulations at this stage due to the limits of the talk-through approach for investigating detail. Simulation would be particularly useful if investigating a scenario involving large crowds, or running several instances of an emergency to understand the distribution of evacuation times or range of possible outcomes (Gwynne et al., 1999).

Once physical properties become available, fire drills are a suitable approach for predicting behaviour (Chapter 8). However, VEs and simulation are also recommended, particularly if the scenario presents a risk of physical injury to participants (Chapter 8; Gwynne et al., 1999).

In addition to design applications, behavioural predictions can also be useful for informing emergency response procedures and training (Meguro et al., 1998; Perry and Lindell, 2003; Kanno et al., 2006; Lawson et al., 2007b; Chittaro and Ranon, 2009; Hsiung et al., 2009; Kinsey et al., 2009a; Tubbs and Meacham, 2009). For developing emergency response procedures, fire drills, VEs, and simulation could all generate useful behavioural information, as shown in Table 9.2. Participant predictions, as used in the talk-through approach, are limited in their ability to consider social behaviours, such as the response to an authority figure (Chapter 8), and therefore are likely to be of less use. Similarly, the use of experts is only recommended for specific measures, and the necessary information to make a prediction may not be available in the literature. Using reports by survivors will be limited by the available access to survivors. For behavioural predictions which are used to inform training programmes, all approaches are suitable, although the cautionary notes for experts, use of literature and reports by survivors above still applies.

Table 9.2. Recommended applications for various approaches for predicting behaviour in emergencies. Yellow indicates caution; red cells indicate approaches which are not recommended.

Approach	Suitability for predicting behaviour to inform various stages of a design			Informing emergency response procedures	Generating behavioural predictions to inform training programmes for emergency response
	Early: Concept phase	**Mid:** CAD models are available	**Late:** Full-size physical properties are available		
Fire drills	Not recommended – approach requires a physical property		Recommended – can provide useful data (see Table 9.3)	Recommended – identified behaviours could be incorporated in response procedures	Recommended for informing training programmes
VE	Not recommended – insufficient detail likely to be available to generate VE	Recommended for measures shown in Table 9.3	Recommended if conditions of emergency situation would otherwise present a risk of physical injury to participants	Recommended – emergency procedures could be investigated in the VE	Recommended for informing training programmes
Participant predictions	Recommended – low resources, yet can provide an indication of expected behaviours	More likely to benefit from VE or simulation	More likely to benefit from fire drills, VE or simulation	Recommended for identifying likely behaviours, but predicting social interaction could be difficult	Recommended for informing training programmes
Experts	Recommended – but limited to measures in Table 9.3				
Literature	Recommended - but measures will be limited to mainly qualitative behaviours or models of behaviour (Table 9.3)				
Simulation	Not recommended – insufficient detail likely to be available to generate a simulation	Recommended - particularly useful for large crowds, or running repeat models (Gwynne et al., 1999)	Recommended to investigate design changes if conditions of emergency situation would present a risk of physical injury or distress to real participants (Gwynne et al., 1999)	Recommended – can be used to investigate alternative response procedures (e.g. Hsiung et al., 2009; Kinsey et al., 2009a)	Recommended – simulation tools can be used to investigate possible outcomes of an evacuation (Kuligowski and Peacock, 2005)
Reports by survivors	Recommended – but gaining access to survivors may be difficult (e.g. Galea et al., 2009)				

9.4.2 Recommended methods and measures for predicting behaviour in emergencies

Table 9.3 shows the recommended methods and measures obtainable for each of the reviewed approaches. The information for this table was derived both from the research conducted for this thesis and also previous literature, from which the most commonly used methods and measures are reported. It was generated mainly with consideration of the criteria *validity* and *reliability*.

This table can guide users in selecting methods to obtain the desired measures/outcomes shown in the final column. It includes recommended data collection and data analysis methods and measures for each approach. Caution may be required in certain instances, for either methods or measures, as indicated by yellow cells. Methods and measures which are not recommended for a particular approach are shown in red. Footnotes are used to explain any concerns or issues. Again, this categorisation was based on the work conducted from this thesis, but also from the review of previous literature in Chapters 2 and 3, during which problems with methods and measures were identified.

A variety of measures can be obtained using fire drills, as used in Study 4 (Chapter 8) and reported in Section 2.6. These can be obtained through the use of observation (video) and post-event questionnaires. However, as mentioned in the standardised comparison, a combination of these methods is recommended: the video provides objectivity (e.g. evacuation times for which participants' estimates were inaccurate) and the post-event questionnaire captures acts which are too detailed to be captured on video. The post-event questionnaire is also required to capture cognitive aspects and participants' perceptions and influences on their evacuation behaviour. However, it is not recommended for self-reported evacuation times due to the lack of correlation with objective measures seen in Study 4.

A combination of screen recordings (or automated event recording) and post-event questionnaire is also recommended when predicting behaviour using VEs. With this approach, caution must be exerted when considering acts which are specific to the technology used such as navigation difficulties (see Section 8.5). Concerns were also raised over exit choice in the VE (Section 8.5; Mantovani et al., 2001) and travel time (Study 4; Yogesh, 2009).

Table 9.3. Recommended methods and measures for predicting behaviour in emergencies. Yellow indicates caution; red cells indicate measures which are not recommended.

Approach	Methods for data collection	Methods for data analysis	Measures/outcome
Fire drills	Observation (video)	Behavioural coding	Frequencies of acts (Study 4) Number and type of acts performed prior to evacuation (Gwynne et al., 2003)
		Sequence analysis	Act transitions (Study 4)
		Timeline analysis	Evacuation times (Study 4; Proulx, 1995; Muir et al., 1996; Shields and Boyce, 2000; Olsson and Regan, 2001; Gwynne et al., 2003; Jong-Hoon et al., 2009; Xudong et al., 2009) Pre-movement times (Study 4; Proulx, 1995; Shields and Boyce, 2000; Olsson and Regan, 2001; Gwynne et al., 2003) Movement speed (Study 4; Proulx, 1995; Kinsey et al., 2009b) Merging behaviour in stairwells (Boyce et al., 2009; Melly et al., 2009)
		Checklist	Exit usage (Study 4; Proulx, 1995; Shields and Boyce, 2000; Kobes et al., 2009; Xudong et al., 2009) Route choice (Kinsey et al., 2009b)
	Questionnaire	Behavioural coding	Frequencies of acts (Study 4) Primary acts (Shields and Boyce, 2000; Xudong et al., 2009)
		Sequence analysis	Act transitions (Study 4)
		Checklists/content analysis	Pre-event activity (Shields and Boyce, 2000) Frequencies of pre-evacuation acts (Proulx, 1995) Commitment to pre-event activity (Shields and Boyce, 2000) Perceived emergency cues (Shields and Boyce, 2000; Xudong et al., 2009) Reasons for exit choice (Shields and Boyce, 2000, Xudong et al., 2009)
		Statistical analysis	Estimates of evacuation time (Study 4) [a] Estimates of pre-movement time (Study 4; Xudong et al., 2009) [a]
		Statistical analysis of rating scales	Perception of danger (Study 4; Kobes et al., 2009) Influences on exit choice (Study 4)

[a] No correlation was found between participants' reported evacuation times and observed evacuation times in Study 4 (Chapter 8)

Approach	Methods for data collection	Methods for data analysis	Measures/outcome
VE	Observation (screen recording)	Behavioural coding	Frequencies of acts (Study 4)[a] Response behaviours (Gamberini et al., 2003; Smith and Trenholme, 2009)[a]
		Sequence analysis	Act transitions (Study 4) [a]
		Time line analysis	Evacuation times (Study 4; Shih et al., 2000; Mantovani et al., 2001; Smith and Trenholme, 2009) Pre-movement times (Study 4)
			Travel time (Study 4; Yogesh, 2009)[b]
		Checklist	Exit usage (Study 4; Shih et al., 2000; Mantovani et al., 2001)[c]
	Automated event recording	Statistical analysis	Evacuation time (Gamberini et al., 2003; Mol et al., 2008)
	Questionnaire	Behavioural coding of retrospective walk-through	Frequencies of acts (Study 4)[a]
		Sequence analysis	Act transitions (Study 4)[a]
		Statistical analysis	Estimates of evacuation time (Study 4) Estimates of pre-movement time (Study 4)
		Statistical analysis of rating scales	Perception of danger (Study 4) Influences on exit choice (Study 4)
Participant predictions	Talk-through with questionnaire/interview	Behavioural coding	Frequencies of acts (Study 4)
		Sequence analysis	Act transitions (Study 4)
		Statistical analysis of time estimates	Evacuation time (Study 4)[d]
			Pre-movement time (Study 4)
		Statistical analysis of rating scales	Perception of danger (Study 4)
			Influences on exit choice (Study 4)[e]
		Checklist	Exit usage (Study 4)
	Vignettes with questionnaire	Statistical analysis	Lift vs. stair use (Heyes and Spearpoint, 2009)[f]
	Internet survey	Statistical analysis	Lift vs. stair use (Heyes and Spearpoint, 2009)[f]

[a] Caution is required as acts may be specific to VEs, e.g. getting lost, navigation difficulties

[b] Greater variability in travel times (i.e. time from starting evacuation to leaving the building) in VEs than in fire drill evacuations (Study 4; Yogesh, 2009)

[c] Study 4 (Chapter 8) identified concerns about exit choice in the VE

[d] High total evacuation time was reported in Study 4 (Chapter 8)

[e] Greater than expected influence of signace (Study 4: Chapter 8)

[f] High variability was reported in the results of this study: caution required

Approach	Methods for data collection	Methods for data analysis	Measures/outcome
	Post-event survey following similar incident	Statistical analysis	Lift vs. stair use (Heyes and Spearpoint, 2009)[a]
Experts	Abstraction hierarchy	Analysis of situation awareness requirements	Identification of the information required for situation awareness in an emergency (Groner, 2009)
	Hierarchical task analysis	HAZOP	Identification of tasks with potential for human error (Zachary Au, 2009) The impact of behavioural issues on emergency response processes (Zachary Au, 2009)
	Quantitative estimates	Statistical analysis	Public compliance (percentage) with an evacuation order (Dombroski, 2009)[a]
	Paired comparisons	Statistical analysis	Likelihood of acts (Study 2)[b]
Literature	Review of scientific literature	Analysis of reported behaviours	Qualitative descriptions of human behaviour in emergencies (Appendix I; Proulx, 2001; Purser and Bensilum, 2001; Glass and Schoch-Spana, 2002; Mawson, 2005) Qualitative predictions of human and social behaviour during building egress (Pan et al., 2006) Qualitative descriptions of perception of danger (Study 4) Qualitative descriptions of preferred exit usage (Study 4; Purser and Bensilum, 2001) Descriptions of the behavioural response to alarms (Proulx, 2007)
			Qualitative descriptions of sequences of behaviour (Study 4)[c]
		Statistical analysis	Evacuation times (Study 4; Purser and Bensilum, 2001)[d] Pre-movement times (Study 4; Proulx, 2001; Purser and Bensilum, 2001)[d] Travel times (Purser and Bensilum, 2001)
		Model/theory building	Model of influential factors during perception and interpretation of a fire situation (Kuligowski, 2009) Stress model (Proulx, 1993) Model of decision making under time pressure and stress (Ozel, 2001) Review of panic theory (Pauls and Jones, 1980b; Sime, 1995; Proulx, 2001) Social attachment model (Mawson, 2005) Qualitative behavioural predictions based on Emergent Norm Theory (Aguirre et al., 1998) Model of behavioural response to terrorism (Fischer, 2002; Fischer 2003) Model of the behavioural response to a nuclear incident (Kanno et al., 2006)
			Analytic model of required safe egress times (Proulx et al., 2006)[e]

[a] High variability was reported in the results of these studies: caution required.

[b] Behaviours did not correlate with those from real emergencies, therefore this measure is not recommended; caution is also required with re-use of this method

[c] Few data sources on sequential behaviour found in Study 4 (Chapter 8)

[d] Generalisability of quantitative data derived from the literature to other situations is unclear

[e] Proulx et al. (2006) recommend caution in the use of the model.

Approach	Methods for data collection	Methods for data analysis	Measures/outcome
Simulation	Behavioural (and partial behaviour) models	Analysis of model output	Visualisation of the evacuation (Kuligowski, 2003; Kuligowski and Peacock, 2005) Exit usage (Kuligowski, 2003) Emergent behaviour e.g. herding, queuing (Pan et al., 2006) Number of occupants trapped (Kuligowski, 2003) Outcome of different evacuation strategies (Hsiung et al., 2009; Kinsey et al., 2009a) Identification of bottlenecks (Kuligowski, 2003)
			Evacuation time (Gwynne et al., 1999; Kuligowski, 2003; Olsson and Regan, 2001; Ko et al., 2007; Christoffersen and Soderlind, 2009; Hsiung et al., 2009; Kinsey et al., 2009a; Purser and Boyce, 2009)[a] Pre-movement times (Olsson and Regan, 2001; Kuligowski, 2003)[a] Movement speed/flow rates (Olsson and Regan, 2001; Ko et al., 2007; Purser and Boyce, 2009; Xu and Song, 2009)
	Movement models[b]	Analysis of model output	Visualisation of the evacuation (Kuligowski, 2003; Kuligowski and Peacock, 2005) Evacuation time (Gwynne et al., 1999; Kuligowski, 2003) Pre-movement times (Kuligowski, 2003) Travel times (Kuligowski, 2003) Exit usage (Kuligowski, 2003) Identification of bottlenecks (Kuligowski, 2003)
Reports by survivors	Interviews[c]	Data coding and analysis	Frequencies of initial response acts (Edelman et al., 1980; McConnell et al., 2009) Quantitative data on perceived cues to emergency (Edelman et al., 1980; McConnell et al., 2009) Factors facilitating or hindering evacuation (Gershon et al., 2007; Averill et al., 2009) Frequency of acts (Canter et al., 1980) Frequency and type of acts undertaken prior to evacuation (Galea et al., 2009) Number of stoppages (Galea et al., 2009) Exit usage (Edelman et al., 1980)
		Statistical analysis	Causal models of delay to initiate evacuation (Averill et al., 2009) Correlation between number of acts completed & delay to evacuation (Galea et al., 2009) Travel speeds (Averill et al., 2009; Galea et al., 2009) Response times (Galea et al., 2009; McConnell et al., 2009)
		Statistical analysis of rating scales	Perceived risk (Galea et al., 2009; McConnell et al., 2009)
		Sequence analysis	Act transitions (Canter et al., 1980)

[a] Differences have been found in evacuation times when applying different simulation models to the same scenario, therefore caution is required with these measures
[b] Movement models do not incorporate human behaviour: this limits their face validity and therefore caution is required in their use
[c] Concerns have been raised about the accuracy of survivors' recall of events (Edelman et al., 1980; Fahy and Proulx, 2005; Proulx and Reid, 2006), therefore caution is required.

Approach	Methods for data collection	Methods for data analysis	Measures/outcome
		Model building	Prevalence of co-operation rather than panic (Drury et al., 2006) Generic models of behavioural sequences (Canter et al., 1980)
	Questionnaires[a]	Content analysis	Factors facilitating or hindering evacuation (Gershon 2009) Frequency of response acts (Proulx and Reid, 2006; Jeon and Hong, 2009) Motivations for evacuating (Proulx and Reid, 2006; Jeon and Hong, 2009) Perception of situation/risk (Proulx and Reid, 2006; Jeon and Hong, 2009) Exit choice (Proulx and Reid, 2006; Jeon and Hong, 2009)
		Statistical analysis	Evacuation time (Proulx and Reid, 2006) Pre-movement time (Proulx and Reid, 2006)
		Model building	Variables associated with first actions, increased evacuation behaviour and movement through smoke (Wood, 1980)

[a] Concerns have been raised about the accuracy of survivors' recall of events (Edelman et al., 1980; Fahy and Proulx, 2005; Proulx and Reid, 2006), therefore caution is required.

With participant predictions, the talk-through method is recommended for a variety of measures, including frequencies of acts, act transitions, exit usage, perception of danger and influences on exit choice. These were deemed suitable based on the development of the talk-through approach (Chapter 7) and the Standardised Comparison (Chapter 8). Caution is required with predicted evacuation times, due to the high total evacuation time reported in Study 4. The approach relies mainly upon questionnaires for data collection, although questions may be implemented in interview format. The vignettes and survey methods from Heyes and Spearpoint (2009) require caution due to variability in the results.

With the use of experts, task analysis based approaches for data analysis are recommended, based on the studies by Groner (2009) and Zachary Au (2009), for revealing situation awareness requirements and identifying the potential for human error. Using the paired comparisons method to understand the likelihood of behaviours as used in Study 2 (Chapter 5) is not recommended. The paired comparisons technique may be appropriate in a different application, although caution is required until proven against the criteria for judging a method.

When using literature analysis, qualitative predictions of behaviour may be possible, but information on sequences of behaviour may be lacking, as found during the standardised comparison investigation. Lack of data may also be an issue with other investigations based on literature analysis. Evacuation times derived from existing emergencies may be limited to the conditions from which they were obtained; their generalisability is unclear.

Concerning computer simulation, behavioural models have been recommended over movement models, as the latter do not attempt to model occupant behaviour. However, concerns over the validation of the simulation models (Gwynne et al., 1999; Kuligowski, 2003) urge caution in their use, particularly when considering evacuation times.

Finally, reports by survivors can be used to obtain rich data on behaviour in emergencies, through interviews and questionnaires. However, uncertainty about the accuracy of survivors' recall of events, in particular quantitative aspects (e.g. Proulx and Reid, 2006) emphasise caution in their use.

9.4.3 Recommendations to address constraints affecting prediction

Table 9.4 presents recommended approaches given constraints which may affect the predictions. These constraints were identified during the research conducted within this thesis, plus those reported in previous investigations. They relate mainly to resource and ethics considerations.

Considering first the constraint of a low budget, VEs, participant predictions, experts and use of literature are recommended, being low cost approaches to implement. While fire drills themselves are relatively inexpensive, there may be a cost associated with disruption to building occupants. Simulation will also have an associated cost – either to purchase the tool, or to hire an expert to conduct the investigation (Kuligowski and Peacock, 2005). Approaching behavioural prediction using reports by survivors was reported to have high associated resources in Section 2.2.

Most of the approaches are suitable given limited time to conduct the prediction, with the exception of reports by survivors; gaining access to survivors can take several months, plus the time taken for transcribing interview data can be considerable (Galea et al., 2009).

Considering potential re-use from the approaches, VEs have high potential for re-use, as the environment can be modified to investigate design changes and re-used. Literature and reports by survivors may be used to identify high-level generalisable models of behaviour which may be re-used across various emergencies. Simulation tools can be run many times with little additional resource. With participant predictions, only the analysis method (i.e. spreadsheet) can be re-used, although this is the greatest resource associated with the approach. Fire drills have limited re-use; a new drill is likely to be required for each investigation, or following any changes to the scenario. Using experts is also likely to have limited re-use – a new investigation would be required for each investigation.

If the scenario under investigation presents a risk of physical danger to the people involved, participant predictions, expert predictions, literature and simulation can all be used without putting participants at any actual risk of injury. Fire drills are not recommended, as it is necessary to protect participants from harm. VEs can cause simulator sickness if the scenario involves rapid movement, and measures should be put in place to protect against this. Caution should be exerted with reports by survivors, as asking them about an event in which they were injured may be upsetting.

Considering scenarios which present a risk of distress, literature and simulation are recommended, as neither of these approaches places participants within the distressing situation. Participant predictions and the use of experts require caution, as while the participants or experts are not required to experience the scenario, they will be required to think about it. Using survivors to report on distressing events may also be traumatic for them. Fire drills and VEs are not recommended, as these would expose the participant to the distressing scenario. While it is only a virtualised representation in VEs, this may cause distress and therefore should be avoided.

Table 9.4. Recommended approaches given constraints on the prediction. Yellow indicates caution; red cells indicate circumstances in which the approaches are not recommended.

Approach	Constraints on prediction				
	Low budget	**Limited time**	**Re-use is required**	**Emergency situation of interest presents a risk of physical danger (e.g. smoke, fire, CBRN, crowd surges, rapid movement)**	**Emergency situation of interest has risk of causing distress to people involved**
Fire drills	Recommended, but may cause disruption to building occupants (Study 4; Pauls and Jones, 1980a)	Recommended – can be implemented quickly (Study 4)	Likely to require new drill for each scenario or investigation.	Not recommended – risk of physical injury (Section 2.6; Study 4)	Not recommended – risk of distress to drill participants (e.g. Brooks et al., 2001)
VE	Recommended – low cost to implement (Study 4; Smith and Trenholme, 2009)	Recommended – reasonably quick to implement, although VE must be developed (Study 4; Smith and Trenholme, 2009)	VE can be modified for design alternatives, and re-used indefinitely (Study 4)	VE can cause simulator sickness if scenario involves rapid movement through a scene – measures required to protect against this (Chapter 6; Study 4)	Not recommended – risk of causing participants distress
Participant predictions	Recommended – low cost to implement (e.g. Study 4)	Recommended – can be implemented quickly (Study 4)	Data analysis approach (main resource) can be re-used	Recommended – participant is not put at risk of physical injury (Studies 3a-3c; Study 4)	Caution required – participants will be asked to consider distressing event (Studies 3a-3c)
Experts	Recommended – low cost to implement (e.g. Study 2)	Recommended – can be implemented quickly if access to experts can be obtained (Study 2)	Likely to require new investigation for each emergency scenario	Recommended – no participants are exposed to a risk of physical danger	Caution required – risk of distress if experts are required to consider traumatic event
Literature	Recommended – low cost to implement (Chapter 2; Study 4)	Recommended – can be implemented quickly (Chapter 2; Study 4)	Generic models of behaviour likely to have value for re-use (Section 2.3)	Predictions can be drawn from literature: no participants involved, therefore no risk of injury	Predictions can be drawn from literature: no participants involved, therefore no risk of distress
Simulation	Simulation tools likely to have an associated cost of purchase or consultancy fees (Kuligowski and Peacock, 2005)	Simulation exercises can be conducted quickly (based on Kuligowski, 2003)	Simulations can be run many times (Gwynne et al., 1999)	Can investigate effects of environmental hazards, even "fatalities" in agents, without any risk to real participants (Gwynne et al., 1999)	Can investigate distressing events, without any risk to real participants (Gwynne et al., 1999)
Reports by survivors	Administering and analysing interview data is resource intensive (Gershon et al., 2007; Galea et al., 2009)	Gaining access to survivors can take considerable logistical preparations (Galea et al., 2009)	Some models (e.g. Drury et al., 2006) may be generalisable across multiple emergencies	Caution required: survivors may have been injured in emergency event of interest; discussing event may prove traumatic (Gershon et al., 2007; Gershon, 2009)	Caution required: survivors' experiences may have been traumatic (e.g. Galea et al., 2007)

9.5 Novel contribution to knowledge

The main contribution of this thesis was illustrated in Table 9.1 – an evaluation of approaches for predicting human behaviour in emergencies. While previous studies from the scientific literature (Chapter 2) have reported evaluations of approaches for predicting behaviour (Gwynne et al., 1999; Kuligowski, 2003; Santos and Aguirre, 2004; Kuligowski and Peacock, 2005; Smith and Trenholme, 2009), none provided as detailed and systematic evaluation of the approaches as investigated for this thesis. In particular, while simulation has received relatively in-depth analysis (Gwynne et al., 1999; Kuligowski, 2003; Santos and Aguirre, 2004), none of the other approaches had previously been evaluated against all the criteria in this thesis, and with consideration for emergency situations. Furthermore, Chapter 2 revealed that concerns raised about the lack of validation in simulation tools (Munley et al., 1996; Gwynne et al., 1999; Shields and Proulx, 2000; Kuligowski, 2003) also applied to other predictive approaches. The research in this thesis addressed these concerns with the investigation of several different aspects of validation for each of the approaches, including frequency and sequence of acts, time to evacuate, perception of danger and exit choice. These aspects were selected based on consideration of the measures reported (although not collectively in any one study) in previous research in human behaviour in emergencies (Canter et al., 1980; Edelman et al., 1980; Proulx, 1995; Shields and Boyce, 2000; Olsson and Regan, 2001; Gwynne et al., 2003; Proulx and Reid, 2006; Heyes and Spearpoint, 2009; McConnell et al., 2009).

The development of a new approach for predicting behaviour in emergencies (talk-through approach), was a further novel contribution. The review of previous research in Chapter 2 identified that participant predictions had received limited attention and analysis for emergency scenarios. Thus, the development of this new approach provides an addition to knowledge.

The standardised comparison (Chapter 8) provides the first systematic comparison of fire drills, VEs, the new talk-through approach and the use of literature. Chapter 2 demonstrated that while previous research has included comparisons of different approaches, such as simulation and fire drills, this was the first comparison of the approaches described above, and using established criteria to evaluate their quality (Olsson and Regan, 2001; Kuligowski, 2003; Gwynne et al., 2005; Christoffersen and Soderlind, 2009). This systematic approach enabled a detailed comparison of their performance against the criteria.

The recommendations in Section 9.4 are another important original contribution to knowledge. Recommendations for HF professionals working in emergency preparedness were not previously available in the scientific literature, as identified during the project work which gave rise to the research conducted for this thesis (Section 1.2). However, from this research, and through analysis of previous work, recommendations have been made for selecting suitable approaches for various HF applications. Recommendations for the choice of methods and techniques for the desired data/outcome are also provided. Finally, recommendations were made based on constraints affecting the prediction. Thus, this collection of recommendations is a novel and valuable contribution which can inform HF practitioners when making predictions of human behaviour in emergencies.

9.6 Limitations of the research

This section addresses the limitations of the research and methodological approach. Recommendations to address these in further work are made in Section 10.4.

The most important limitation of this research was that the predictive validity of the approaches was determined by their ability to predict data reported in reference studies, rather than their ability to forecast behaviour in future events. This distinction is important as it could be potentially dangerous to assume greater confidence in the behavioural predictions than that supported by the research work conducted for this thesis.

A related limitation was that even with significant correlations between the behaviours predicted in this study and the reference studies, the frequency or sequences of any one of the behaviours may be considerably different. This was illustrated in the third development study (3c) of the talk-through approach (Section 7.4). A significant correlation and large effect size were found between the frequency of predicted acts and those in the reference study (Canter et al., 1980) (r_s=0.572, N=51, p<0.001). Yet, act 24 *Leave immediate area* accounted for 11% of the total number of predicted acts, and only 4% of the total number of acts in the reference study. Thus, it is recommended that the predictions are considered indicative of the behaviours in the reference data set, and that the individual values are treated with caution. Furthermore, the absence of normality in the majority of the data sets led to the use of non-parametric testing, which evaluates the rank order of the entities, rather than the values. Thus, the testing was limited to the order of the behaviours (predicted and actual) rather than their magnitude.

For both of the above limitations, the main implication for practitioners is that the limits of the predictive approaches must be recognised during subsequent use. This cautionary note has been mentioned several times throughout the thesis.

Other limitations arose through the limited availability of data sources for validation. The study by Canter et al. (1980) was found to be one of the most detailed, quantified investigations of human behaviour in fire, and hence the extended use of this for validation. However, the descriptions of the dwellings investigated by Canter et al. (1980) were limited, with little more detail than “domestic, multiple occupancy and hospital fires”. Thus, the specific contexts from which the behaviours were captured were unknown. Furthermore, Canter et al. (1980) investigated behaviours using in part interviews with survivors of fires. These could have been subject to recall issues (Edelman et al., 1980; Aguirre et al., 1998; Gershon et al., 2007). To address the problems of limited validation studies, attempts were made by the author of this thesis to gain access to more detailed incident data through contacts in the emergency services and by applying to a database of survivors accounts from the World Trade Center (HEED)(Galea et al., 2009), but these attempts were unsuccessful. While these would have improved the quality of the validation studies during phase 1, the problem was somewhat overcome during the standardised comparison, as the process of triangulation was used to create internal data sources for validation.

Attempts were made to avoid experimenter bias in this research, through experimental design, reviewing the approaches taken with other researchers and an inter-rater reliability check. However, the behavioural predictions made were not “blind”; by necessity the author knew the reference study well and therefore unknown sources of experimenter bias may have influenced the outcomes. Furthermore, the experimental work reported in this thesis was conducted entirely by the author, or with significant input from the author. Recommendations to address this concern, including application of the approaches by other researchers to ensure replicability, are made in Section 10.4.

Considering the broad range of possible emergency situations (e.g. CBRN incident, natural disaster, battlefield crises, and train or aeroplane crash) the approaches have been investigated in a relatively limited number of scenarios. The research focussed on building evacuations and fire drill studies, driven in part by the greater availability of data for validation in this area. Consequently, this limits the generalisability of the findings; to be truly useful to HF practitioners, the approaches need to be investigated in a much broader range of emergency situations.

A further limitation which may be addressed in future work was that each approach was only investigated for a small number of instances. Gwynne et al. (1999) reported that a distribution of evacuation times would be expected on repeat applications, something which can realistically be achieved using simulation tools, but which practicality and resource limitations restrict for approaches involving participants. It is feasible to imagine that the behaviours identified by each approach would also demonstrate variability on repeat applications, which has only been partially investigated in this research. Understanding the distribution of evacuation times would increase confidence in the generalisability of the findings.

To re-iterate a concern raised in Chapter 8, a particular limitation of the standardised comparison was differences in group sizes between the different approaches in the standardised comparison. To avoid disruption to other building users, the fire drill had to be conducted in one instance, thus all 22 participants were evacuated at the same time. However, bandwidth and computer/lab space availability restricted the number of participants per group in the VE to a maximum of four participants. Several authors recognise the importance of social behaviours in emergencies (Pauls and Jones, 1980b; Sime, 1995; Aguirre et al., 1998; Mawson, 2005; Pan et al., 2006), and in particular Aguirre et al. (1998) found that occupants in small groups responded quicker if the situation was perceived as dangerous. Thus, the differences in group sizes may be a confounding variable in the standardised comparison. However, the participants in the standardised comparison did not perceive the situation as very dangerous, and the inclusion of groups allowed for some investigation of social behaviours.

The talk-through approach was developed and assessed with participants making individual predictions. Social and group behaviours were investigated only through interactions which were predicted by the participant; the anticipated response of the other party was determined by the experimenter. As mentioned above, social behaviours are an important influence on the outcome of an emergency, although their impact on the talk-through approach was not analysed in this research. However, given the generally favourable conclusions drawn about concurrent validity of the approaches in the standardised comparison regardless, the impact of this issue is limited.

Most of the VE research conducted for this thesis used standard desktop computers or laptops. These were chosen due to practicality and accessibility as no VEs with greater sophistication were available for this research at that time. Furthermore, during the development of the emergency training simulator within the DiFac EU-project (Section 6.3) a 3D stereoscopic rear projection screen on early prototypes was rejected in favour of a desktop solution. This was due to the lower cost and wider availability of the latter, which were important requirements of the target user-group. Despite the use of PCs and laptops for the research conducted for this thesis, it must be recognised that VEs with greater sophistication exist which may provide potential benefits for further research in this area. For example, CAVEs and projection walls provide greater levels of immersion (Bowman and McMahan, 2007) than desktop systems. Higher levels of immersion, which indicate an increase in the sensory fidelity of a VE (Slater, 2003), have resulted in improved spatial understanding and awareness (Bowman and McMahan, 2007; Schuchardt and Bowman, 2007; Sowndararajan et al., 2008). Increased spatial awareness may be particularly beneficial during evacuation research as it may affect navigation from the building, and in particular resolve some of the issues reported in VE during the standardised comparison. Several studies within Chapter 2 were reported to have used head-mounted displays as an alternative VE (Meguro et al., 1998; Mantovani et al., 2001; Gamberini et al., 2003), although these have been associated with an increase in symptoms of simulator sickness (Sharples et al., 2008).

Finally, it is worth re-iterating some of the problems experienced with the Second Life VE. This was selected due to low cost and ease of learning for development, as well as the available support at the University of Nottingham. However, several problems were experienced with the software, including:

- Some avatars consistently looked towards the ground due to a bug
- The VE sometimes appeared “jerky” due to delays loading the environment
- The voice communications were sometimes poor quality

These issues may have affected the realism implied by the VE by focussing participants on technology problems rather than on the evacuation task. An alternative to Second Life was to use a gaming development toolkit (Smith and Trenholme, 2009). This would have been likely to address the issues described above. However, the development time may have been longer for someone without prior experience, and no technical support was available at the University of Nottingham. Furthermore, using a different development program to Smith and Trenholme (2009) added to the original contribution of the work, and provided additional understanding of the use of VEs for investigating behaviour in emergencies.

9.7 Chapter summary

This chapter has presented a summary of the research findings, and discussed them with consideration given to previous research. Discussion was made on the overall performance of the predictive approaches against the criteria for judging their quality. Based on the research findings, recommendations have been made for practitioners involved with making behavioural predictions in emergency situations. A review is made of the novel contribution to knowledge. Finally, the limitations of the research have been discussed.

10. Conclusions and future work

10.1 Chapter overview

The chapter presents the main conclusions drawn from this work. Evidence is provided for achievement of the aims, which were first introduced in Chapter 1. Future work is proposed to address the limitations of this research.

10.2 Conclusions

This thesis has presented an analysis of approaches for predicting human behaviour in emergencies. In particular, the validity, reliability, sensitivity, ethics and resources of each approach were evaluated. This work has begun to address the lack of structured and comprehensive assessments of the approaches used in previous studies. These often focussed on the predicted behaviours themselves, rather than the methods and approaches used to generate the predictions.

The approaches were evaluated for use by human factors professionals, for applications such as designing a building for safe egress, informing training programmes and developing emergency response procedures. This was due to the growing recognition that human factors can contribute to improving safety in emergency situations. It was also in response to criticism that human factors professionals do not always pay sufficient attention to the quality of the methods they use.

Selection of the approaches was based on a review of previous literature, which revealed that fire drills, expert predictions, virtual environments and participant predictions had not previously been analysed against all the criteria mentioned above, yet they offered potential for predicting behaviour in emergencies. Tests of the approaches were conducted, in which virtual environments and fire drills were seen to perform sufficiently well against the criteria to be investigated further. In contrast the results from a study of expert predictions demonstrated low predictive validity when compared to reports of actual human behaviour in real fires.

A talk-through approach was also developed in which participants were asked to describe the actions they would take in response to an emergency scenario. The predicted actions were found to be significantly correlated with behaviours demonstrated in real fires. The approach required low resources and presented a low risk of injury or distress which indicated its potential value for behavioural prediction in human factors applications.

A systematic comparison was subsequently conducted of the behaviours obtained from a fire drill, a VE study, and the talk-through approach when applied to an evacuation from a university building. The results were grounded with a review of the scientific literature. The behaviours observed in the fire drill demonstrated comparability with those predicted from the other approaches. These results can inform the on-going debate about the accuracy of behaviours demonstrated during fire drills. The behaviours obtained through the VE study also demonstrated comparability with those from the other approaches. However, some concerns were raised about acts related to navigation in the VE. Again, considering the wider context of this research, several recent papers have investigated other aspects of the use of VEs for behavioural investigation in emergencies; the studies conducted for this thesis further contribute to an understanding of their strengths and weaknesses. Finally, the standardised comparison provided additional evidence to support the use of the talk-through approach for predicting human behaviour in emergencies. It also provided behaviours which were comparable with those from other approaches, was conducted quickly, and posed no risk of physical injury. Minor drawbacks were indications that it may not be sufficiently sensitive to investigate some aspects of design detail, and social processes were difficult to incorporate.

The limitations of any behavioural prediction require careful consideration, and in this research prediction referred to the ability of an approach to produce specific data, under the conditions investigated. The predictive validity of the approaches, and means by which it has been assessed, must be understood before they are used in any further applications. Another important limitation was a focus on building evacuations throughout the research. Future work is required to investigate the strength of these approaches in other scenarios and for different types of emergencies.

Recommendations were made for the selection of approaches, based on the purpose of the prediction. The talk-through approach, VEs and fire drills can all contribute to design applications, but at different stages depending on the level of the design's development and its representation. Both fire drills and VEs were recommended for informing emergency response procedures. For developing training programmes, predictions derived from fire drills, VEs and the talk-through approach could prove useful. While the limitations of the research must be considered when using these recommendations, this consolidated guidance did not previously exist in the scientific literature. Thus, in conjunction with the recommended future work, the information provided in this thesis can be used to inform the practice of making predictions of human behaviour in emergency situations.

10.3 Review of aims

The aims, originally described in the introduction Chapter 1, are shown in boxed text below. The achievements associated with each of the aims are discussed.

Aim 1: *To analyse approaches for predicting human behaviour in emergencies against established criteria for assessing their quality.*

Detailed and systematic analysis has been conducted of several approaches, namely: fire drills, VEs, a talk-through approach, the use of literature and expert predictions. Other approaches exist for predicting human behaviour in emergencies which were not analysed, for example simulation and analysis of reports by survivors. However, a review of the previous literature in Chapter 2 revealed that the former had received considerable attention by other authors (see Section 2.4) whereas the latter can only be conducted following an emergency incident, and if access to the survivors is possible. Thus, the selection of approaches was based on those with potential for predicting human behaviour in emergencies, with the feasibility and opportunity to be investigated further.

Each of the approaches has been assessed against established criteria for assessing the quality of a human factors approach, based on those provided by Wilson (2005). This addressed a general concern about ergonomists paying insufficient attention to the methods they use (Annett, 2002). Section 1.2 of this thesis explained that a systems engineering approach, such as can be found in human factors, can help improve safety in emergency situations, yet it was unclear how traditional HF approaches fare in the analysis and prediction of human interaction with systems in emergency situations. Thus, this analysis of predictive approaches can be used by human factors professionals to make an additional contribution to the discipline. Of particular relevance is the criterion *Sensitivity* which was used to evaluate the appropriateness of the data produced by each method for human factors applications.

Assessment of the approaches was made with consideration to their application by human factors professionals to evacuation scenarios. Future research should investigate other emergencies, for example CBRN incidents (Section 4.3).

Finally, practicality prevented specific use of all of the criteria listed by Wilson (2005), in particular non-reactivity, acceptability and feasibility of use. However, aspects of these were reviewed within the applied criteria. A more notable exclusion was generalisability, which was discussed as a limitation in Section 9.6.

Aim 2: *To develop new approaches for predicting human behaviour in emergency situations.*

The most notable achievement was the development of the talk-through approach reported in Chapter 7, and applied in the standardised comparison in Chapter 8. While drawing upon existing techniques (*talk-through* Kirwan and Ainsworth, 1992 and *sequential analysis* Bakeman and Gottman, 1986), the novelty was the combination of these methods and application to emergency scenarios. Furthermore the approach showed potential for behavioural predictions as it requires low-resources, and yet the results demonstrated comparability with those from a study of behaviour in real fires. However, the approach needs testing in emergencies other than building evacuation (see Sections 9.6 and 10.4).

The approach taken for the expert prediction was also unique, using the highly regarded paired comparisons technique (Sinclair, 2005) for predicting the likelihood of behaviours in emergencies. However, the approach demonstrated limited promise, showing no significant correlations with the acts from a reference study. While improvements may be made (for example through the use of a more detailed reference scenario), to be useful the approach also needs to be developed to address the lack of behavioural sequence data.

Aim 3: *To make a systematic comparison of approaches for predicting human behaviour in emergencies.*

The standardised comparison (Chapter 8) allowed the approaches to be compared using the same scenario for behavioural prediction. This enabled greater control of the variables, and therefore a more accurate assessment against the criteria for judging their quality. For example, a detailed analysis was made of the resources required to apply each approach, which would not have been possible had they been applied individually to different scenarios. The comparison was limited to a building evacuation; future work should investigate a range of different scenarios.

Aim 4: *To develop recommendations and guidance for HF professionals responsible for behavioural predictions in emergency situations.*

Guidance has been inherent through the assessment of the methods against the criteria for judging the quality of a method, and in particular a review of the *Sensitivity* referring to the appropriateness of the data for human factors applications. Detailed recommendations and guidance for practitioners was provided in Section 9.4, which was based on the results of this research and those from previous investigations of human behaviour in emergencies. Recommendations for further work are presented in Section 10.4.

10.4 Future work

To understand the generalisability of the approaches to scenarios other than building evacuations they should be investigated for a range of emergencies. Furthermore, they should also be investigated in a range of contexts, including industrial scenarios, urban environments, boats, aeroplanes and trains. The research should give consideration to the quality of the approaches through assessment against the criteria used in this thesis. Taking future work in this direction would extend the analysis of the approaches' performance to these different scenarios, thus augmenting the guidance available to practitioners who need to make behavioural predictions for emergency situations.

Future work should investigate the validity of the approaches against incident reports and survivors of an emergency, if access can be gained to these rich sources of data. This work would address the concerns raised about the reference study used for this research (Canter et al., 1980, Section 9.6). The standardised comparison allowed for a controlled investigation of the behaviours in a process of triangulation, but an understanding of the predictive validity of the approaches would be enhanced with data pertaining to a real emergency event.

Another important future step is for a number of researchers to apply the predictive approaches to further investigate their reliability. This work should be conducted without the researchers having knowledge of the reference data set prior to conducting their behavioural predictions to avoid any experimenter bias.

Considering future work for the specific approaches, one avenue to develop the talk-through approach is to investigate its ability to predict social interactions. This could be done by inviting groups of participants to predict their actions in response to an emergency scenario. Thus, their social interactions and group behaviours could be studied as their predicted response to the event unfolds. This may increase the validity of the social behaviours obtained using the approach.

Another opportunity to enhance the talk-through approach may be to adopt techniques from the cognitive interview. This method has proven successful in increasing eye witness recollection and description of events during police interviews. The cognitive interview uses a variety of techniques to assist the interviewees in recreating a mental representation of the event. For example, they are probed on the environmental and contextual details, which may have influenced their reactions. These details are in turn fed back to the interviewee before they are asked about the events (Geiselman et al., 1985; Fisher et al., 1989; Memon et al., 1997). These techniques may also help participants envisage their hypothetical behaviours in a predictive application.

The feedback from participants may also be improved by adapting elements of the cooperative evaluation approach (Wright and Monk, 1991). Originally applied to the evaluation of user interfaces, this approach requires evaluators to interact with participants during a trial. The participants are encouraged to ask questions if confused, as these can reveal weaknesses in the system (Wright and Monk, 1991). A similar technique could be used in the talk-through approach, in which participants are encouraged to interact with the researcher, to highlight key decision stages or to provide an understanding of confusion at points in the act sequences.

For the VE, future work could investigate behaviours using a more reliable platform and one in which larger group sizes can be studied when co-located without problems associated with bandwidth limitations. However, using more than approximately 10 participants is likely to be impractical anyway, as managing the participants and their avatars would become difficult. It may also be difficult to keep the participants in the same physical location to administer the experiment yet avoid real-world communication while they are using the VE. One possible solution is to keep a small numbers of participants, but include autonomous agents within the VE. Thus, the participant would feel that they are part of a larger crowd in the VE. A similar approach to this was also suggested by Mol et al. (2008).

Another opportunity to improve the VE would be to incorporate measures to improve navigation, such as the guidelines tested by Smith and Marsh (2004) mentioned in Section 9.3. This would involve placing pictures and features on walls in the VE to give an indication of the area off screen when obscured from the avatar's viewpoint by walls in the foreground. Consideration would need to be given to the implementation of these measures, as they may require artefacts to be introduced to the VE which were not present in the real world. Therefore, transferability of the findings from the VE to the real world would need to be evaluated. However, they may provide a solution to address the concerns raised about navigation in the VE raised in Chapter 8.

10.5 Chapter summary

Research has been conducted into approaches for predicting human behaviour in emergency situations. Investigations into the quality of the approaches were conducted through a series of tests and a systematic comparison using a building evacuation scenario. A new approach has also been developed, in which participants are asked to describe how they would respond to an emergency scenario. The research findings provide previously unavailable guidance for human factors professionals in the selection and application of approaches for predicting human behaviour in emergencies.

11. References

Aguirre, B.E., Wenger, D., and Vigo, G. (1998). A test of the Emergent Norm Theory of collective behavior. *Sociological Forum,* 13(2), pp.301-320.

Ambulance HART. (2009). *Mass Decontamination Unit exercise for West Midlands HART team at RAF Cosford*. Retrieved 18 October 2009, from http://www.ambulancehart.org.uk/editorial.asp?page_id=17&n_id=31

Annett, J. (2002). A note on the validity and reliability of ergonomics methods. *Theoretical Issues in Ergonomics Science*, 3(2), pp.228-232.

Averill, J., Peacock, R., Kuligowski, E., and Reneke, R. (2009). Federal investigation of the evacuation of the World Trade Center on September 11, 2001. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.7-14.

Bakeman, R., and Gottman, J. M. (1986). *Observing Interaction: An Introduction to Sequential Analysis.* Cambridge: Cambridge University Press.

Banyard, P and Flanagan, C. (2005). *Ethical Issues and Guidelines in Psychology*. Hove: Routledge.

Banyard, P. and Grayson, A. (2000). *Introducing Psychological Research* (2nd Ed.). Basingstoke: PALGRAVE.

BBC News. (2001). *US attack hero caught on film.* Retrieved 18 October 2009, from http://news.bbc.co.uk/1/hi/uk/1564229.stm

Boer, L.C. (2005). Getting passengers out – evacuation behaviours. In: J.R. Wilson, B.J. Norris, T. Clarke, A. Mills (Eds.), *Rail Human Factors: Supporting the Integrated Railway.* London: Ashgate Publishing, pp.462–472.

Bohannon, R.W. (1997). Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. *Age and Ageing*, 1997(26), pp.15-19.

Bosley, K. (2003). *Domestic Firefighting.* Retrieved 8 January 2008, from http://www.rmd.communities.gov.uk/project.asp?intProjectID=10887

Bowman, D.A., and McMahan, R.P. (2007). *Computer*, 40(7), pp36-43

Boyce, K.E., Purser, D., and Shields, J. (2009). Experimental studies to investigate merging behaviour in a staircase. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.111-122.

Brooks, C.J., Muir, H.C., and Gibbs, P.N.G. (2001). The basis for the development of a fuselage evacuation time for a ditched helicopter. *Aviation, Space and Environmental Medicine,* 72(6), pp.553-561.

Burns, W. J. (2007). *Risk perception: A review.* Retrieved 14 October 2010, from http://create.usc.edu/research/54570.pdf

Business Link (2009). *Fire Safety and Risk Assessment.* Retrieved 18 October 2009, from http://www.businesslink.gov.uk/bdotg/action/detail?type=RESOURCES&itemId=1073793917

Cabinet Office Civil Contingencies Secretariat (2005). *Emergency Response and Recovery.* Retrieved 18 October 2009, from http://www.cabinetoffice.gov.uk/media/238642/err-guidance-120809.pdf#page=62

Canter, D., Breaux, J., and Sime, J. (1980). Domestic, multiple occupancy, and hospital fires. In: D. Canter (Ed.), *Fires and Human Behaviour*. New York: John Wiley and Sons, pp.117-136.

Chittaro, L., and Ranon, R. (2009). Serious games for training occupants of a building in personal fire safety skills. In: *Proceedings of the 1st IEEE International Conference in Games and Virtual Worlds for Serious Applications (VS-GAMES 2009).* Coventry 23-24 March 2009. pp.76-83.

Christoffersen, B., and Soderlind, C. (2009). Comparison of two egress models and a full-scale experiment. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.573-578.

Cobb, S.V.G., Nichols, S., Ramsey, A., and Wilson, J.R. (1999). Virtual Reality-Induced Symptoms and Effects (VRISE). *Presence,* 8(2), pp.169-186.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Earlbaum Associates.

Colgan, P. W., and Smith, T.J. (1978). Multidimensional contingency table analysis. In P.W. Colgan (Ed.), *Quantitative Ethology.* New York: Wiley, pp.145–174.

Communities and Local Government, (2009a). *New Dimension mass decontamination photos.* Retrieved 18 October 2009, from http://www.communities.gov.uk/fire/resilienceresponse/newdimensionequipping/newdimensionmass/

Communities and Local Government, (2009b). *Introduction to the New Dimension Project.* Retrieved 18 October 2009, from http://www.communities.gov.uk/fire/resilienceresponse/newdimensionequipping/introduction/

Comptdaer J., Bourguignon, D., and Chiva, E. (2005). A new microscopic approach to crowd modelling applied to urban crisis management training. Paper presented at *the 2005 Conference on Behavior Representations in Modeling and Simulation – BRIMS.* Universal City, CA 16-19 May 2005.

Cornwell, J. B., Silverman, B. G., O'Brien, K., and Johns, M. (2002). A demonstration of the PMF-extraction approach: modeling the effects of sound on crowd behavior. In: *Proceedings of the Eleventh Conference on Computer Generated Forces and Behavioral Representation*. Orlando 2002. pp.107-113.

Dane, F.C. (1990). *Research methods.* Belmont: Brooks/Cole Publishing Co.

Deere, S.J., Galea, E.R., and Lawrence, P.J. (2009). A systematic methodology to assess the impact of human factors in ship design. *Applied Mathematical Modelling,* 33(2), pp.867-883.

DiFac consortium (2006). *Annex I: Description of Work.* DiFac (IST5-035079). Milan: Istituto di Tecnologie Industriali e Automazione.

Dombroski, M., Fischhoff, B., and Fischbeck, P. (2006). Predicting emergency evacuation and sheltering behaviour: a structured analytical approach. *Risk Analysis* 26(6), pp.1675-1688.

Drury, C.G. (2005). Designing ergonomics studies. In: J.R. Wilson, and N. Corlett, (Eds.), *Evaluation of Human Work. 3rd ed.* London: Taylor & Francis, pp.39-60.

Drury, J. (2004). No need to panic. *The Psychologist,* 17(3), pp.118–119.

Drury, J., Cockling, C., and Reicher, S. (2006). Every man for himself - or for the group? How crowd solidarity can arise in an emergency: An Interview Study of Disaster Survivors. Paper presented at: *Group and Intergroup Relations Pre-Conference, Society for Personality and Social Psychology 7th Annual Meeting.* Palm Springs, CA January 2006.

Dunbar, G. (2005). *Evaluating research methods in psychology: A case study approach.* Oxford: BPS Blackwell.

Edelman, P., Herz, E., and Bickman, L. (1980). A model of behaviour in fires applied to a nursing home fire. In: D. Canter (Ed.), *Fires and Human Behaviour*. New York: John Wiley and Sons, pp.181-203.

Fahy, R.F., and Proulx, G. (2005). *Analysis of published accounts of the World Trade Center evacuation.* NIST report NCSTAR 1-7A. Washington: NIST National Institute of Standards and Technology.

Fahy, R.F. (2005). Available data and input into models. In: E.D. Kuligowski and R.D. Peacock, (Eds.), *Workshop on Building Occupant Movement During Fire Emergencies*, National Institute of Standards and Technology, NIST SP 1032, Gaithersburg, MD, pp. 62–67.

Farrington-Darby, T., Wilson, J. R., Norris, B. J., and Clarke, T. (2006). A naturalistic study of railway controllers. *Ergonomics*, 49(12&13), pp.1370-1394.

Fischer, H.W. (2002). Terrorism and 11 September 2001: does the "behavioural response to disaster" model fit? *Disaster and Prevention Management,* 11(2), pp.123-127.

Fischer, H.W. (2003). Behavioral and organizational responses to terrorism: A model based on 9/11. *Journal of Emergency Management,* 1(3), pp.42-48.

Fisher, R.P., Geiselman, R.E., and Amador, M. (1989). Field test of the cognitive interview: enhancing the recollection of actual victims and witnesses of crime. *Journal of Applied Psychology,* 74(5), pp.722-727.

Galea, E.R., Hulse, L., Day, R., Siddiqui, A., and Sharp, G. (2009). The UK WTC 9/11 evacuation study: an overview of the methodologies employed and some analysis relating to fatigue, stair travel speeds and occupant response times. In: *Proceedings* *of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.27-40.

Galea, E.R., Shields, J., Canter, D., Boyce, K., Day, R., Hulse, L., Siddiqui, A., Summerfield, L., Marselle, M., Greenhall, P.V. (2007). The UK WTC 9/11 evacuation study: methodologies used in the elicitation and storage of human factors data. In: *Proceedings of the 11th International Fire Science & Engineering Conference, Interflam 2007.* London 3-5 September 2007. pp.169-181.

Gamberini, L., Cottone, P., Spagnolli, A., Varotto, D., and Mantovani, G. (2003). Responding to a fire emergency in a virtual environment: different patterns of action for different situations. *Ergonomics*, 46(8), pp.842-858.

Geiselman, R.E., Fisher, R.P., MacKinnon, D.P., and Holland, H.L. (1985). Eyewitness memory enhancement in the Police interview: cognitive retrieval mnemonics versus hypnosis. *Journal of Applied Psychology,* 70(2), pp.401-412.

Gershon, R.R.M. (2009). The World Trade Center evacuation study: factors associated with evacuation time and injury. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.15-26.

Gershon, R.R.M., Qureshi, K.A., Rubin, M.S., and Raveis, V.H. (2007). Factors associated with high-rise evacuation: qualitative results from the World Trade Center evacuation study. *Prehospital and Disaster Medicine*, 22(3), pp.165-173.

Glass, T.A., and Schoch-Spana, M. (2002). Bioterrorism and the people: how to vaccinate a city against panic. *Confronting Biological Weapons*, 34(2), pp.217-223.

Groner, N. (2009). A situation awareness requirements analysis for the use of elevators during fire emergencies. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire*. Cambridge 13-15 July 2009. London: Interscience Communications, pp.61-72.

Gwynne, S. (2009). The standardisation of human egress data. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.481-492.

Gwynne, S., Galea, E.R., Owen, M., Lawrence, P.J., and Filippidis, L. (2005). A systematic comparison of buildingEXODUS predictions with experimental data from the Stapelfeldt trials and the Milburn House evacuation. *Applied Mathematical Modeling* 29(9), pp.818-851.

Gwynne, S., Galea, E.R., Parke, J., and Hickson, J. (2003). The collection and analysis of pre-evacuation times derived from evacuation trials and their application to evacuation modelling. *Fire Technology*, 39, pp.173-195.

Gwynne, S., Galea., E.R., Owen, M., Lawrence., P.J., and Filippidis, L. (1999). A review of methodologies used in the computer simulation of evacuation from the built environment. *Building and Environment,* 34(6), pp.741-749.

Haag, A.B. (1992). Ergonomics standards, guidelines, and strategies for the prevention of back injuries. *Occupational Medicine*, 7(1), pp.155-165.

Haslam, C. (2010). *Smoke alarms and detectors.* Retrieved 17 September 2010, from www.fireservice.co.uk

Heyes, E., and Spearpoint, M. (2009). Lifts for evacuation – human behaviour considerations. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire*. Cambridge 13-15 July 2009. London: Interscience Communications, pp.73-84.

Home Office (2009). *Chemical, biological, radiological, nuclear and explosives (CBRNE).* Retrieved 18 October 2009, from: http://security.homeoffice.gov.uk/cbrn-resilience/decontamination/

Howitt D. and Cramer, D. (2011). *Introduction to Research Methods in Psychology* (3rd Ed.). Harlow: Pearson Education Ltd.

Hsiung, K-H., Chein, S-W., Tseng, C-H., and Huang, P-T. (2009). A study on high rise building fire evacuation strategies for Taipei 101 financial centre. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.51-60.

Jarrett, C. (2009). Get a second life. *The Psychologist,* 22(6), pp.490-493.

Jeon, G., and Hong, W. (2009). Characteristic features of the behavior and perception of evacuees from the Daegu subway fire and safety measure in an underground fire. *Journal of Asian Architecture and Building Engineering,* 8(2), pp.415-422.

Jong-Hoon, K., Woon-Hyung, K., Sam-Kew, R., Duck-Hee, L., and Woo-Sung, J. (2009). Experiments of the subway car egress. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.619-622.

Kady, R.A., and Allen, K. (2009). School bus evacuation: research to practice. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.623-626.

Kady, R.A., Gwynne, S., and David, J. (2009). A review of the sources of occupant performance data used in building evacuation models. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.471-480.

Kanno, T., Shimizu, T., and Furuta, K. (2006). Modeling and simulation of residents' response in nuclear disaster. *Cognition, Technology and Work,* 8(2), pp.124-136.

Kecklund, L., Anderzen, I., Petterson, S., Haggstrom, J., and Wahlstrom, B. (2009). Evacuation in trains – the railway safety challenge. Paper presented at *the Third International Conference on Rail Human Factors*. Lille 3-5 March 2009.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology,* 3(3), pp.203-220.

Kieras, D.E., and Meyer, D.E. (2000). The role of cognitive task analysis in the application of predictive models of human performance. In: J.M. Schraagen, S.F. Chipman and V.L. Shalin (Eds.), *Cognitive Task Analysis*. Mahwah: Lawrence Erlbaum Associates, pp.237-260.

Kinsey, M.J., Galea, E.R., and Lawrence, P.J. (2009a). Investigating the use of elevators for high-rise building evacuation through computer simulation. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.85-96.

Kinsey, M.J., Galea, E.R. and Lawrence, P.J. (2009b). Extended model of pedestrian escalator behaviour based on data collected within a Chinese underground station. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.173-185.

Kirwan, B. (1994). *A Guide to Practical Human Reliability Assessment*. London: Taylor and Francis.

Kirwan, B., and Ainsworth, L. K. (1992). *A Guide to Task Analysis.* London: Taylor and Francis.

Ko, S., Spearpoint, M., and Teo, A. (2007). Trial evacuation of an industrial premises and evacuation model comparison. *Fire Safety Journal,* 42(2), pp.91-105.

Kobes, M., Oberije, N., Groenewegen, K., Hesloot, I., and Vries, B. (2009). Hotel evacuation at night; an analysis of unannounced fire drills under various conditions. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.219-230.

Kuligowski, E. (2003). *The evaluation of a performance-based design process for a hotel building: the comparison of two egress models.* MSc thesis. The University of Maryland.

Kuligowski, E. (2009). The process of human behaviour in fires. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.627-632.

Kuligowski, E.D., and Peacock, R.D. (2005). *A review of building evacuation models.* Technical Note 1471. Washington: NIST National Institute of Standards and Technology.

Lancaster and Morecambe College (2008). Decontamination exercise at LMC. Retrieved 18 October, 2009, from http://www.lmc.ac.uk/home/news/2009/decontamination.php.

Landis, J.R., and Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, pp.159-174.

Laughery, R. (2005). Simulation and modelling for analysing human performance. In: J.R. Wilson, and N. Corlett, (Eds.), *Evaluation of Human Work. 3rd ed.* London: Taylor & Francis, pp.219-238.

Lawson, G., and D'Cruz, M. (2008). *User-group evaluation results, Turin June 08.* DiFac (IST-5-035079). Internal report. Nottingham: The University of Nottingham.

Lawson, G., and D'Cruz, M. (in press). Ergonomics methods and the digital factory. Accepted for publication in: L. Canetta, C. Redaelli, and M. Flores (Eds.), *DiFac: Digital Factory for Human Oriented Production System.* Berlin/Heidelberg: Springer.

Lawson, G., D'Cruz, M., and Bourguignon, D. (2008). *End-user evaluation of the emergency training simulator.* DiFac (IST-5-035079). Internal report. Nottingham: The University of Nottingham.

Lawson, G., D'Cruz, M., Bourguignon, D., and Pavlopoulos, C. (2007a). *D13 Methodologies for Ergonomic Programming, Observation and Analysis.* Report produced for DiFac (IST-5-035079). Nottingham: The University of Nottingham.

Lawson, G., D'Cruz, M., Bourguignon, D., and Pentenrieder, K. (2007b): Training in the Digital Factory. Paper Presented at *IFAC Workshop on Manufacturing, Modelling, Management and Control.* Budapest 14-16 November 2007.

Lawson, G., Sharples, S., Clarke, D., and Cobb, S. (2009a). The use of experts for predicting human behaviour in fire. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.493-500.

Lawson, G., Sharples, S., Clarke, D., and Cobb, S. (2009b). Development of a technique for predicting the human response to an emergency situation. In: D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics, Proceedings of the 8th International Conference, EPCE 2009, HCI International 2009.* San Diego 19-24 July 2009. Heidelberg: Springer, pp.22-31.

Lawson, G., Sharples, S., Cobb, S., and Clarke, D. (2009c). Predicting the human response to an emergency. In: P.D. Bust (Ed.), *Contemporary Ergonomics 2009, Proceedings of the Ergonomics Society's Annual Conference.* London 22-23 April 2009. London: Taylor and Francis, pp.525-532.

Li, H., Tang, W., and Simpson, D. (2004). Behaviour based motion simulation for fire evacuation procedures. In: *Proceedings of the Theory and Practice of Computer Graphics.* Bournemouth 8-10 June 2004. pp.112-118.

Linden Research Inc. (2009). *Second Life.* Retrieved 18 October 2009, from http://secondlife.com/

Linden Research Inc. (2010). *Second Life.* Retrieved 5 July 2010, from http://secondlife.com/

Mantovani, G., Gamberini, L., Martinelli, M., and Varotto, D. (2001). Exploring the suitability of virtual environments for safety training: signals, norms and ambiguity in a simulated emergency escape. *Cognition, Technology and Work,* 3(1), pp.33-41.

Mawson, A.R. (2005). Understanding Mass Panic and Other Collective Responses to Threat and Disaster. *Psychiatry,* 68(2), pp.95-113.

McConnell, N.C., Boyce, K.E., and Shields, T.J. (2009). An analysis of the recognition and response behaviours of evacuees of WTC 1 on 9/11. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.659-669.

Meacham, B., Lord, J., Moore, A., Fahy, R.F., Proulx, G., and Notarianni, K. (2004). Investigation of uncertainty in egress models and data. In: *Proceedings of the 3rd International Symposium on Human Behaviour in Fire.* Belfast, 1-3 Sept 2004. pp.419-428.

Meguro, K., Haga, Y., Yamazaki, F. and Katayama. T. (1998). Application of Virtual Reality to Human Evacuation Behavior, Structural Safety and Reliability. In: *Proceedings of the International Conference on Safety Structure and Reliability ICOSSAR'97*. Kyoto 24-28 November 1997, pp.1603-1610.

Melly, M., Lennon, P., and Lennon, R. (2009). Who defers to whom? Deference Behaviour on stairs. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.135-146.

Memon, A., Wark, L., Holley, A., Bull, R., Koehnken, G. (1997). Eyewitness performance in cognitive and structured interviews. *Memory,* 5(5), pp.639–55.

Miller, G.A. (1994). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 101(2), pp.343-352.

Mintz, A. (1951). Non-adaptive group behaviour. *Journal of Abnormal and Social Psychology,* 46(2), pp.150-159.

Mol, A.C.A., Jorge, C.A.F., and Couto, P.M. (2008). Using a Game Engine for VR Simulations in Evacuation Planning. *IEEE Computer Graphics and Applications,* 28(3), pp.6-12.

Muhdi, R., Davis, J., and Blackburn, T. (2006). Improving occupant characteristics in performance-based evacuation modeling. In: *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting-2006*. San Francisco October 16-20, 2006. pp.1199-1203.

Muhdi, R.A. (2006). Evacuation Modeling: Development, characteristics, and limitations. In: *National Occupational Research Agenda (NORA) Conference Proceedings,* 2006. pp.87-92.

Muir, H.C. (1996). Research into the factors influencing survival in aircraft accidents. *Aeronautical Journal,* May 1996, pp.177-181.

Muir, H.C., Bottomley, D.M., and Marrison, C. (1996). Effects of motivation and cabin configuration on emergency aircraft evacuation behaviour and rates of egress. *The International Journal of Aviation Psychology,* 6(1), pp.57-77.

Munley, G.A., Mortimer, R.J., Oyinloye, J.O., Worth, S.V., and Williams, J.C. (1996). Developing and validating Escapaid: a computer model of evacuation from complex spaces. In: S.A. Robertson (Ed.), *Contemporary Ergonomics 1996.* London: Taylor and Francis, pp.171-176.

Murakimi, Y., Sugimoto, Y., and Ishida, T. (2005). Modeling human behavior for virtual training systems. In: *Proceedings of the 20th National Conference in Artificial Intelligence 2005*. Pittsburgh. pp.127-132.

Neale D. C. (1997), Factors influencing spatial awareness and orientation in desktop virtual environments. In: *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*. Albuquerque 22-26 September 1997. pp.1278-1282.

Neale, H., and Nichols, S. (2001). Theme-based content analysis: a flexible method for virtual environment evaluation. *International Journal of Human-Computer Studies,* 55(2), pp.167-189.

Nichols, S., and Patel, H. (2002). Health and Safety implications of virtual reality: a review of empirical evidence. *Applied Ergonomics,* 33(3), pp.251-271.

Nilsson, D., and Uhr, C. (2009). Complex Systems – A holistic approach for understanding and modelling fire evacuation behaviour. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire*. Cambridge 13-15 July 2009. London: Interscience Communications, pp.525-530.

Nunnally, J.C. (1967). *Psychometric Theory.* New York: McGraw-Hill.

Olsson, P.A., and Regan, M.A. (2001). A comparison between actual and predicted evacuation times. *Safety Science,* 38(2), pp.139-145.

Ozel, F. (2001). Time pressure and stress as a factor during emergency egress. *Safety Science,* 38(2), pp.95-107.

Pan, X., Han, C.S., and Law, K.H. (2005). A multi-agent based simulation framework for the study of human and social behaviour in egress analysis. In: *Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering. American Society of Civil Engineers.* Cancun, Mexico 12-15 July 2005, pp.969-980.

Pan, X., Han, C.S., Dauber, K., and Law, K.H. (2006). Human and social behavior in computational modeling and analysis of egress. *Automation in Construction,* 15(4), pp.448-461.

Pauls, J.L. (1999). A personal perspective on research, consulting and codes/standards development in fire-related human behaviour, 1969-1999, with an emphasis on space and time factors. *Fire and Materials,* 23(6), pp.265-272.

Pauls, J.L., and Jones, B.K. (1980a). Building evacuation: research methods and case studies. In: D. Canter (Ed.), *Fires and Human Behaviour*. New York: John Wiley and Sons, pp.227-250.

Pauls, J.L., and Jones, B.K. (1980b). Research in human behaviour. *Fire Journal,* 74(3), pp.35-41.

Pentenrieder, K., Constantinescu, C., Redaelli, C., Pavlopoulos, C., Decker, F., and Lungu, R. (2007). *D19 New media for collaborative virtual manufacturing.* Report produced for DiFac (IST-5-035079). Munich: METAIO.

Perneger, V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal,* 316(7139), pp.1236-1238.

Perry, R.W., and Lindell, M.K. (2003). Preparedness for emergency response: guidelines for the emergency planning process. *Disasters,* 27(4), pp.336-350.

Pheasant, S. (1996). *Bodyspace, 2nd ed.* London: Taylor and Francis.

Pires, T, T. (2005). An approach for modeling human cognitive behaviour in evacuation models. *Fire Safety Journal,* 40(2), pp.177-189.

Proulx, G. (1993). A stress model for people facing a fire. *Journal of Environmental Psychology,* 13(2), pp.137-147.

Proulx, G. (1995). Evacuation time and movement in apartment buildings. *Fire Safety Journal,* 24(3), pp.229-246.

Proulx, G. (2001). *Occupant behaviour and evacuation.* In: Proceedings of the 9th International Fire Protection Symposium. Munich 25-26 May 2001. pp.219-232.

Proulx, G. (2007). Response to fire alarms. *Fire Protection Engineering,* 33(Winter 2007), pp.8-14.

Proulx, G., and Reid, I.M.A. (2006). Occupant behavior and evacuation during the Chicago cook county administration building fire. *Journal of Fire Protection Engineering,* 16(4), pp.283-309.

Proulx. G., Cavan, N.R., and Tonikian, R. (2006). *Egress times from single family houses.* Research Report: IRC-RR-209. Ontario: National Research Council Canada.

PSCA International (2009). *A New Dimension.* Retrieved 18 October, 2009, from http://www.publicservice.co.uk/article.asp?publication=Home%20Affairs&id=384&content_name=Civil%20contingencies&article=11667

Purser, D.A., and Boyce, K. (2009). Implications of modelling and experimental studies of evacuation behaviour on stairs for multi-storey building design. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire*. Cambridge 13-15 July 2009. London: Interscience Communications, pp.147-160.

Purser, D.A., Bensilum, M. (2001). Quantification of behaviour for engineering design standards and escape time calculations. *Safety Science,* 38(2001), pp.157-182.

Reason, J. (2000). Human error models and management. *Western Journal of Medicine,* 172(6), pp.393-396.

Ren, A., Chen, C., and Luo, Y. (2008). Simulation of emergency evacuation in virtual reality. *Tsingua Science and Technology,* 13(5), pp.674-680.

Rowe, G., and Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting,* 15(4), pp.353-375.

Rymaszewski, M., Au, W.J., Wallace, M., Winters, C., Ondrejka, C., and Bastone-Cunningham., B. (2007). *Second Life: the official guide.* Indianapolis: Wiley publishing.

Saieh, N. (2008). *Loisium Hotel / Steven Holl.* Retrieved 8 May 2011, from http://www.archdaily.com/5524/loisium-hotel-steven-holl/

Santos, G., and Aguirre, B.E. (2004). A critical review of emergency evacuation simulation models. In: *Proceedings of NIST Workshop on Building Occupant Movement during Fire Emergencies.* 9-10 June 2004. Gaithersburg: National Institute of Standards and Technology, pp.27-52.

Schuchardt, P., and Bowman, D.A. (2007). The benefits of Immersion for spatial understanding of complex underground cave systems. In: *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*. 5-7 November 2007 Newport Beach, California. New York: ACM Press, pp. 121-124.

Shadbolt, N. (2005). Eliciting expertise. In: J.R. Wilson, and N. Corlett, (Eds.), *Evaluation of Human Work*, *3rd ed*. London: Taylor and Francis, pp.185-218.

Sharma, S. (2009). Simulation and modeling of group behavior during emergency evacuation. In: *Intelligent Agents, 2009. IA'09. IEEE Symposium on Intelligent Agents*. Nashville 30 March-2 April 2009. pp.122-127.

Sharma, S., and Gifford, S. (2005). Using RFID to evaluate evacuation behavior models. In: *Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society 2005. (NAFIPS 2005).* Detroit 26-28 June 2005. pp.804-808.

Sharma, S.B., Tabak, V., Brocklehurst, D., Sagun, A., and Bouchlaghem, D. (2009). A comprehensive modern approach to developing evacuation data capture/analysis and simulation tools for real world fire engineering. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire*. Cambridge 13-15 July 2009. London: Interscience Communications, pp.195-205

Sharples, S., Cobb, S., Moody, A., and Wilson, J.R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays,* 29(2008), pp. 58-69.

Shaw, R. (2001). Don't panic: behaviour in major incidents. *Disaster Prevention and Management*, 10(1), pp.5-10.

Shields, T.G., and Boyce, K.E. (2000). A study of evacuation from large retail stores. *Fire Safety Journal,* 35(1), pp.25-49.

Shields, T.J., and Proulx, G. (2000). The science of human behaviour: past research endeavours, current developments and fashioning a research agenda. In: *Proceedings of the Sixth International Symposium on Fire Safety Science.* Poiters, France 5-9 July 2000. pp.95-114.

Shih, N.J., Lin, C.Y., and Yang, C.H. (2000). A virtual-reality-based feasibility study of evacuation time compared to the traditional calculation method. *Fire Safety Journal,* 34(4), pp.377-391.

Silverman, B. G., Johns, M., Cornwell, J., and O'Brien, K. (2006). Human behaviour models for agents in simulators and games: part I - enabling science with PMFserv. *Presence: teleoperators and virtual environments*, 15(2), pp.139-162.

Silverman, B. G., Might, R., Dubois, R., Shin, H., Johns, M., and Weaver, R. (2001). Toward a human behaviour models anthology for synthetic agent development. In: *10th Computer Generated Forces (CGF) Proceedings.* Norfolk, VA 15-17 May 2001. New York: IEEE.

Sime, J.D. (1980). The concept of panic. In D. Canter (Ed.), *Fires and Human Behaviour*. New York: John Wiley and Sons, pp.63-82.

Sime, J.D. (1995). Crowd psychology and engineering. *Safety Science* 21(1), pp.1-14.

Sinclair, M. A. (2005). Participative assessment. In: J.R. Wilson, and N. Corlett, (Eds.) *Evaluation of Human Work. 3rd ed.* London: Taylor & Francis, pp.83-111.

Slater, M. (2003). A note on Presence terminology. *Presence-Connect* 3.

Slovic, P. (2002). Perception of risk posed by extreme events. Paper presented at *Risk Management Strategies in an Uncertain World.* Palisades, NY 12-13 April 2002.

Smith, S.P. and Du'Mont, S. (2009). Measuring the effect of gaming experience on virtual environment navigation tasks. In: *Proceeding of the IEEE Symposium on 3D User Interfaces, 2009 (3DUI 09)*. Layfayette 14-15 March 2009. pp.3-10.

Smith, S.P. and Marsh, T. (2004). Evaluating design guidelines for reducing user disorientation in a desktop virtual environment. *Virtual Reality,* 8(1), pp.55-62.

Smith, S.P. and Trenholme, D. (2009). Rapid prototyping a virtual fire drill environment using computer game technology. *Fire Safety Journal,* 44(4), pp.559-569.

South Yorkshire Fire and Rescue (2007). *Emergency mass decontamination exercise.* Retrieved 18 October 2009, from http://www.syfire.gov.uk/2E79599A64824507BE2207B739F1D6C1.html

Sowndararajan, A., Wang, R., Bowman, D.A. (2008). Quantifying the benefits of immersion for procedural training. In: *Proceedings of the 2008 workshop on Immersive projection technologies/Emerging display technologies,* 9-10 August 2008, Los Angeles, California. New York: ACM Press

Stanton, N., Hedge, A., Brookhuis, K., Salas, E., and Hendrick, H. (2004). *Handbook of human factors and ergonomics methods.* Boca Raton, FL: CRC Press.

Stanton, N., Salmon, P., and Walker, G. (2005). *Human Factors Methods: A Practical Guide for Engineering and Design.* Aldershot: Ashgate Publishing Group.

Stanton, N.A., and Annett, J. (2000), Future directions for task analysis. In: J. Annett, and N.A. Stanton (Eds.). *Task Analysis.* London: Taylor & Francis, pp.229-234.

Stanton, N.A. (1999). *Guide to methodology in ergonomics*. London: CRC Press.

Tubbs, J., and Meacham, B. (2009). Selecting appropriate evacuation strategies for super tall buildings: current challenges and needs. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.41-50.

Wilson, J. R. (2005). Methods in the understanding of human factors. In: J.R. Wilson, and N. Corlett, (Eds.), *Evaluation of Human Work. 3rd ed.* London: Taylor & Francis, pp.1-31.

Wood, P.G. (1980). A survey of behaviour in fires. In: D. Canter (Ed.) *Fires and Human Behaviour*. New York: John Wiley and Sons, pp.83-96.

Wright, P.C., and Monk, A.F. (1991). A cost-effective evaluation method for use by designers. *International Journal of Man-Machine Studies,* 35, pp.891-912.

Xu, X., and Song, W. (2009). Staircase evacuation modeling and its comparison with an egress drill. *Building and Environment,* 44(5), pp.1039-1046.

Xudong, C., Heping, Z., Qiyuan, X., Yong, Z., Hongjiang, Zhang., and Chenjie, Z. (2009). Study of announced evacuation drill from a retail store. *Building and Environment,* 44(5), pp.854-870.

Yogesh, P. (2009). *Testing the use of a fire drill scenario developed in Second Life.* MSc thesis. The University of Nottingham.

Zachary Au, S. (2009). How well will your emergency plan work? – A technique to assess human errors and human behaviour in emergency response. In: *Proceedings of the 4th International Symposium on Human Behaviour in Fire.* Cambridge 13-15 July 2009. London: Interscience Communications, pp.443-446.

Zachary Au, S., and Gilroy, J. (2009). Playing safe: ergonomics influences on sporting venues. *The Ergonomist.* March 2009, pp.4-5.

Zhao, L. (1999). A new approach for modeling the occupant response to a fire in a building. *Journal of Fire Protection Engineering,* 10(28), pp.28-38.

Appendix I: Summary of behaviours and influences on behaviour from previous literature

A summary of the main behaviours which were predicted or analysed by the various approaches reviewed in Chapter 2 is shown in Table I – Table III. This gives an indication of what type of behaviours are exhibited in emergency situations. It is only intended as a summary: for full details of the behaviours, in particular the conditions under which they were obtained or can be expected, reference should be made to the original research.

Table I. Behaviours/influences on behaviour associated with a positive effect on evacuation and response to emergencies

Behaviour	Context	Approach	References
Individuals in crowds retain their rationality; social structures exist; altruistic and group-oriented behaviours predominate; panic is unlikely	Crowd behaviours; fires, WMDs	Literature	Pauls and Jones, 1980b; Sime, 1995; Shaw, 2001; Proulx, 2001; Fischer, 2002; Glass and Schoch-Spana, 2002; Drury, 2004; Pan et al., 2006.
People with a higher perceived risk have been seen to respond more quickly or have a greater likelihood of leaving. The level of action taken is appropriate to their perception of the environment.	High rise (WTC); Factory, residential and institutional buildings; Nursing homes	Reports by survivors	Edelman et al., 1980; Wood, 1980; Galea et al., 2009.
Sensory cues have contributed to initiation of evacuation. The cues have included perceiving a fire cue, extensive smoke spread, a higher number of consistent, unambiguous cues, social cues consistent with understanding of fire, cues from an official familiar source	High rise buildings; Factory, residential and institutional buildings; Nursing homes; Community disasters and building fires	Reports by survivors; Literature	Edelman et al., 1980; Wood, 1980; Proulx and Reid, 2006; Gershon et al., 2007; Kuligowski, 2009.
Occupants engaging with second (or subsequent) actions associated with the evacuation are more likely to define the situation as a fire	Community disasters and building fires	Literature	Kuligowski, 2009
Evacuation may be supported by previous emergency preparedness training and experience, and communication of emergency plans to employees	High rise buildings (WTC) Community disasters and building fires	Reports by survivors; Literature	Gershon et al., 2007; Kuligowski, 2009.
Raising the alarm/organising the evacuation is associated with more frequent training or instruction on what to do in a fire	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980

Behaviour	Context	Approach	References
Previous experience of an emergency has been associated with a more rapid response	High rise buildings (WTC)	Reports by survivors	Gershon et al., 2007
In contrast to the above, lack of previous involvement in a fire has been associated with increased evacuation	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
Defining the situation as a fire may increase the likelihood of defining the risk to self/others	Community disasters and building fires	Literature	Kuligowski, 2009
Clear direction from a perceived authority figure has been seen to support evacuation	High rise buildings (WTC)	Reports by survivors	Gershon et al., 2007
Those with a position of responsibility are more likely to act quickly	Building fires	Literature	Proulx, 2007
Cooperativeness in groups who know each has been reported to decrease the time it takes to join the evacuation	High rise buildings	Literature	Aguirre et al., 1998
Smaller groups have been reported to respond quicker than larger groups if the situation is perceived as dangerous	High rise buildings	Literature	Aguirre et al., 1998
Social activities, e.g. cheering, praying have helped an evacuation progress	High rise buildings (WTC)	Reports by survivors	Gershon et al., 2007
Assistance from co-workers and emergency responders have supported evacuation	High rise buildings (WTC)	Reports by survivors	Averill et al., 2009
Physical safety features (e.g. lighting, handrails, reflective tape, floor lighting) may aid evacuation progress	High rise buildings (WTC)	Reports by survivors	Gershon et al., 2007
Complete familiarity with building has been associated with increased evacuation	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
Providing information to evacuees can reduce stress	Public buildings	Literature	Proulx, 1993
Home environment (vs. work environment) has been associated with increased evacuation	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
Gender (female) has been associated with increased perception of fire cues, warning others, leaving the building immediately, requesting assistance, evacuating family, and increased evacuation.	Factory, residential and institutional buildings; Community disasters and building fires	Reports by survivors; Literature	Wood, 1980; Kuligowski, 2009.
Younger age is associated with increased evacuation	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
Speaking the same language as others increases likelihood of perception of a cue, as does frequent interaction with family	Community disasters and building fires	Literature	Kuligowski, 2009
Providing risk information as part of the warning message increases the likelihood of the situation being defined as a fire	Community disasters and building fires	Literature	Kuligowski, 2009
Information from private channels is more likely to be acted upon than public channels;	Nuclear incident	Literature	Kanno et al., 2006

Table II. Behaviours/influences on behaviour which hinder evacuation or response to emergencies

Behaviour	Context	Approach	References
Occupants are often reported to complete non-evacuation activities before evacuating, which may include a combination of: seeking information, collecting belongings, taking emergency equipment, finding a pet, putting on jackets/getting dressed, providing verbal instructions to evacuate, finding children, making phone calls, shutting down computers, securing items, changing footwear, seeking permission to leave or disengaging socially.	High rise buildings (including WTC) buildings; Mixed occupancy residential; University buildings	Reports by survivors; Literature; Fire drills	Proulx, 1995, 2001; Purser and Bensilum, 2001; Gwynne et al., 2003; Proulx and Reid 2006; McConnell et al. 2009.
Alarms may not prompt immediate evacuation; without additional cues such as sight or smell of smoke, hearing shouts of “fire”, or instruction by staff an alarm will rarely trigger evacuation	Nursing home; Building fires; Hospital	Reports by survivors; Analysis of literature; Fire drill	Edelman et al., 1980; Gwynne et al., 2003; Proulx, 2007.
If a group is split up, they may re-form the group before exiting as a whole	Group behaviours; Aeroplane evacuation; WMDs; Retail stores	Literature; Fire drill; Experimental	Muir et al., 1996; Shields and Boyce, 2000; Fischer, 2003; Pan et al., 2006.
Highly individualistic crowd members are more likely to demonstrate non-adaptive behaviours	Building evacuation	Literature	Pan et al., 2006
Individuals are likely to experience higher levels of stress in a crowded situation			
Higher crowd density, environmental constraints (such as dim lighting, too narrow exits, poor signage), and high emotional arousal can cause non-adaptive behaviours			
Evacuees have been reported to act passively in response to an emergency	Subway fire	Reports by survivors	Jeon and Hong, 2009
Commitment to prior activities and ambiguous early cues of a fire may delay response	Public buildings	Literature	Purser and Bensilum, 2001
Information to be processed in an emergency is mainly ambiguous; Ambiguous information causes uncertainty, fear, worry and confusion	Public buildings	Literature	Proulx, 1993
The appearance of a problem will take longer to perceive in an emergency; immediate situations will receive more focus than future scenarios	Building evacuation	Literature	Pan et al., 2006
In aeroplane evacuations, trial participants have been noted stepping on or climbing over others and climbing over the backs of seats to evacuate	Aeroplane evacuation	Experimental	Muir et al., 1996
People who disregarded an emergency have reported a lower perception of risk	High rise buildings (WTC)	Reports by survivors	McConnell et al., 2009
Attempting to save personal effects is associated with less than complete familiarity with the building	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
Unfamiliarity with the building, including locating fire exits and poor signage, has been reported to hinder evacuation	High rise buildings (WTC)	Reports by survivors	Gershon, 2009

Behaviour	Context	Approach	References
Lack of participation in drills is associated with slower evacuation			
Many occupants will have difficulty descending a large number of stairs	Super tall buildings	Literature	Tubbs and Meacham, 2009
Previous experience of a fire incident is reported to decrease the likelihood of leaving immediately	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
The following have all been reported to hinder evacuation:			
Certain footwear (e.g. high heels) Structural damage Debris and glass in the lobby Smoke and water in the stairs Locked doors Lack of managers/leadership Concerns about ability to walk down stairs	High rise buildings (WTC)	Reports by survivors	Gershon et al., 2007
Congestion on stairway Fire-fighter counter flow Slow moving occupants	High rise buildings (WTC)	Reports by survivors	Averill et al., 2009
A reluctance to pass disabled evacuees on escape routes	Retail store	Fire drill	Shields and Boyce, 2000
Poor visibility and lack of recognition of the guide lights	Subway fire	Reports by survivors	Jeon and Hong, 2009
Role: visitors in a building are more likely to wait or expect instructions	Building fires	Literature	Proulx, 2007
Habituation with environment Experience of false alarms Perceptual disability Older age Higher stress/anxiety level Sleeping Higher dose of toxic gases	Community disasters and building fires	Literature	Kuligowski, 2009
Pre-existing social relationships, due to greater efforts to help friends and colleagues. Greater resources available to groups, due to the time required to process the information. Cooperativeness in groups of strangers, as this increases the time it takes to join the evacuation. This may be due to the time it takes to search for meaning in the situation and take action. Interaction with different groups of people.	High rise buildings	Literature	Aguirre et al., 1998

Table III. Other notable behaviours/influences on behaviour in emergencies

Behaviour	Context	Approach	References
Group/social behaviours			
Pre-existing social relationships will continue to exert influence during an emergency	Bioterrorism	Literature	Glass and Schoch-Spana, 2002
Natural leaders have been seen to emerge in emergencies	High rise buildings (WTC)	Reports by survivors	Gershon et al., 2007
Evacuees may copy the behaviour of others, following those who evacuate	Drivers in a tunnel; nuclear incident; building	Fire drill; literature	Boer, 2005; Kanno et al., 2006; Pan et al., 2006.

Behaviour	Context	Approach	References
	evacuation		
Exit choice			
Evacuees are likely to use the main exit, the most familiar or the exit directed by staff; they may not use the optimum evacuation route or evaluate all alternatives.	Nursing home; subway fire; Building fires; Retail stores	Reports by survivors; Literature; Fire drill	Edelman et al., 1980; Shields and Boyce, 2000; Ozel, 2001; Pan et al., 2006; Jeon and Hong, 2009; Xudong et al., 2009.
Evacuees have been reported to use lifts during evacuation	High rise buildings	Reports by survivors	Proulx and Reid, 2006
Congestion may limit an evacuee's ability to chose stairs or escalator	Underground stations	Observation (categorised as fire drill)	Kinsey et al., 2009b
Without smoke, the majority of evacuees in a hotel trial were seen to use the main exit; with smoke the majority used the nearest exit.	Hotel	Fire drill/ laboratory study	Kobes et al., 2009
Decisions tend towards the less risky option under time pressure, which may contribute to the selection of familiar exit routes	Building fires	Literature	Ozel, 2001
At a higher floor level, a larger percentage of people are predicted to chose the lifts rather than stairs; this declines as waiting time increases	High rise buildings	Participant predictions	Heyes and Spearpoint, 2009
A common response to a threat is to seek the proximity of familiar people and places	Disasters	Literature	Mawson, 2005
Pre-movement and egress times			
Pre-movement times for high rise buildings have been reported as 5 minutes (SD=4.7, range=0-30min)	High rise buildings	Reports by survivors	Proulx and Reid, 2006
Mean pre-movement time across four retail stores has been reported as 25 to 37s, maximum pre-movement time ranged from 55 to 100s. Total evacuation times ranged from 131 to 240s.	Retail store	Fire drill	Shields and Boyce, 2000
In another study of retail stores, pre-movement time figures saw: 36% starting evacuation within 60-120s, 13% within 120-180s, and 9% took over 180s to start evacuation	Retail store	Fire drill	Xudong et al., 2009
Pre-movement time has been reported as less than 2 mins for well managed buildings, it could extend to 4 mins.	Offices, shop waiting rooms and assembly spaces	Literature	Purser and Bensilum, 2001
80% of WTC evacuees responded within 9 minutes of impact	High rise buildings (WTC)	Reports by survivors	Galea et al., 2009; McConnell et al., 2009.
With good fire safety management systems, pre-movement times are shorter and have less variation	Offices, shop waiting rooms and assembly spaces	Literature	Purser and Bensilum, 2001
Pre-movement times take a positive skew, with some people taking much longer to evacuate	Offices, shop waiting rooms and assembly spaces	Literature	Purser and Bensilum, 2001
The range of best and worst case egress times for single family houses has been estimated at 2 minutes to 16 minutes 10 seconds.	Single family houses	Literature	Proulx et al., 2006

Behaviour	Context	Approach	References
Travel speeds in stairwells have been reported between 0.2-0.29 m/s; with 85% of participants stopping at least once (43% for congestion, 8% for rest)	High rise buildings (WTC)	Reports by survivors	Averill et al., 2009; Galea et al., 2009.
The total evacuation time of a retail store was recorded at 490s	Retail store	Fire drill	Xudong et al., 2009
Sequences of behaviour			
Perceiving a fire cue leads to either a misinterpretation sequence or an investigation sequence. Seeing smoke leads to one of three preparatory sequences, either instruct, explore or withdraw. After the preparatory sequence, occupants enter one of four "act" sequences, namely evacuate, fight, warn or wait.	Domestic, multiple occupancy and hospital fires	Reports by survivors	Canter et al., 1980
When shopping with accompanier, the most common first action has been reported as "leave immediately", followed by "search for accompanier then leave together"	Retail store	Fire drill	Xudong et al., 2009
Shopping without an accompanier, only 19% were reported to leave immediately as their first action.	Retail store	Fire drill	Xudong et al., 2009
Re-entry to a building is associated with gender: male, daytime, any presence of smoke and previous involvement in a fire	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980
The majority of shoppers have reported little or no commitment to the activity being undertaken at the time of the alarm, and do not complete it.	Retail store	Fire drill	Shields and Boyce, 2000
Specific responses to CBRN incidents			
The first people affected by a chemical or biological agent may not realise it at first; the people affected may treat themselves, go to GP or a hospital.	WMDs	Literature	Fischer, 2003
Trust in information will remain low if it is received from a smaller number of media sources. More reactive behaviour is predicted for elderly people. Farmers with land or animals tend to be reluctant to evacuate. Bad weather is likely to cause people to be reluctant to evacuate. Distance from loudspeakers and weather will affect information acquisition. Evacuation instructions will increase the recognition of urgency. Seeing and hearing ambulance or fire engines will increase recognition of an incident. Men are less likely to pass on obtained information than women. People associated with others who need care or help are expected to evacuate earlier.	Nuclear incident	Literature	Kanno et al., 2006
Rioting and looting would be rare following a WMD.	Bioterrorism	Literature	Glass and Schoch-Spana, 2002

Behaviour	Context	Approach	References
Compliance with instructions by authorities			
Approximately half to two-thirds of the population would refuse to evacuate following a WMD	WMDs	Literature	Fischer, 2003
Most of the population would cooperate with quarantine orders; however many will be outside the city area before symptoms appear; others will evacuate before the announcement to quarantine is made; some will successfully evade the quarantine.	WMDs	Literature	Fischer, 2003
Delay in symptoms may mean people treat themselves, go to a GP or to hospital	WMDs	Literature	Fischer, 2003
70-80% compliance has been predicted with an order to evacuate. Higher compliance is expected with an order to shelter at home or evacuate from work. Lower compliance is expected with an order to evacuate from home or shelter at work. 10% reduction in compliance is expected if the media are sceptical. Seeing or hearing the explosion is not predicted to have a great effect on compliance rates. 10-15% improvement in compliance could be achieved through workplace drills. 60-70% compliance is expected with an order to shelter in the current location.	Radiological device	Expert predictions	Dombroski et al., 2006
Occupants may not follow instructions to stay in place, particularly if they have seen other people evacuating.	Super tall buildings	Literature	Tubbs and Meacham, 2009
Individual factors			
Fighting the fire or minimising the risk is associated with previous experience of a fire incident. Gender (male) is associated with minimising the risk or fight the fire. Men are also more likely to move through smoke. Gender (female) is associated with warning others, leaving immediately, requesting assistance, or evacuating family. Previous experience of a fire incident is associated with fighting the fire and minimising the risk.	Factory, residential and institutional buildings	Reports by survivors	Wood, 1980

Appendix II: Consent form and post trial questionnaires used in the fire drill evacuation in the standardised comparison

CONSENT FORM: POST-EVACUATION

My name is Glyn Lawson and as part of my research I am investigating the behaviours and actions people take during Fire Drills. During today's Fire Drill, your behaviours have been captured on video camera. I would like to analyse them as part of my research, but to do so I need your consent. I would also like to analyse the data from the various questionnaire you have completed.

Although I hope to publish the results of my research, it will be entirely anonymous – your name will never be associated with the study. Only researchers within the Human Factors Research Group working on this project will be able to gain access to the video footage or questionnaires, which will be stored securely in a locked cabinet.

Please tick the appropriate box below and sign your name. Please also answer some basic questions about your appearance so I can identify you in the video footage.

Remember, the video footage and questionnaire data will only be used to support analysis of the trial. It will be stored securely in a locked cabinet. It will always be treated as anonymous data.

Please note that you will receive your £10 gratuity payment whichever box you tick.

☐ I confirm that I have read and understood the above description of the study and **agree** for the video footage of my actions and questionnaire data to be analysed.

☐ I confirm that I have read and understood the above description of the study and **do not agree** for the video footage of my actions or questionnaire data to be analysed.

Name:... Date:.................….............

Participant ID number :....................... Age (optional):……………

Email................................ Degree course:..

How long have you studied at the University of Nottingham:.....................years

Please give a few details about your appearance to help me identify you in the video footage:

Gender: Male/female

Hair colour:…………………………. Hair length/style:……………

Type of top (e.g. collared shirt, t-shirt, jumper, jacket): …………………

Colour/any distinguishing logos: ……………………………..

Were you carrying a bag? If so, please describe it…………………………

Any other distinguishing aspects of your appearance which may help identify you:…………………………..

To receive your payment, I need you to fully complete all of the forms on the following pages.

In the following table, I would like you to describe what actions you took after hearing the fire alarm. Please be very detailed: for example, you might write "I walked to the door of the computer room", or "I put on my jacket". Please also write the location for each action. For example, this could include "computer room", "corridor", "stairwell", "foyer", "outside building".

Please write in this column what actions you took after hearing the fire alarm. Please write them **in order**, starting with the first action you took.	Please write in this column the location where the action took place (e.g. computer room, corridor...)
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	
11.	
12.	
13.	
14.	
15.	
16.	
17.	
18.	
19.	
20.	

Please review your answers to check they are complete. For example, if you put on clothes, or took any belongings, please check that these are included in the list. Add rows if necessary.

Please answer all the following questions:

1. How familiar are you with the Psychology building? (please circle the appropriate number)

Not at all familiar				Very familiar
1	2	3	4	5

2. Please rate how influential each of the following was on your choice of exit from the building (please circle the appropriate number):

a. Familiarity with route

Not at all influential				Very influential
1	2	3	4	5

b. The behaviour of other participants

Not at all influential				Very influential
1	2	3	4	5

c. Instruction from authority figure

Not at all influential				Very influential
1	2	3	4	5

d. Emergency exit signs

Not at all influential				Very influential
1	2	3	4	5

e. Distance (i.e. the shortest route to exit the building)

Not at all influential				Very influential
1	2	3	4	5

f. Other (please specify):..

Not at all influential				Very influential
1	2	3	4	5

3. Do you remember seeing any emergency exit signs during your evacuation? If so, please indicate where they were (tick all that apply).

☐ In the computer room A5
☐ Hanging from roof in basement
☐ Above the first set of double doors in the corridor
☐ On door into parking area
☐ On the door leading to the corridor stairwell exit

Question 3 continued overleaf...

Question 3 continued...

- ☐ In the exit stairwell (accessed from the corridor)
- ☐ Above the second set of double doors leading to the reception area
- ☐ Other (please specify):

4. How familiar are you with the emergency response procedures for your school? (Please circle):

Not at all familiar				Very familiar
1	2	3	4	5

5. How closely did you follow the emergency response procedure for the school? (Please circle):

Not at all				Followed it exactly
1	2	3	4	5

6. Please rate your perception of danger when you first heard the alarm:

No danger				Great danger
1	2	3	4	5

7. Please rate your perception of danger when you decided to leave the computer room:

No danger				Great danger
1	2	3	4	5

8. Please rate your perception of danger as you exited the building:

No danger				Great danger
1	2	3	4	5

9. How long do you think it took to leave the computer room?......................

10. How long do you think it took to leave the building?.......................

Many thanks! Remember that your responses will be treated anonymously, and the data will be stored confidentially. Please return your form to the experimenter, Glyn Lawson.

Appendix III: Consent form and post trial questionnaires used in the VE study in the standardised comparison

CONSENT FORM

Investigating behaviour in Emergency Situations

My name is Glyn Lawson and as part of my research within the Human Factors Research Group I am carrying out a study into behaviours demonstrated in virtual environments. In this experiment your actions and conversations will be recorded in the Second Life virtual environment in response to certain nominal and emergency situations.

Your name will not be used in conjunction with the study and all information you provide will be treated anonymously. Your data will be recorded (and stored on a secure computer) as Participant X, not your real name. I hope to publish the results of my work, but it will not be possible to identify you from any data used in reports. The raw data will be destroyed following completion of my PhD research.

This consent form is the only record of your name, and it will be stored in a locked cabinet will also be destroyed following completion of the PhD.

The experiment will not take more than half an hour and you will be paid £5 for your time. If at any time during the experiment you feel distressed, upset or any other unwanted emotions please indicate this to me, and the trial will be ended. Similarly, if you feel any adverse effects of using the virtual environment, please stop immediately. You will be paid for your time regardless.

Participation in the study is voluntary and you may stop at any point if you do not wish to continue.

I confirm that I have read and understood the above description of the study and agree to take part.

Name (sign):..(print):..

Date:.................…............

SECONDLIFE VIRTUAL ENVIRONMENT BUILDING EVACUATION STUDY: QUESTIONNAIRE

Please complete the following form. Please note that all answers will be treated anonymously, i.e. your data will be recorded as "participant X".

Name:	
Date:	
Name of avatar:	
Age (optional):	
Email:	
Degree course:	
How long have you studied at the University of Nottingham:	
Gender:	Male/female

To receive your payment, I need you to fully complete all of the forms on the following pages.

In the following table, I would like you to describe what actions you took after hearing the fire alarm. Please be very detailed: for example, you might write "I controlled my avatar to the door of the computer room". Please also write the location of your avatar for each action. For example, this could include "computer room", "corridor", "stairwell", "foyer", "outside building".

Please write in this column what actions you took after hearing the fire alarm. Please write them **in order**, starting with the first action you took.	Please write in this column the location where the action took place (e.g. computer room, corridor...)
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	
11.	
12.	
13.	
14.	
15.	
16.	
17.	
18.	
19.	

Please review your answers to check they are complete. Add rows if necessary.

Please answer all the following questions:

1. How familiar are you with the Psychology building? (please circle the appropriate number)

Not at all familiar				Very familiar
1	2	3	4	5

2. Please rate how influential each of the following was on your choice of exit from the building (please circle the appropriate number):

a. Familiarity with route

Not at all influential				Very influential
1	2	3	4	5

b. The behaviour of other participants

Not at all influential				Very influential
1	2	3	4	5

c. Instruction from authority figure

Not at all influential				Very influential
1	2	3	4	5

d. Emergency exit signs

Not at all influential				Very influential
1	2	3	4	5

e. Distance (i.e. the shortest route to exit the building)

Not at all influential				Very influential
1	2	3	4	5

f. Other (please specify):...

Not at all influential				Very influential
1	2	3	4	5

3. Do you remember seeing any emergency exit signs during your evacuation? If so, please indicate where they were (tick all that apply).

- ☐ In the computer room A5
- ☐ Hanging from roof in basement
- ☐ Above the first set of double doors in the corridor
- ☐ On door into parking area
- ☐ On the door leading to the corridor stairwell exit
- ☐ In the exit stairwell (accessed from the corridor)
- ☐ Above the second set of double doors leading to the reception area
- ☐ Other (please specify):

4. How familiar are you with the emergency response procedures for your school? (Please circle):

Not at all familiar				Very familiar
1	2	3	4	5

5. How closely do you think your actions in the virtual environment followed the emergency response procedure for the school? (Please circle):

Not at all				Followed it exactly
1	2	3	4	5

6. Please rate your perception of danger when you first heard the alarm *:*

No danger				Great danger
1	2	3	4	5

7. Please rate your perception of danger when you decided to leave the computer room:

No danger				Great danger
1	2	3	4	5

8. Please rate your perception of danger as you exited the building:

No danger				Great danger
1	2	3	4	5

9. How long do you think it took your avatar to leave the computer room?......................

10. How long do you think it took your avatar to leave the building?.......................

11. Please rate your experience of computer gaming:

Not at all experienced				Very experienced
1	2	3	4	5

Many thanks! Remember that your responses will be treated anonymously, and the data will be stored confidentially. Please return your form to the experimenter, Glyn Lawson.

Appendix IV: Consent form and post trial questionnaires used in the talk-through study in the standardised comparison

CONSENT FORM

Investigating behaviour in Emergency Situations

My name is Glyn Lawson and as part of my research within the Human Factors Research Group I am carrying out a study into whether people are able to predict accurately how they would behave during an emergency situation. This experiment will involve me asking how you would behave in a particular scenario. You will be asked what you would do, and your responses will be used for subsequent analysis. Your name will not be used in conjunction with the study, and all information you provide will be treated anonymously. Your data will be recorded (and stored on a secure computer) as Participant X, not using your real name. While I hope to publish from this work, it will not be possible to identify you from any data used in reports. The raw data will be destroyed following completion of my PhD research.

This consent form is the only record of your name, and it will be stored in a locked cabinet and will be destroyed following completion of my PhD research.

The experiment will not take more than half an hour and you will be paid £5 for your time. If at any time during the experiment you feel distressed, upset or any other unwanted emotions please indicate this to me, and the trial will be ended. You will be paid for your time regardless.

Participation in the study is voluntary and you may stop at any point if you do not wish to continue.

I confirm that I have read and understood the above description of the study and agree to take part.

Name:...

Date:..................…............

Age:................................

EMERGENCY EVACUATION STUDY: QUESTIONNAIRE

Name:	
Date:	
Participant ID:	
Age (optional):	
Email:	
Degree course:	
How long have you studied at the University of Nottingham:	
Gender:	Male/female

I would like you to imagine that you are working in the A5 computer lab (shown on building plan below). Imagine there are about 20 people in the lab in total. It is a weekend morning. Imagine you have not yet logged in, and are completing a paper-based personality questionnaire, as part of a job application.

I would like you to describe what actions you would take after hearing the fire alarm. Please be very detailed, I may probe you for more detail if necessary.

Please also tell me the location for each action. For example, this could include “computer room”, “corridor”, “foyer”, “outside building”.

What actions would you take after hearing the fire alarm? Please state them **in order**, starting with the first action you would take.	Where would each action take place? (e.g. computer room, corridor...)
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	
11.	
12.	
13.	
14.	

1. How familiar are you with the Psychology building? (please circle the appropriate number)

Not at all familiar				Very familiar
1	2	3	4	5

2. Please rate how influential each of the following would be on your choice of exit from the building (please circle the appropriate number):

a. Familiarity with route

Not at all influential				Very influential
1	2	3	4	5

b. The behaviour of other participants

Not at all influential				Very influential
1	2	3	4	5

c. Instruction from authority figure

Not at all influential				Very influential
1	2	3	4	5

d. Emergency exit signs

Not at all influential				Very influential
1	2	3	4	5

e. Distance (i.e. shortest route to exit building)

Not at all influential				Very influential
1	2	3	4	5

f. Other (please specify):..

Not at all influential				Very influential
1	2	3	4	5

3. Are you aware of any emergency exit signs in the building? If so, please indicate where you think they are (please tick all that apply).

- ☐ In the computer room A5
- ☐ Hanging from roof in basement
- ☐ Above the first set of double doors in the corridor
- ☐ On door into parking area
- ☐ On the door leading to the corridor stairwell exit

☐ In the exit stairwell (accessed from the corridor)
☐ Above the second set of double doors leading to the reception area
☐ Other (please specify):

4. How familiar are you with the emergency response procedures for your school? (Please circle):

Not at all familiar				Very familiar
1	2	3	4	5

5. How closely do you think your predicted actions follow the emergency response procedure for the school? (Please circle):

Not at all				Followed it exactly
1	2	3	4	5

6. Please rate your anticipated perception of danger when the alarm first sounded:

No danger				Great danger
1	2	3	4	5

7. Please rate your anticipated perception of danger at the point when you decided to leave the computer room:

No danger				Great danger
1	2	3	4	5

8. Please rate your anticipated perception of danger when you exited the building:

No danger				Great danger
1	2	3	4	5

9. How long do you think it would take you to leave the computer room?

10. How long do you think it would take to leave the building?