THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

A Human Factors Evaluation of Auditory Displays in Medical Electrical Equipment

Alexandra Nai-Khee Wee
Bachelor of Information Technology (Honours)
Diploma of Information Technology

A thesis submitted for the degree of Master of Philosophy at
The University of Queensland in 2014

School of Information Technology and Electrical Engineering

Abstract

Auditory alarms have, for the most part, gone unchanged in human technological developmental history. This is largely due to their success and simplicity. As they take advantage of the primitive human principle of fight or flight, they have been found to be a significant attractor of attention, are generally easy to understand, omi-directional, and to the point.

Medical staff, for instance, use sounds in many aspects of patient monitoring—the steady beeping of the pulse oximeter, or insistent alarms from equipment when a patient is going critical. However, due to the complex nature of patient care, the sounds that demand the most attention are not always the most useful, and are sometimes detrimental to work, causing annoyance and confusion.

Thus the problem is determining the best way to alert medical and nursing staff to changes in a patient's state, provide them with enough information about this change when needed, and not affect their ability to address it and other immediate concerns. For about three decades, the design of effective auditory alarms and displays in the operating room and intensive care unit (ICU) environment has received a significant amount of attention from human factors researchers. Contributing studies on the physical limitations of the auditory channel (Neurobiology and Psychoacoustics), how this information is processed (Cognitive Psychology), and how best to represent the information (Display Engineering) provide a valuable source of information which should inform the design of such displays. How much effect this has had on the industry, however, is debatable, as the design and development of the majority of auditory displays currently in operating rooms seems to have largely ignored these findings.

Instead, human factors researchers have found it necessary to perform their evaluations on an ad hoc basis. Often well after the medical systems have been produced and put into use.

This thesis presents the findings of a study involving a set of informative melodic medical alarms which has been presented as a design standard for implementation in the International Electrotechnical Commission standard for Medical Electrical Equipment. The goal in providing this design was to provide alarms which not only direct attention, but

convey simple information encoded within the melody, which communicates to medical practitioners which organ system has entered an abnormal state.

In the first 2 experiments, the practicability of the presented alarms as an informative auditory display was evaluated on both novice and professional (practicing nurses) populations. The use of mnemonic learning aids was also examined as these were central to the fundamental design intention of the alarm melodies. The results confirmed participant testimonies that the provided mnemonics did not aid them in identifying alarm labels when presented with the alarm sound, even after prolonged testing over 2 days. Expert knowledge in the domain and knowledge of the context of the menmonics also did not provide any performance benefit on accuracy which was poor in both populations. However, participants' prior exposure to formal musical training was found to greatly affect their accuracy in identification of presented stimuli.

Having established the ineffectiveness of the provided standard to improve on existing auditory alerts in support safety critical patient monitoring by communicating meaningful information to medical practitioners, it was necessary to identify what aspects of the melodic alarms were the source of the poor performance produced by the participant pool. This exploration was conducted in an effort to isolate useful or disadvantageous aspects of auditory display design to provide guidelines and suggestions for future informative alarm development.

In analysing the data from the previous experiments it was determined that participant selection errors were non-random events. When a particular stimuli was played participants immediately eliminated those sounds they knew were not correct, leaving them with 2 or 3 possibilities to choose from. These patterns were quite consistent across all participants regardless of any prior musical experience. The alarm groups were also classified into melodic shapes which appeared to be the major cause of participants' inability to accurately identify the alarms.

A subsequent study was set up to test this phenomenon. A confusion cluster was decoupled (by removing one half of the confusion pair from the alarm set) and the experiment was run again to ascertain if the overall accuracy of the alarm left would be improved by a value significantly better than chance.

A final study was then run to explore if the alarm recommendations, as suggested by the IEC standard, could in any way be stretched to provide a useful set of informative alarms, given the strict limitations on musical variation. A frequently confused alarm was re-designed within the specifications outlined in the standard but in such a way that it would be as different to the other alarms as the standard would allow. The results of this study showed that participants performed no better than in previous experiments when prior musical training was taken into account.

This result provides strong support to the view that the IEC recommended standard for melodic alarms in medical electrical equipment, with its current design limitations, will not produce a set of alarms which can provide any informative benefits to medical practitioners over and beyond current arbitrary alerting alarms that simply direct attention to the piece of equipment producing the sound.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the *Copyright Act 1968*.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Conference Papers

Wee, A., and Sanderson, P. (2005, 12-16 September). Testing New Alarms for Medical Electrical Equipment. Paper presented at the Tenth IFIP TC13 International Conference on Human-Computer Interaction (Interact2005), Rome. - Ms Wee was responsible for 90% of drafting and writing, conducting the entire participant study and presentation of the paper. Prof Sanderson supervised the conception and design of the study and was responsible for 30% of the analysis and interpretation of data as well as providing feedback of the draft material.

Sanderson, P. M., Seah, E., Lacherez, P., Wee, A., Thompson, C., and Watson, M. (2006). Melodic medical equipment alarms: Are they safe? (Abstract). Proceedings of the Annual Scientific Meeting of the Australian and New Zealand College of Anaesthetists (ANZCA ASM 2006), 13-17 May, Adelaide, SA. – Prof. Sanderson was responsible for drafting and writing as well as supervision of the overall research program. Ms. Wee conducted 2 of the experimental studies described in the paper including 80% of the analysis and interpretations as well as provided additional assistance to Ms. Seah in conducting her contributed study. Dr. Lacherez provided technical help in setting up the experimental computer program.

Wee, A., and Sanderson, P. (2006). Effects of mnemonics in learnability of melodic alarms with registered nurses. Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society. 16-20 October, San Francisco, CA. - Ms Wee was responsible for 90% of drafting and writing, conducting the entire participant study and presentation of the paper. Prof Sanderson supervised the conception and design of the study and was responsible for 30% of the analysis and interpretation of data as well as providing feedback of the draft material.

Sanderson, P., Wee, A., Seah, E., and Lacherez, P. (2006). Auditory alarms, medical standards, and urgency. Proceedings of the 12th International Conference on Auditory Display (ICAD2006). Queen Mary University of London. 20-23 June. - Prof. Sanderson was responsible for drafting and writing as well as supervision of the overall research program. Ms. Wee conducted 2 of the experimental studies described in the paper including 80% of the analysis and interpretations as well as provided additional assistance to Ms. Seah in conducting her contributed study. Dr. Lacherez provided technical help in setting up the experimental computer program.

Journal Articles

Wee, A., and Sanderson, P. (2008). Are melodic medical equipment alarms easily learned? Anesthesia and Analgesia, 106(2), 501-508. - Ms Wee was responsible for 100% of drafting and writing, conducting the entire participant study and presentation of the paper. Prof Sanderson supervised the conception and design of the study and was responsible for 25% of the analysis and interpretation of data as well as providing feedback of the draft material.

Sanderson, P., Wee, A., and Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. Anaesthesia, 61, 142-147. - Ms Wee was responsible for 90% of drafting and writing, conducting the entire participant study. Prof Sanderson supervised the conception and design of the study and was responsible for 30% of the analysis and interpretation of data as well as providing feedback of the draft material. Dr. Lacherez provided technical help in setting up the experimental computer program.

Publications included in this thesis

No publications included.

Contributions by others to the thesis

Some of the Csound programming to create the stimuli as well as some of the experiment coding was carried out by Philippe Lacherez and Phil Cole at the University of Queensland, Australia. Philippe wrote the Csound files that contained the basic parameters of the sound stimuli. Philippe, Phil and I coded the experiment program to present the stimuli to participants and collect the data of their responses. Professor Penelope Sanderson recommended the initial statistical procedure and assisted with carrying out the analysis and interpreting of the results of the first two studies. Dr. Richard Morey provided statistical advice on the latter two experiments.

Collection of responses for Experiment 3 was carried out by Pennie Gibbins at the University of Newcastle. Collection of responses for half of the participants in Experiment 4 was carried out by Babette Rae also at the University of Newcastle.

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

The research presented in this thesis could not have been completed without the support of some key people. Firstly, I would like to thank my parents, Michael and Cheng Lee for giving me this extraordinary opportunity for further studies, as well as for being patient with my distractions from work and completing of this research.

I would also like to thank the people who have provided research insight and guidance; my supervisors Dr. Mikael Boden, A/Prof Scott Brown, Professor Penelope Sanderson for their extreme patience, advice, encouragement and inspiration. My immense gratitude and admiration to Dr. Richard Morey for help in understanding statistics and making me consider alternatives to the techniques that I have sometimes stubbornly clung to.

For their patience and goodwill, the staff of the University of Queensland Usability Labs (UQUL) for putting up with requests for resources in the early days of my research.

My thanks to the staff and researchers at the University of Newcastle, especially Prof. Andrew Heathcote, for allowing me to run my studies using their resources and participant pool. Special thanks to Pennie Gibbins and Babette Rae helping to run participants in as well as allowing me to engage in discussions about the work, solidifying ideas and asking questions which helped me clarify my explanations.

I was able to undertake this research with financial support from a University of Queensland Postgraduate Award Scholarship. I am also grateful to the research ethics council and staff of Queensland Health for allowing studies to be conducted on site and encouraging staff participation.

The greatest contributor both in the form of moral support and actual labour assistance has come from my husband and best friend, Derek Morey. Without whom I would have given up on this work in frustration. His insightful questions, expertise in data analysis and presentation and countless reviews of my work always brought me back from confusing tangents and helped convey the issue as clearly as I could. And for my son Darius, this work has been submitted in spite of your efforts of distraction. And lastly for my soon-to-arrive daughter, Aurelia, thank you for waiting until this work has been submitted.

The errors, idiocies and inconsistencies of this work however, remain my own.

Keywords

auditory alarms, displays, iec60601-1-8, learning, medical, melodic, mnemonics

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 080602 Computer-Human Interaction 100%

Fields of Research (FoR) Classification

FoR code: 1702 Cognitive Sciences 100%

Table of Contents

List of Figures

List of Tables

List of Abbreviations used in the thesis

Acronym	**Definition**
ANZCA	The Australian and New Zealand College of Anaesthetists
ANZSRC	Australian and New Zealand Standard Research Classification
ASM	Annual Scientific Meeting
ASTM	American Society for Testing and Materials
AUD	Australian Dollar
CEN	The European Committee for Standardization (CEN, French: Comité Européen de Normalisation)
CV	Cardiovascular Alarm
GE	General Alarm
GHZ	Gigahertz
HSD	Honestly Significant Difference
ICU	Intensive Care Unit
IEC	International Electrotechnical Commission
IEEE	The Institute of Electrical and Electronics Engineers
IN	Infusion Alarm
ISO	International Organization for Standardization
JCAHO	Joint Commission on Accreditation of Healthcare Organisations

Acronym	Definition
JWG	Joint Working Group
JZS	Jeffrey-Zellner-Siow (Bayes Analysis)
LTM	Long Term Memory
MRI	Magnetic Resonance Imaging
NBC	National Broadcasting Company
OX	Oxygen Alarm
PE	Perfusion Alarm
PF	Powerfail Alarm
SPL	Sound Pressure Level
TE	Temperature Alarm
VN	Ventilation Alarm (Note that in Appendices this has been referenced as VE as a result of analysis conducted earilier)
UQUL	The University of Queensland Usability Laboratory

1 INTRODUCTION

This thesis is a presentation of a research program that evaluates the effectiveness of a set of alarms currently being used in practice. It identifies what aspects contribute to the effectiveness or ineffectiveness of the alarms based on the characteristics of the sounds themselves, as well as contributing attributes of the individuals who would be expected to use the alarms to aid them in performing a safety critical task.

This thesis focuses on the problems of alarms in the operating theatre and the applications of melodic organ groupings of alarms to reduce some of the problems that are attributed to alarms in high stress, safety critical work environments.

Before we begin looking at the specifics of auditory alarms, it is important to understand why we use them and auditory alarms have useful properties that are beyond supporting a visual equivalent. The auditory input channel is useful for conveying critical information because hearing is a primary warning sense (Patterson, 1990). The omnidirectional and obligatory properties of auditory displays and alerts make them very useful for increasing performance in vigilance or monitoring tasks; characteristics not shared by their visual counterparts.

Over the last few decades, the demands of many (increasingly) time critical tasks have seen the advent of new and often varied machines and equipment; all in the name of helping us keep better informed. Because of this, a new problem in vigilance and monitoring has arisen. Woods (1995), citing Lees (1993) identified several factors in auditory alarm systems that contribute to difficulties in fault management. These are: nuisance alarms (persistent, often non-critical alarms), ambiguous or underspecified alarms messages, alarm inflation (sudden increase in the number of alarms heard) and alarms that indicate expected events rather than system anomalies.

In ordinary day to day situations, alarms can cause irritation and annoyance, such as when a car alarm sounds at two o’clock in the morning, or when an over-sensitive fire alarm triggers unnecessary evacuations in a building and disrupting productivity. However, in safety critical domains the “Alarm Problem”, as Woods (1995) calls it, can cause distraction or disruption in fault management and identification tasks, especially during periods of high cognitive load and task criticality.

The alarm problem consists of several issues in which alarms systems can impair a person's ability to do his job effectively and efficiently. These include:

- Too many false alarms;
- Overly loud and intrusive alarms;
- Alarms so quiet that they are missed;
- Alarms that indicate a known event;
- An overload of alarms sounds which leads to confusion; and
- Excessive reliance on alarms when better information signals might be more appropriate.

It is for this reason that for over two decades there has been a lot of interest in the research into the management of auditory alarms in such domains as air traffic control, operating rooms, intensive care units (ICU) and power stations.

1.1 Research Aims and Rationale

This research program emerged from the observation that most, if not all, patient monitoring and hospital equipment are fitted with auditory alarms to direct operator attention to unacceptable values. A major drive behind this is that medical machinery developers are avoiding the potential for litigation in the high risk realm of patient safety. Developers seldom consider the human factors consequences alarm systems in the highly variable domain of medical equipment. It is important to note that prior to start of the 21st century, manufacturers developed many of their machines in isolation, as they often had no standard to operate in, and no fore-knowledge of what other equipment may be in use in the same vicinity. This combination of independently developed systems with a wide variety and frequency of alarms from numerous equipment, combined with the background noises and distractions of a highly trafficked environment, makes it difficult for medical personnel to be certain of what an alarm is indicating (Watson and Sanderson 2003), thus failing the primary intent of the designers who incorporated these alarms in their equipment designs in order to direct attention to the intended area. As a consequence, the ambiguity of these alarms systems is considered a nuisance or distraction to medical staff and the people around them.

Some domains such as aerospace (Sarter, 2000) and nuclear power stations (Kemeny, 1979) have had some degree of human factors techniques applied to the design of both the visual and auditory displays. This is largely due to the concern that errors and accidents which result from these domains would cause large scale incidents which exceed the general population's risk appetite.

On the other hand, the field of medical equipment and design has largely relied on engineering input as the driving force of their designs and has not applied many lessons learnt from the field of Human Factors or Cognitive engineering. The latest standard for medical electrical equipment reflects this engineering driven bias.

In 2000 the ISO and IEC convened a Joint Working group on alarm signals to draft the IEC 60601 standard for medical electrical equipment. This full standard was published in 2003 and 2006. Included in all 3 variations, were a set of melodic alarms differentiated by physiological systems in order to more easily convey information to users.

The underlying goal of the alarms was to make use of the ability to encode additional information within the alarm sound (Block, 1994), in this case the identity of the untoward physiological parameter. In order to achieve this goal, the alarm has to attract the focus of the users (clinical and medical staff), be able to convey urgency, which is what traditional alarms currently do. In addition to this the physiological system to which the alarm refers to should also be readily identifiable by trained users. Identification of the alarms should also not add any undue workload onto users who already operate under heavy cognitive and mental workloads.

Prior to the studies conducted in this thesis (and published as a result of), there appeared to be no indication that the alarms as proposed in these standards have been formally tested to ascertain if they satisfy the abovementioned requirements in the areas of general usefulness, learnability, discernability and memorability. This would seem to be a considerable oversight given that the proposed alarms are part of an international standard for a safety critical domain.

The experimental program consists of a series of experiments that assesses the performance of novice users and medical staff while using the melodic alarms proposed in IEC 60601-1-8 in association and recall tasks.

Performance measures include:

- Accuracy of identification;
- Time taken for participants to become competent with alarm identification;
- Time taken to respond to presented stimuli; and
- Ability to correctly and quickly identify the alarms whilst performing a cognitively engaging task.

The aim of the research presented in this thesis was to develop a better understanding of which aspects of learning and musical qualities would lead to a more useful and informative set of auditory alarms. In particular:

- To examine the effect of various learning conditions and expertise which may enable better performance in alarm identification and reaction time;
- To determine the necessary attributes an individual needs to possesses in order achieve better performance in melodic alarm identification;
- To determine if the IEC 60601-1-8 guidelines provide the most effective design for melodic alarms; and
- To recommend design conditions and properties to improve the existing design, thereby increasing the effectiveness of the melodic alarms.

In the process of achieving the abovementioned aims, several secondary investigations will also be carried out:

- The effect of training, learning aids and learning conditions
- The effect of musical and sound properties in creating confusion.
- Using lessons learnt to improve melodic alarms in an attempt to achieve greater accuracy and confidence in identification.

1.2 Developmental Structure of the Thesis

The thesis argument has been structured with the aid of Goal Structuring Notation (GSN).

GSN is an argument or logic based methodology that traditionally represents all aspects of a safety argument in an elegant and logical diagram (Kelly & Weaver, 2004). This includes all requirements, claims, evidence and context. GSN has become a well-respected and often used technique in the safety critical system industry to present comprehensive evidence supported safety arguments which is readily understood between stakeholders. GSN has found use in a number of significant projects and across a number of discrete domains, including avionics, railways, marine and air traffic control.

In constructing this thesis argument, I found this method to be useful to organise and structure the various supporting literature, hypothesis, strategies and outcomes, because the notation was able to capture the relationships between the thesis elements and present a coherent argument where all claims and hypothesis presented could be easily traced through and communicated to the reader or examiner.

The technique itself is not difficult to comprehend, and has gained some recognition in industry practice for the management of safety-related elements (including Human Factors) requirements in safety critical systems(Rich, Blanchard, & McCloskey, 2007). However, due to its specific terminology and lack of general use outside of the Safety Case development environment, I believe it is necessary to briefly explain the symbology and structure of GSN.

1.2.1 GSN Symbology

The reason for using this notation is to demonstrate that a top-level claim is valid. A top level claim is supported by an argument, and is broken down into a number of further sub-claims until the point is reached where a sub-claim can be supported directly by evidence.

A general example of GSN Symbology is presented in Figure 1-1. The purpose of the dissertation or goal of the thesis is presented as a top level *Goal (*Goal 0*)*. This can be broken down into sub goals which combined meet the requirements of the top goal. Sub goals can subsequently be broken down further until they can be practically and directly supported by prior research, experimental testing or logical reasoning. The *Strategy* is the plan and the method by which the goal is to be tackled. A *strategy* needs to be supported in some why via some rationale, *assumption* or *justification*. Goals and Strategies may be based on known or established research or common practice based on prior industry or academic literature which are presented as *Context*. Ultimately the structure of the argument needs to be supported with *evidence* in the form of literature reference, analytical test, experimental or statistical results.

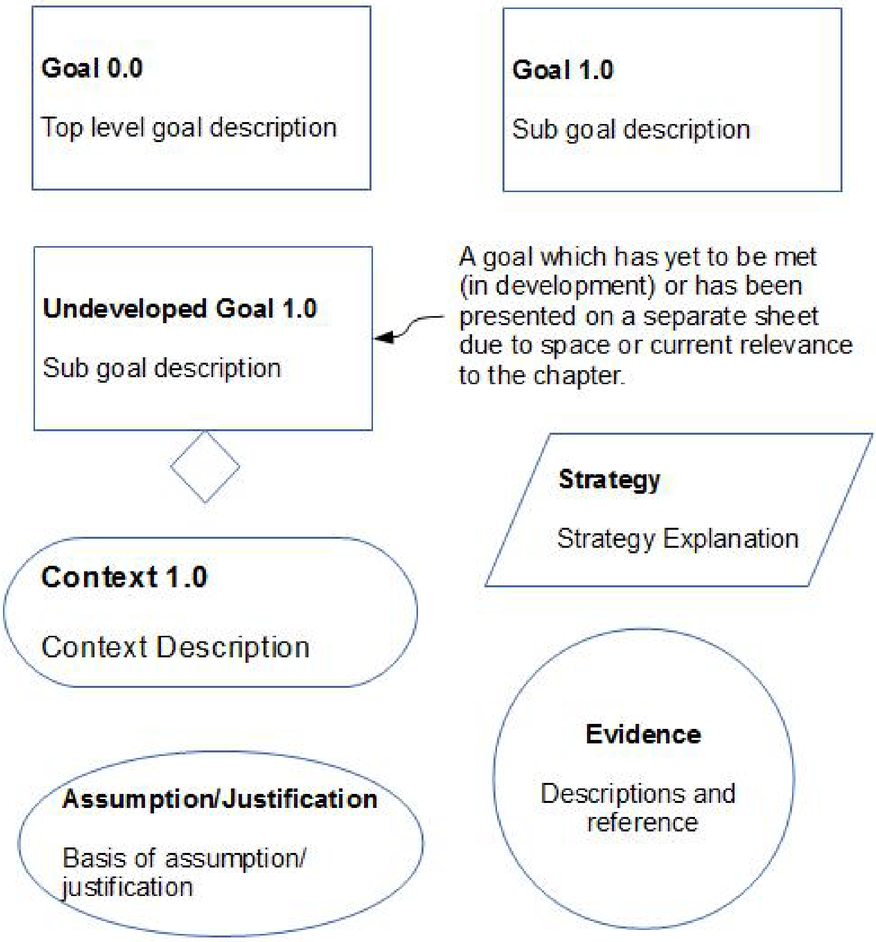


Figure 1-1: GSN Symbology

1.3 Written structure of the thesis

In this Chapter, I have identified and described the problem that provided the motivation for this research program as well as detailed the developmental technique used to develop the structure of the research program. Although the issues associated with the development of useful alarms exists in many complex safety critical domains, I have chosen to focus my work in the domain of medical alarms due to the body of research available in the literature and the availability of auditory alarm standards for medical equipment.

In Chapter 2, I provide a review into the literature, first by presenting an overview of the broad context of related scientific disciplines and narrowing it down to highlighting the usefulness and problems of alarms as an action or activity directing mechanism. I will also review some studies that have aimed to reduce or eliminate the alarm problem by the use of alternative forms of auditory displays such as continuous auditory streams, voice alarms, intelligent or knowledge based alarms and auditory icons (earcons).

In Chapter 3, I provide an overview of alarms in this medical environment and the history and rationales as to the current state of the alarm standard based on research and personal communication with members of the IEC 80801-1-8 standards committee.

Chapter 4 introduces the research program. This includes the motivation and aims of the thesis and describes the studies that were run, as well as list and detail the rational for the hypotheses that were tested.

Chapters 5 to 7 details the subsets of the experiments that were meant to test the hypotheses listed in Chapter 4. The results of the individual studies are analysed and presented here.

In the final 2 chapters, I discuss the overall conclusions as they relate to responding to the questions that were posed in Chapter 4.Also included are further implications of the results and considerations of possible directions in which the research may take.

2 RESEARCH CONTEXT

The following sections present an overview of the field of enquiry from the broad context of the related scientific disciplines to a more focused examination of the specific domain in which the research is contained. The multidisciplinary nature of the research question requires that the problem be approached with an understanding of the underlying fundamentals of Physiology, Psychology and Design in order to build onto the higher level investigations of Auditory Alarm Design in the Patient Monitoring Environment.

Figure 2-1[1] presents the relationship between the high level scientific disciplines and how it relates to the research question as discussed in this chapter and provides a useful roadmap when navigating the reasons why these scientific areas were explored.

This thesis and the research contained herein, fits broadly within the context of 3 distinct scientific disciplines. These are:

1. Neurobiology;
2. Cognitive Psychology; and
3. Engineering.

2.1 Neurobiology

Neurobiology is the branch of biology specifically dealing with the nervous system, and applies, in part, because how we hear sounds implicitly limits the value of intra-aural stimulation as a medium for conveying information. In this section, I will provide a brief overview of the following topics:

- How do we hear?
  - What frequencies are we sensitive to?
  - What decibels (dB) are we limited to?
- What is masking?
- What is Neuroplasticity and how does musical training alter our brains?

[1] Due to page size restrictions, this figure may not show up clearly in print form. It is recommended that the reader access the digital image hosted at https://dl.dropboxusercontent.com/u/68065118/Literature%20Review.pdf

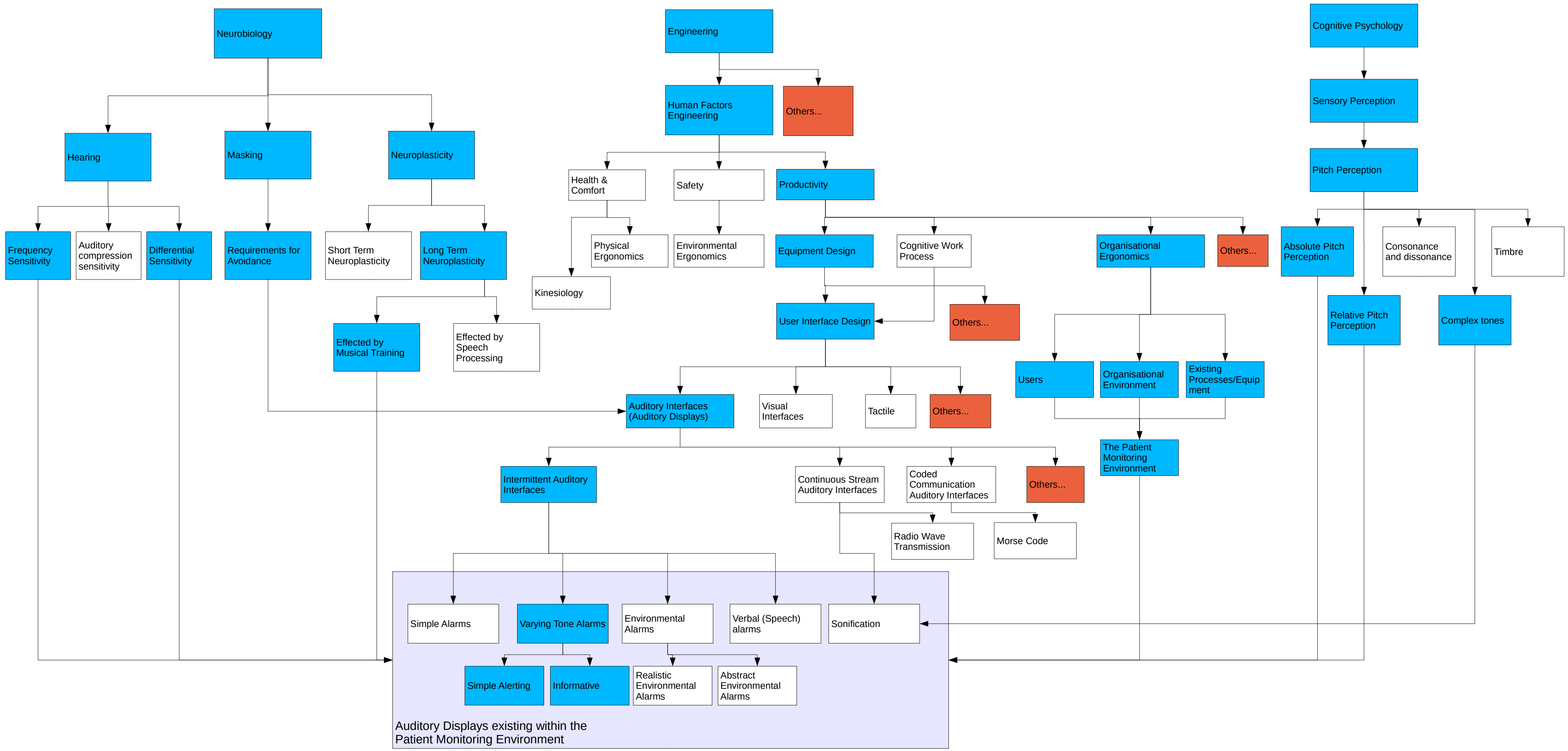


Figure 2-1: Literature Review - Relationship between the high level scientific disciplines and how it relates to the research domain (bounded by purple box). Blue coloured boxes represent aspects which directly contribute to the research question. Boxes in white have been included to provide additional context but do not directly contribute to the research question. Also available at full resolution at https://dl.dropboxusercontent.com/u/68065118/Literature%20Review.pdf

2.1.1 How do we hear?

Primates (humans included), and most other mammals, hear by virtue of three key body parts: the tympanic membrane, the ossicles, and an organ called the cochlea (which is innervated by the auditory nerve). In short, it is the role of the tympanic membrane (eardrum) to receive the auditory signal and transmit the wave through the ossicles (the three small bones in your ear) to the cochlea. The wave then travels through the basilar membrane (BM), and into to an area called the Organ of Corti. Here, stereocilia (small hair cells) vibrate in response to the movement of the BM and stimulate the auditory nerve.

The physical limitations of the auditory pathway fall into 3 key categories:

Frequency sensitivity – although there are 15,000 stereocilia in the human ear, not all of these stimulate the release of neurotransmitters. In fact, only about half of these hair cells are innervated; these are known as the inner hair cells (IHC). IHCs in humans are ideally sensitive to frequencies of .02 to 20 kHz, and age is the most pronounced variable affecting frequency sensitivity with both range and speed of processing being affected.

Presbyacusis is the progressive deterioration of auditory sensitivity associated with aging and is the most common cause of adult auditory deficiency (Seidman, Ahmad, & Bai, 2002). The audible range of frequencies declines from .02 -> 10+ kHz at 20 dB (age 20), to .02->1.5 kHz at 20 dB (age 60) (Chou, Dana, Bougatsos, Fleming, & Beil, 2011).

There is also an age-related difference in the rate of frequency change that can be processed by the auditory cortex (Mendelson & Ricketts, 2001). The processing of rapid, complex acoustic cues, such as rapid changes in frequency and amplitude, are considered to be the basis of speech discrimination loss in the elderly.

Auditory compression sensitivity – this refers to the ear’s ability to endure stress and perceive soft noises. Outer hair cells (OHC) respond to wave pressure by stimulating the basilar membrane. As the basilar membrane vibrates, the OHCs will feed this back through the membrane, amplifying low level vibration by as much as 50 dB (Bacon, 2006).

OHCs are less responsive under high compression. As demonstrated in the figure below, hearing loss due to damaged OHCs not only means that lower dB stimulus are inaccessible, but also that the perceived rate of change in the stimulus level is impaired (Bacon, 2006).

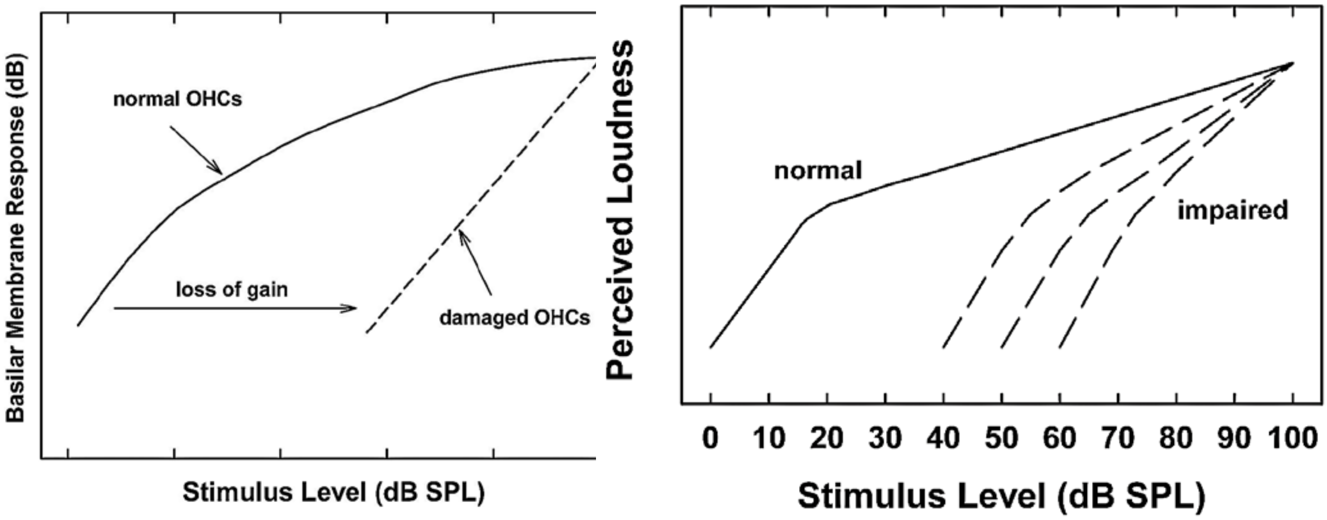


Figure 2-2: From Bacon (2006) (Left) The effects of OHC damage on the basilar membrane. (Right) Perceived loudness grows as a function of stimulus level in individuals with normal hearing. The dashed lines show loudness growth for individuals with different amounts of hearing loss

Differential sensitivity – often referred to as the difference limen, this is the smallest perceivable difference between either two different frequencies, or the same frequency at different intensities (dB).

- Intensity differentiation –(Viemeister & Bacon, 1988) found that the difference limen of intensity varied between 10 and 50 dB depending on frequency, with higher frequencies requiring higher variations for perceptibility

- Frequency differentiation – the ability to discern changes in frequency is dependent on both the frequency of the primary tone (f_0) and its intensity (dB), with the smallest variability (~1 Hz) detected at low frequencies over 40 dB (Jesteadt, Wier, & Green, 1977). The difference limen was greater than 100 Hz at frequencies over 4 kHz.

2.1.2 What is masking?

In short, masking is when the perception of one sound is affected by the presence of another sound (Gelfand, 2004). Auditory Masking is measured by whether the threshold of audibility (i.e. 8 kHz at 40dB) is altered by the presence of another pure tone, and if so, by how much. As a general rule, low frequencies mask higher frequencies, and high frequencies have little effect on low frequencies (see Figure 2-3).

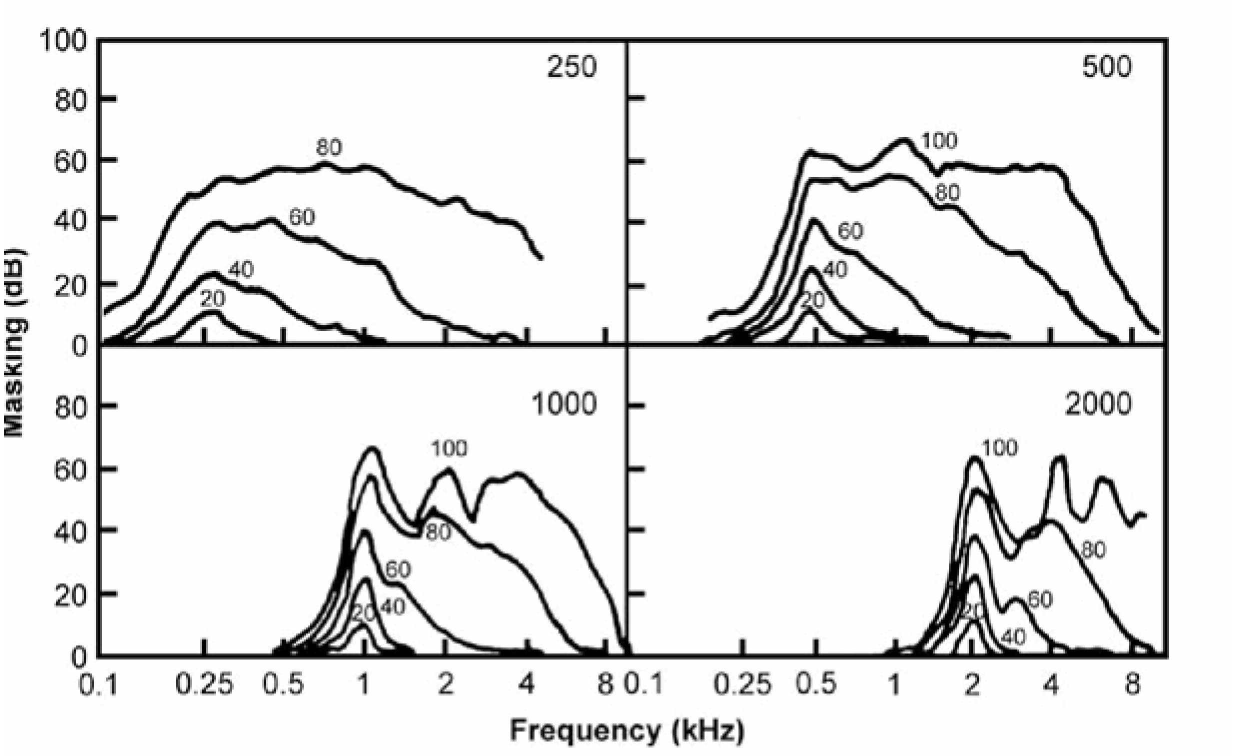


Figure 2-3: Masking patterns produce various pure tone markers (Masker frequency indicated in each frame) from Gelfand (2004) citing Ehmer (1959)

Masking frequencies (in Hz) are shown in the top right corners of each chart, with the masking effect (in dB) indicated by the line.

2.1.3 What is Neuroplasticity and how does musical training alter our brains?

Neuroplasticity refers to alterations of the brain's structure related to experience, degradation (due to age or illness), neurogenesis, or injury. These alterations can impact memory, learning, signal transmission time, and response time, and fall into two categories:

- **Short term neuroplasticity** – which has been shown to occur during sound localization tasks (Macleod & Carr, 2005), where binaural cues such as inter-aural level differences, or inter-aural timing differences are present. Here, the type of information (timing vs. intensity) being transmitted by different subsets of auditory nerve were altered in the presence of stimulus, demonstrating the ability of one nerve to facilitate the efforts of another. Most of us would recognize this effect in straining to hear something happening across the room, and while it might improve our perception of specific things, it is both limited in duration, and takes receptors from other tasks to accomplish this.

- **Long term neuroplasticity –** speech processing is possibly the most common way in which auditory stimulus alters the structure of the brain but it is not the only way. Musical training also significantly alters the brain's structure (Hyde et al., 2009). Children subjected to 15 month of musical training showed significant brain deformations (as compared to the control) in motor control areas (Right Precentral Gyrus, and Corpus Collosum), auditory region (Heschl's Gyrus), and outside the auditory/motor areas in the bilateral frontolateral and frontomesial regions and a left posterior pericingulate region. Fundamental auditory abilities such as the preattentive encoding of spectral and temporal features have also been observed to be enhanced in musically trained individuals when compared to those without musical training (Van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004).

2.1.4 Why this is relevant to the research

Physical sensitivities define the boundaries of acceptable alarm design, particularly when taken in environmental context. Given that the design in question focuses almost exclusively on pitch variation, the assumptions we make about the user demographic (age, medical history, proclivity for attending rock concerts, etc), and the environment (other people, other machines, tasks, etc.) become critical. We will touch on these points throughout the research, and go into some detail in recommendations for later research (Chapter 10).

2.2 Cognitive Psychology

Cognitive Psychology is the branch of study focused on the behavioural implications of specific neural processes. Here I will present how the biological structures, and changes to these structures (e.g. aging and neuroplasticity), alter the auditory task performance of individuals. Specifically, I will address the following topic of pitch and the perception of tones through the traits of absolute and relative pitch.

2.2.1 Pitch

Typically used interchangeably with the term frequency, pitch (in the psychoacoustical sense), and pitch perception, is the ordering of frequencies by audible intervals called 'pitch scales'. These scales are the direct result of human's differential sensitivity to frequency variation (see Section 2.1.1). It should come as no surprise then, that pitch perception does not follow frequency modulation (Hz). Instead, studies have shown that as the frequency increases above 1 kHz the rate of perceived pitch change (from 1 kHz to 2.1 kHz), required an actual frequency adjustment of more than 6 times that amount (1.3 kHz to 8 kHz) (Zwicker & Fastl, 1999).

The ability to correctly identify an individual pitch (e.g. D# in a musical scale) is not a common trait, and individuals displaying this skill fall into one of two (or both) categories: absolute pitch, and relative pitch.

Absolute pitch (AP) is the ability to name specific musical notes without the aid of a reference note. This ability has not been shown to manifest through training in adults, but is instead believed to occur as a result of environmental influences in children (Ohnishi et al., 2001). Physiologically, musically trained individuals possessing of AP have been shown to have larger planum temporale regions in general and, when compared to other non-AP musicians, the right planum temporale showed significant deviations in absolute size.

Performance wise, Absolute Pitch has been shown to increase both the speed and accuracy of pitch identification in both absolute and relative frequency identification tasks, as well as enhanced auditory sequence recollection (Ross, Herdman, & Pantev, 2005) but performed no better than other musicians in identification of timbre variations, or when both pitch and timbre vary.

Relative Pitch is traditionally defined as the ability to accurately identify pitch patterns in melodic sequences (known as contours), but this may not be totally correct. In fact, Relative Pitch has also been shown to relate to other tonal variations other than pitch, and may instead be described as the general ability to discern alterations of contour, including volume and timbre (McDermott, Lehr, & Oxenham, 2008). One critically important distinction between Relative Pitch and Absolute Pitch is that Relative Pitch can be learned at any age (provided the individual has healthy ears).

Other psychological features of auditory perceptions which need to be addressed are consonance and dissonance, complex tones, and timbre.

Consonance and dissonance are the perceived ‘pleasantness’ or ‘unpleasantness’ of two or more frequencies played together.

Complex tones are pure tones with an added harmonic, creating the perception of a new tone that may be perceived to be ‘lighter’, or more ‘mellow’ than the pure tone alone.

Timbre is the quality of the note produced and is what distinguishes one instrument from another, be it within a group (i.e. a Stradivarius vs a common violin) or between groups (brass vs string instruments). Even though each instrument plays the same notes, they will all have distinct qualities.

2.3 Engineering

The science of engineering contributes largely to the application portion of this thesis. Based on concepts from the previous 2 broad scientific disciplines, Engineering attempts to conceive and model an appropriate solution to a given problem or objective.

2.3.1 Human Factors Engineering

The primary objective of human factors engineering is to breakdown the interactions between humans and the systems with which they work (Wickens, Gordon, & Liu, 1998) with the focus on achieving:

- Health and Comfort;
- Safety; and
- Productivity.

Human Factors (proper), therefore, is the study of the necessary factors, development of tools and measurement criteria to facilitate the achievement of these goals.

As systems can be cognitive, physical, mechanical, operational and virtual in nature, the study of human factors is by necessity a multidisciplinary field, incorporating psychology, engineering, biomechanics, mechanobiology, industrial design, statistics, graphics design operations research and anthropometry (Wickens et al., 1998).

Historically speaking, human factors grew from a concern for human interaction with physical devices (typically military or industrial). The US Department of Defence's Handbook for Human Engineering Design Guidelines (MIL-HDBK-759C, 1995) lists in its general guidelines that:

> *The design of systems, equipment and facilities should conform to the capability and limitations of the fully-equipped individual to operate, maintain, supply and transport the material in its operation environment consistent with tactical criteria and logistic capabilities. Accordingly design-induced workload, accuracy, time constraints mental processing and communications requirements should not exceed operator, maintainer or controller capabilities.*

2.3.2 Productivity and Design

The subject of this thesis comes under the domain of human factors, primarily because it concerns the ability of individuals to perform a safety-critical function (monitoring patient health) correctly and effectively. This requires that the task itself be within the capability and limitations of a typical "fully-equipped" operator, using the system for the purpose in which it was intended in the design (MIL-HDBK-759C, 1995).

As much of Human Factors design engineering is concerned with ensuring equipment fits with the limitations, both physical and cognitive, of the human animal, it is for this reason that Neurobiology (focusing on physical aspects of listening and hearing) and Cognitive Psychology (Interpretation and perception of what we hear) were introduced in sections 2.1and □in the chapter. By understanding these physical and cognitive limitations, useful recommendations be made to the design and improvement of Auditory Displays in the Patient Monitoring Environment.

2.3.3 User Interface Design

One of the key elements to managing productivity is the effective design of the Human Machine interface. Displays (or User Interfaces) to machinery or equipment are typically human-made artifacts designed to support the organic perception of relevant system variables, and facilitate the further processing of that information (Wickens et al., 1998). These can exist in a variety of forms (e.g. visual, auditory, tactile, etc.).

There are three steps in effective display design:

1. Determining the nature of the task that said display is intended to support (alerting, informing, reminding, supporting, etc).
2. Determining what resources (physical, cognitive, information, etc) are necessary to carry out the task.
3. Determining the best physical representation of the display (sensory mode/scale/colour/size/etc).

These steps form the basis of some of the fundamental principles of Display Design which are separated into four distinct categories:

- Perceptual Operations – Usage of the senses to perceive display input;
- Mental Model – Understanding of what the information represents;
- Human attention – Direction of focus to where it is needed; and
- Human memory – Retaining information about what the display means and how it can be used;

2.3.3.1 Perceptual Operations

Perceptual Operators deal with the ability of an individual to perceive the information being conveyed by the display. These include:

Absolute Judgements: Non-contexualised, single sensory valuables such as colour, pitch, size, loudness or line lengths. On this, research has suggested that it is advisable to avoid requiring a user to judge between too many levels of a represented variable on the basis of single sensory variables, as humans are generally bad at making these judgements on variables which contain more than 5 to 7 possible levels. (Rouder, Morey, Cowen, & Pfaltz, 2004; Wickens et al., 1998).

Top-down processing: (e.g. Red means stop; arrows pointing to the right mean forward, etc) Individuals perceive and interpret signals in accordance with past experience. This is why signals that are contrary to expectations will meet with a greater probability of errors or confusions.

Redundancy Gain: Messages that confirm or reinforce other messages will more likely be interpreted correctly. Examples of these could include an alarm which reflects an abnormal value displayed numerically on a visual display.

Discriminability: If 2 signals are similar they will likely be confused when they are perceived. Levels of similarity in visual displays can be easily calculated as the ratio of similar features but in Auditory displays this required more understanding of the similarity of features such as Pitch, Rhythm, Timbre and Tone.

2.3.3.2 Mental Model Principles

The mental model is a cognitive representation of how a system actually works and how it's used (Norman, 1988). Mental models are useful in that when users share a similar mental model of a display representation (through standardisation for example), it is a powerful tool for learning and recall. However, mental models also differ in terms of whether they are personal or similar across large groups of people (population stereotype). Even designs developed to cater for a population stereotype should be mindful of the implications if the same interface is used across different countries cultures and language zones.

One method of ensuring a shared cognitive mental model is by using a model that everyone already shares; that of how we interact with the real world (e.g. Gravity, light/dark, Up/Down, etc). 'Ecological Interface design' is the adherence to the principle of realism and correspondence with the environment (Rasmussen & Vicente, 1989). Because the display mirrors real world situations with which all individuals have experience, they are particularly useful as displays because they invoke a strong sense of recall.

2.3.3.3 Human Attention Principles

The greater the number of interfaces, the greater the number of items there are trying to claim the attention of the operator. User interface designers have, by necessity had to accept that the attention of the user is going to be divided and to direct attention only when necessary. By presenting multiple compatible resources (such as Visual and Auditory displays) in tandem, information processing can be facilitated.

2.3.3.4 Memory Principles

Working memory is vulnerable because of its limited capacity. The capacity of working memory is generally considered to be around seven elements (Miller, 1956) with adjustments for memory grouping or age of the individual. Therefore, it is not recommended that user displays rely on this limited capacity especially in high pressure safety critical environments. Various techniques may be employed to aid memory and recall of display aspects such as predictive displays, usage of previous knowledge of the world (Norman, 1988) and consistency of approach.

2.4 Auditory Interfaces

The aspects of User Interface Design (described above) are critical to the design of effective auditory displays. In the following sections, I will discuss the different types of Auditory Displays commonly used, along with their advantages and disadvantages.

2.5 Auditory Displays and their purpose

In recent times, our experiences of auditory information displays have grown significantly, from simply hearing unpleasant warning sounds when an alarm sounds or an error occurs on a computer system, to an almost continuous stream of beeps and bops from the numerous pieces of technology that surround us at almost every point in the day. As technology advances, sound generating technology improves with it, producing more useful and accessible auditory displays. The obvious down side to this advancement is that humans simply aren’t able to handle the amount of information being conveyed. Even if you could eliminate conflicting or duplicate streams (e.g. moving when you realise that the person sitting next to you has selected the same ring/text tone as you), the volume of information coming through the auditory input channel is continually making it difficult to distinguish between useful information and unimportant noise (Brewster, 1997).

2.5.1 Types of Auditory Displays

On a very basic level, an auditory display is the use of sound to communicate information (Mcgookin & Brewster, 2004) from the monitoring source to the user. In fact, the term 'Auditory Display' can be used to refer to almost any kind of process or device that uses sound to provide information. There are many types of auditory displays and each has its advantages, disadvantages and appropriate uses. In order to accurately evaluate any auditory system it is necessary to understand the capabilities and limitations of each type. This would ensure that the right type of display was used for the right type of job.

The following sections present a summary of the four types of auditory displays (Generic Auditory Alarms, Earcons, Intelligent Alarms and Sonification) currently being used or considered for use in the patient monitoring context along with their strengths and weaknesses.

2.5.1.1 The Generic Auditory Alarm

The most common form of auditory display is an auditory alarm. These bursts of occasionally annoying and surprising sounds are commonly considered as having the role of alerting and providing warnings of potentially dangerous or aberrant states.

Auditory alarms are frequently seen as 'a significant attractor of attention' (Stanton, 1994), are commonly related to warnings of danger, and take advantage of the primitive human principle of fight or flight. As such, auditory alarms have been used throughout human history largely unchanged in the form of cries for help, ringing of the town hall's bells to alert an oncoming siege, or tornado sirens to inform people to seek shelter. These alarms are simple and easy to understand.

Increasing technological advances, however, and information onslaught from multiple sensory inputs, has resulted the generic alarm losing much of its urgency. The modern ear, upon receiving a generic auditory input now has to decipher if the alarm is conveying any of a number of pieces of information from a possible fire alarm (or fire drill), a car alarm (or false alarm), a mobile phone or an alert that the battery on a device is low. It is because of the overuse of auditory alarms to convey mundane or false events that people have become numb to the alerting properties of generic loud sounds. Therefore, more recent developments into computer or integrated control design have given rise to a newer more informative type of alarm, which aids in distinguishing one alert from another in a system where multiple parameters or functions are being monitored.

2.5.1.1.1 The Artificial Tonal Alarm

The artificial tonal alarm is the typical alarm that has been used since Alarms were put into use in equipment through the use of synthesised electronic sound. Research into the use of artificial tonal sounds to convey information about abnormal events is largely based in the more areas of military and industrial applications. Patterson (1982) presented a protocol for civil aircraft based on pitch differences with urgency mappings built into the melodies. This protocol was then adapted for the medical domain (Patterson, Edworthy, Shailer, Lower, & Wheeler, 1986) where six rhythmically distinct sounds were proposed which related to the six organ systems. The problem with the artificial tonal alarm is that apart from the natural tendency of all alarms for directing attention, meaning and labels are difficult to associate without extensive training and practice (like learning a new language). This makes artificial tonal alarms difficult to remember and learn (Leung, Smith, Parker, & Martin, 1997), as well as limiting the number of them which can be used in a particular environment.

2.5.1.1.2 The Verbal Alarm

Speech-based (or Verbal) auditory alarms are an effective way of communicating abnormal events because there is little or no added instruction necessary. Although in some cases they can be effective if sparingly used (Hakkinen & Willignes, 1984; Leung et al., 1997), verbal alarms are more easily masked by ambient noises and conversation, can take a long time to convey necessary information (Block, 1994; Leung et al., 1997; Wilcox, 2011) as well as eliciting a reactive rather than a considered response by the medical practitioner (i.e. following the instructions as an order as opposed to considering what issues the alarm could be indicating) (Block, 1994; Stanton & Barber, 1997; Wilcox, 2011).

Verbal alarms are also affected by limitations such as language and cultural barriers where words and tones may convey different meanings to different people. There are also necessary considerations that some alarm announcements when generally broadcast may cause unnecessary fear and panic amongst those who may misunderstand their meanings (i.e. hospital patients, members of the general public, etc).

2.5.1.1.3 The Environmental Alarm

Environmental Alarms as described by Bellas and Howard (1987) as a sound type which differs from an abstract sound in that it is associated with real or every day events. Leung et al. (1997) refer to these sounds as auditory icons which should not be confused with the more abstract tonal patterns described by Blattner, Sumikawa, and Greenberg (1989) as also referred to as Auditory Icons or “Earcons” (described later in Section 2.5.1.2).

Environmental alarms far more easily learnt because they build upon an individual’s existing lexicon of sound associations. This gives the alarms context to which meaningful labels can be easily assigned. Participants in the Leung et al. (1997) study indicated that environmental alarm associations were more easily learnt and retained than those from purely abstract sounds and associations as well as being learnt as readily as those involving verbal alarms.

However, it is necessary to note that an individual’s ability to easily make sound associations based on environmental cues is going to be largely based on their current exposure to those sounds. For example, the participants of the Leung et al. (1997) were students and staff of the Aeronautical and Maritime Research Laboratory and therefore would be familiar with the sound of a Fog Horn and Sonar beep (two of the sounds used in the Auditory Icon Set). Block (1992) similarly presented a list of alarms based on the melodies of popular songs such as “I left my heart in San Francisco” to convey associated physiological system labels such as the Cardiovascular System. These methods can be easily shown to be effective information transmitters and subjects readily learn the associations, however, they are not always transferable across international and cultural divisions.

2.5.1.2 The Informative Auditory Alarm (Earcon)

The informative type of alarms, or ‘Earcons’, are short informative tones structured to created auditory messages. These were developed by Blattner and colleagues (Blattner et al., 1989; Sumikawa, 1985; Sumikawa, Blattner, Joy, & Greenberg, 1986) as an auditory counterpart of a computer's graphical icon which, at the time, represented a significant advancement of being able to convey a great deal of visual information concisely and succinctly.

The basic structure of an earcon is the 'motive', which is a short rhythmic sequence of pitch variations which are combined in different ways. Two important key features of Earcon design were highlighted in page 5 of Sumikawa et al. (1986):

> *"A motive is a brief succession of pitches arranged in such a way as to produce a tonal pattern sufficiently distinct to allow it to function as an individual recognisable entity"*
>
> *"The eloquence of motives lies in their ability to be combined to create larger recognisable structures. The repetition of motives, either exact or varied, or the linking of several different motives produces larger, more self sufficient patterns."*

Like all auditory displays, the key feature variations of Earcon motives are: Rhythm, pitch, timbre, register and dynamics. These variations are manipulated to form different and identifiable elements in order to convey a message.

Earcons are not usually naturally occurring sounds unlike the Leung et al. (1997) environmental sounds, but instead, are tones which a user, with frequent contextual usage, can learn and associate meaning with. Mappings of events to sounds are usually arbitrary and symbolic (Gaver, 1997), which requires that the user gain familiarity or will have to be trained in recognising the association.

This is one of the major disadvantages of earcons as an information transmission medium. Studies have found that even in low stress conditions humans do not readily learn and retain more than eight sounds (Patterson & Milroy, 1980).

2.5.1.3 The Intelligent Alarm

All the alarms described so far in Sections 2.5.1.1 and 2.5.1.2 all share one thing in common: these alarms are, by nature, indiscriminate. By this, we mean that the alarm will trigger as soon as a pre-defined value is reached. The machine does not consider whether what it perceives as abnormal could, or should, be expected. This contributes greatly to false and nuisance alarms (Woods, 1995).

Intelligent alarms (Blom, 1988; Loeb, Brunner, Westenskow, Feldman, & Pace, 1989; Mylrea, Orr, & Westenskow, 1993; Navabi, Mylrea, & Watt, 1989; Orr & Westenskow, 1994; Rheineck-Leyssius & Kalkman, 1999a, 1999b; Tsien, 1997; Westenskow & Orr, 1992) on the other hand, use knowledge-based networks to decide if an abnormal parameter is an indication of an event that requires the operator's attention.

The implementation of intelligent alarms requires the implementation of sophisticated decision theoretic frameworks (Huang & Shachter, 1997) and knowledge databases which, not only have to comply with strict industry software safety regulations, but also be reliable enough to be accepted by the general public. Intelligent alarms are currently more prevalent in high-volume and high-automation industries where variables are finite and human interaction is limited such as, Distributed Temperature Sensing, Security, Data Mining and automated factories.

2.5.1.4 Sonification

An alternative auditory display to the traditional alarm is the continuous auditory display or Sonification. Sonification is the representation of multidimensional data parameters in the form of a continuous sound stream (Fitch & Kramer, 1994; Kramer, 1994; Loeb & Fitch, 2000; Sanderson, Crawford, Savill, Watson, & Russell, 2004; Seagull, Wickens, & Loeb, 2001; Watson, Russell, & Sanderson, 1999; Watson & Sanderson, 2001, 2004) or Data Auralisation (Gaver, 1997). This method uses the musical dimensions of sound such as pitch, timbre, harmonics, loudness and rhythm, as well as sound element components such as attack, sustain and decay to map multidimensional data, and presents them as a single auditory stream. Changes in data are reflected in changes in the sound's parameters.

Sonified information can be classified as either realistic or abstract. Realistic sonification, such as the ones used in (Fitch & Kramer, 1994), imitate the actual sounds which they are conveying the information about (i.e. breathing sounds for ventilation). Abstract sonifications are sounds which are not obviously perceived as being associated with real-world sounds (e.g. manipulation of a realistic breathing sound's pitch to represent the level of carbon dioxide in the patient.)

As alerting displays, sonifications and alarms have the benefits of being ubiquitous, which allows for the operator's visual attention to be elsewhere while the operator is still being informed of any important event (Sanderson, Watson, & Russell, 2005). The added benefit of sonification over alarms is that sonification provides continuous information which gives the operator a chance to observe gradually changing status rather than only being alerted when a problem becomes serious, and allow pre-emptive action to be taken before critical events. The continuous nature of the sonification display is that the 'normal' condition is continuously presented so that even a relatively small, but significant deviations, can be readily detected by a trained operator prior to the onset of an alarm event.

Sonification is already currently used in medical equipment. The pulse oximeter presents the information (heart rate and arterial blood oxygenation) continuously, thus allowing the operator to have their focal attention on completing another task while still having timely alerts to when the patient's cardiovascular conditions deviates from what is expected. The sonification of other patient parameters such as respiration (respiration rate, end tidal carbon dioxide and tidal volume) has been extensively investigated in laboratory settings (Sanderson et al., 2004; Sanderson, Watson, et al., 2005; Seagull et al., 2001; Watson, Sanderson, & Russell, 2004) and is showing promise in improving patient care; however, more ecologically relevant research has yet to be completed

2.5.1.5 Advantages of auditory displays

The ubiquitous and attention-directing property of auditory signal displays makes them appear to be the most appropriate method of conveying information in a work environment where the visual focus of individuals is not on visual display monitoring equipment. Areas in medical care, such as the operating room and ICU, are such examples because the attention of the medical practitioner is mainly focused on the patient, or may be diverted towards any number of essential duties and not constantly focused on the visual monitors of each patient. It is not surprising, therefore, that nearly all patient monitoring equipment used in the medical field has been fitted with some form of auditory signal display to direct the attention of medical practitioners towards unacceptable and abnormal values of patient monitored parameters (Watson et al., 2004).

2.5.1.6 Disadvantages of auditory displays

Studies into the disadvantages of auditory displays have mainly concentrated on the problems of alarms in fault detection and critical care systems. (Woods, 1995)

Advances in technology have resulted in the proliferation of safety critical equipment in almost every manufacturing, transportation, industrial, and medical domain, This equipment is produced by a variety of manufacturers who, as (Edworthy & Meredith, 1997) mention, would prefer to avoid the potential for any litigation associated with not providing all their equipment with enough auditory alarms. Momtahan, Hetu, and Tansley (1993) citing Wiklund and Hoffman (1998) point out that medical devices, for example, are usually purchased separately over a period of time with little consideration for compatibility with other products. A typical piece of medical equipment might produce six or more different alarm sounds whereas an anaesthetic work station with several add-on modules might conceivably produce 20 or more different alarm sounds (Edworthy & Hellier, 2005).

The sheer quantity of alarm sounds in the critical care environment has resulted in researchers such as (Loeb, Jones, Leonard, & Behrman, 1992; Momtahan et al., 1993) addressing the problems of people not being able to differentiate reliably between alarms in their work environment, even though they would be expected to be familiar with the alarms. To provide an example; (Loeb et al., 1992) found that 21 anaesthetic clinicians could correctly identify only 34% of the 19 recorded alarms from anaesthesia devices. They also suggest that the ability to distinguish between alarms in the operating room is somewhat more reliable than this number because experienced clinicians use other information to foresee potential events and make pre-emptive preparations. Momtahan et al. (1993) had similar findings from 26 alarms recorded from the ICU and operating room of a Canadian hospital. Their participants were able to identify only about 10 to 15 of the 26 operating room alarms and 9 to 14 of the 23 ICU alarms. (Loeb et al., 1992) also note that anaesthetists mentally group alarms by function and that it may be possible to reduce the maximum number of distinct alarm sounds in an operating room to lessen alarm confusion as well as decrease response time to potential problems.

In an observational study conducted by (Seagull & Sanderson, 2001) it was found that anaesthetists' responses to alarms could be classified into four categories, *correction, expected, ignore, and reminder*. The most common response to the alarms was to ignore them (48% of audible alarms). Of the remaining alarms, only 33% elicited a corrective response and 6% were expected by participants. A later observational study by (Watson et al., 1999; Watson et al., 2004) reported that anaesthetists found alarms in clinical environments to be confusing, intrusive, uninformative, and sometimes to be more of a nuisance than help. (Guillaume et al., 2005) conducted an analysis of auditory warning signals in the operating room and found that only 9% of the signals led to an action on the patient, whereas 20% were considered useful alarms in the control of anaesthesia procedures. This indicates that the remaining 71% of the signals recorded were considered not useful by the operating room staff.

The high quantity of alarms has long been criticised as contributing to increased noise, workload and annoyance in mission critical systems (Woods, 1995). Despite this, alarms still play a vital role in directing attention to important events and cannot be dismissed. The ubiquitous and obligatory properties of auditory alarms make them invaluable in safety critical domains such as anaesthesia, air traffic control, power stations, fire warnings and military operations. In such domains, there can be many alarms with different meanings which prompt different actions. Identification and response to abnormal conditions is considered a priority as errors could result in loss of life or revenue or both.

Various studies across multiple industries, such as into alarms in anesthesia environments (Sanderson, Anderson, & Watson, 2000; Sanderson et al., 2004; Seagull & Sanderson, 2001; Seagull, Xiao, Mackenzie, & Wickens, 2000), in the Three Mile Island incident (Kemeny, 1979) and in flight deck automation (Sarter, 2000), have indicated that the uninformative nature and overuse of alarms has not helped in compromising situations, sometimes going so far as to make matters worse. Moreover, medical alarms in the ICU and operating room have been criticized as too numerous and often inappropriate, distracting and confusing (Meredith & Edworthy, 1995).

2.6 Scope and Context of the Patient Monitoring Environment

The Patient Monitoring environment is a term that can be used to describe any aspect of health services which involves active tracking of patient status. This term could apply to areas of the operating room, centres of acute care and even patient home monitoring. In all of these areas, patient monitoring is a highly demanding field in which technological developments have significantly increased in the last century. This has had the tandem effect of both improving the practitioner’s efficiency as well as adding significant elements of complexity to their work environment.

2.6.1 Users of the Patient Monitoring Equipment

The primary users of the patient monitoring equipment are the health practitioners namely doctors and nurses. Patients and their support group may be considered secondary users of the system, especially as improved displays, improved self-support equipment and information availability mean that these individuals may be able to provide additional support and feedback when the primary users are unavailable.

Primary users who are intimately involved with some of the most sophisticated patient monitoring equipment are specialised doctors known as anaesthetists. These highly trained professionals are usually responsible safeguarding the patient (i.e. sustaining life and preventing injuries and providing necessary surgical conditions (i.e. immobility and consciousness) (Xiao, 1994) in dynamic and complex environments with varying or high workload and risk levels (Gaba & Howard, 1995). It is for this reason that a vast amount of patient monitoring research has traditionally been focused on the anaesthesia environment (Boquet, Bushman, & Davenport, 1980; Loeb et al., 1989; Mylrea et al., 1993; Navabi et al., 1989; Seagull & Sanderson, 2001; Thompson, Watson, & Sanderson, 2005; Watson et al., 1999; Williams & Beatty, 2005) with more focuses on nursing staff (Potter et al., 2005) and home care being reviewed in recent times particular due to the developed world’s increasingly aging population.

2.6.2 Patient Monitoring Environment

The patient monitoring setting involves a complex series of physical and cognitive activities over prolonged timescales and while under rapidly changing conditions. Both doctors and nurses are required to respond and anticipate changing clinical conditions. This is supported by a number of pieces of equipment which may include:

1. Drug delivery systems;
2. Breathing systems that mechanically ventilate the patient, or allow the practitioner to manually ventilate the patient;

3. Integrated or autonomous automated monitoring systems that detect equipment failures and monitors the patient for abnormal vales on vital parameters that may lead to organ damage or death; and

4. Delivery systems for oxygen and other gases.

Due to the diverse nature of the equipment, large and varied number of manufacturers and a fairly general IEC standard for equipment display design, medical practitioners are often having to familiarise themselves with a number of pieces of equipment, thus increasing the workload that they are under.

2.6.3 Auditory Displays in the Patient Monitoring environment

Most, if not all patient monitoring equipment is fitted with alarms to direct attention to unacceptable values, A major drive behind this is that anaesthetic machine developers are avoiding the potential for litigation in the high risk realm of patient safety (Watson et al., 2004). As a result, manufacturers will have a tendency to bias their systems in favour of “false positives” rather than “false negatives. Until recently, developers seldom considered the human factors consequences of alarm systems in the highly variable domain of patient monitoring. The combination of independently developed systems with the wide variety and frequency of alarms from numerous equipment, makes it difficult for a medical practitioner to be certain of what an alarm is indicating, thus failing the primary intent of the equipment designers who use alarms in their designs, which is to direct attention to the intended area.

2.6.3.1 Types of Auditory Displays used in Medical Equipment

In the operating theatre, it is a basic requirement of medical staff to continuously monitor the status of a patient while that patient is undergoing an operation. In the distant past this simply consisted of checking a patient’s pulse and breathing. With the sophistication of modern medicine, however, a patient may be attached to several monitoring devices, each with its own visual and auditory displays and alarms. A typical visual display will continuously indicate the patient’s status by means of one or more visual information signals, and is capable of generating additional visual alarm signals should the need arise.

While a typical auditory ‘display’ may include an auditory information signal (continuously indicating the value of a particular parameter), its more fundamental purpose is to alert the operator to the existence of abnormal situations – and their severity – by means of auditory alarm signals (Watson & Sanderson, 2004). This is mainly due to the fact that medical staff, such as anaesthetists and nurses, do not consistently have their visual attention focused on the monitoring equipment; rather, their primary visual focus is on the patient, preparing charts, or drawing up medication. Weinger and Englund (1990) note that Boquet et al. (1980) found, after analysing 16 hours of filmed anaesthetic activities, that 60% of the visual activity was directed at either the patient or the surgical field, with only 5% directed at the monitors.

Auditory Displays are especially important in safety critical systems such as patient monitoring, where the identification and response to abnormal conditions is a priority, and focused visual attention is not easily sustained. Auditory displays in the patient monitoring environment have been designed, not so much to support patient monitoring, but more to prevent the operator from missing crucial events at times when his/her attention is diverted from the patient monitoring task. Recent alarm standards require that auditory alarms indicate not only the priority of the alarm but also its causal category; they are no longer only designed to simply attract a human observer’s attention (Block, Rouse, Hakala, & Thompson, 2000; IEC60601-1-8, 2003-08-14).

Different types of auditory displays described in section 2.5.1 can be found in medical applications. The most common and prolific would be the generic alarms as they require very little instruction and are simply for the purposes of directing attention to abnormal or concerning events which would require additional visual verification or application of knowledge (i.e. a patient’s call buzzer). A major shortcoming of these alarms is that they do not provide much information and when occurring in high pressure situations in tandem with periods of high alarm proliferation could add stress to medical personnel. However, various other forms of auditory displays can also be found in the medical environment and each has a different level of complexity, usefulness, problems, contribution to workload and ability to annoy.

Intelligent alarms have also been proposed as being a viable solution to solving some of the inherent problems of proliferation, false, nuisance and ambiguous alarms in the patient monitoring and operating room domain (Blom, 1988; Loeb et al., 1989; Mylrea et al., 1993; Navabi et al., 1989). However, due to the high number of human interactions and wide variability of possible events and environment, intelligent alarms that use intensive knowledge-based reasoning have not been yet successfully applied in Operating Rooms and the patient monitoring context, causing the interest in intelligent alarms to lessen significantly and research into the area to become increasingly scarce. The reliance on intelligent alarms for assistance in higher order decision making is also something that medical professionals as well as members of the general public, find difficult to accept.

Over the last decade, several researchers have also examined the usefulness of continuous auditory displays (or sonification) in helping medical practitioners monitor patient physiological parameters in the operating room (Fitch & Kramer, 1994; Loeb & Fitch, 2000; Sanderson et al., 2004; Seagull et al., 2001; Watson et al., 1999; Watson & Sanderson, 2001, 2004). It is important to consider the effects of sonification in the design of auditory alarms because of their common use in medical applications (e.g. pulse oximetry and fetal monitors).

The benefit of pulse oximetry in helping in patient monitoring has resulted in researchers looking into using sonification to represent patient parameters other than the heart rate and oxygenation. (Fitch & Kramer, 1994) sonified eight physiological parameters and investigated the ability of college students to monitor and detect changes. They found that the sonification out-performed the visual and the combined visual and sonification display. However, the sounds that were used bore similarities to realistic sounds. For example, the respiration sound resembled natural breathing, which has been criticized as having the potential to be mixed up with real patient sounds (Watson and Sanderson 2004). Natural sounds are also more likely to be masked by ambient noise.

In order to identify if monitoring patients with sonified parameters would help in a practical application where a user would typically be occupied with multiple tasks, Seagull et al. (2001); Seagull et al. (2000) conducted studies where patient parameters were sonified and monitored by non-medically trained participants under timeshared-task conditions to examine how well the sonification did at alerting participants to abnormal conditions when they were occupied with something else. Accuracy performance in the timeshared task in the combined condition was worse than in the visual or sonification condition. This finding should be taken into account when evaluating the usefulness of any proposed display, particularly within the context of the essential attributes of the intended use environment (Wilcox, 2011).

Any new auditory display should be designed with the fact that this form of continuous sound will, increasingly, be in the background.

2.6.3.2 Alarms vs. Sonification

Due to the success of the pulse oximetry in aiding patient monitoring, Sonfication has been considered as a possible solution to the problems experienced with traditional alarms by making traditional alert type alarms unnecessary (Fitch & Kramer, 1994; Sanderson et al., 2000). A previously conducted study, by the author, was concerned with investigating whether sonification in patient monitoring could make alarms redundant by alerting anaesthetists to potential problems before alarms sounded (Wee, 2003). This was done by comparing three conditions in a within-subjects design.

- Visual display with alarms and sonification;
- Visual display with alarms; and
- Visual display with sonification.

Several limitations were noted at the completion of the study (Wee, 2003) and were further expounded by Philp (2004). One of the main criticisms was that the frequency and number of probes was insufficient to explore the full benefits of early abnormality detection with the sonification. In addition, under the heavy workload of having to report all vital signs at regular intervals, the participants tended to be hypervigilant in order to be able to correctly answer all the questions. The participants tended not to trust the auditory displays and treated each condition as if it were the visual condition (Wee, 2003).

Although the alarm limits used were representative of those used currently in the operating room, the alarms only sounded once during an abnormal event and no follow-on alarms were sounded, thus reducing the occurrence of nuisance and repetitive alarms. In addition, the publication of IEC 60601-1-8 has made it clear that sonification should be compared with melodic alarms as well as general alarms. In conclusion, the Wee (2003) study should be considered to be premature in attempting to test alarms vs. sonification without reflecting on the current state of alarm design.

It is one of the goals of this thesis to examine the current state of the standard in patient monitoring auditory alarms in order to provide a more informed answer to the relative benefits of melodic alarms in the patient monitoring environment.

2.7 Evaluation Techniques

The previous sections determined the theoretical and application context of the design of auditory displays in the domain of patient monitoring and the scientific and research knowledge that should be taken into consideration when determining an effective alarm design. In addition, the means of determining if the IEC 60601-1-8 alarms are effective needs to be broken down into measureable aspects. This is where the domain of Usability and Utility studies can be applied.

Usability describes both the learnability and ease of use of man-made objects such as software, tools, machines, and technology, with a goal towards increasing efficiency, productivity and satisfaction of users of the system. The study of Usability is primarily focused on psychological matters of the user experience and is highly relevant to the evaluation of information displays and monitoring devices in a human factors context.

The meaning of usability and methods of effectively measuring usability are varied and largely dependent on the industry or the type of device being assessed. Large amounts of usability research comes from assessing of graphical user interfaces and websites which places a large amount of focus on aesthetics and visual characteristics to guide users towards achieving their goals.

Shackel (1991) describes Usability and Likability as 2 of the 3 major dimensions of human factors, with Utility (does it fulfil the intended purpose) as the third. Other authors such as Sauro (2010) and Nielsen (2001), group Usability and user satisfaction together under measures of usability along with success rate, time taken and number of errors made.

The measurement techniques used in the experimental phases of this thesis have been modelled after these descriptions and the following types of data was collected and analysed:

- Accuracy;
- Timely responses;
- Analysis of errors made; and
- Overall user satisfaction.

3 DEVELOPMENT OF MEDICAL ALARM STANDARDISATION

Many researchers have called out for the need for standardisation of alarms in medical electrical equipment (Loeb et al., 1992; Meredith & Edworthy, 1995; Momtahan et al., 1993; Schmidt & Baysinger, 1986; Weinger & Englund, 1990). Thus far, the engineering approach has directed the design and standardisation of auditory displays in most domains, patient safety included. (Weinger & Englund, 1990) and (Block, 1992) have specifically called for a human factors approach to design of alarms in the operating room. Healthcare organisations such as the Joint Commission on Accreditation of Healthcare Organisations (JCAHO) have made clinical alarm safety one of its patient safety goals in recognition of the fact that patients continue to be injured or killed because of ineffective alarms (Edworthy & Hellier, 2005). Standardized alarms have also been considered by international standards groups such as the International Organization for Standardization (ISO), European Committee for Standardization (CEN) and the American Society for Testing and Materials (ASTM) standards (Block, 1992).

In earlier efforts to standardize alarm sounds in the operating room, (Patterson et al., 1986) proposed six rhythmically distinct sounds which related to the six organ systems. The alarm sounds were based on Patterson's work in civil aviation alarms. These alarm sounds represented six kinds of devices or monitors in which abnormal conditions requiring alarm alerts might lead to permanent injury or death.

However, the original Patterson et al. (1986) alarms were actually quite acoustically rich and more complex than the ones suggested by the IEC 60601-1-8 standard nearly a decade later. Weinger (1991) reports that the American and International Standards organisations failed to reach a consensus on alarm tone standards for medical monitoring equipment (when authoring ISO-9703 (1992)), adding that "Monitoring equipment manufacturers stated that that may be required to design their equipment to comply with the European standard in order to sell their products in that market" suggests more of a reluctance of manufacturers to try and replace their existing soundcards to be able to play these complex melodies and an attempt to push the standards committees to approve a simpler design. This is somewhat understandable given that it occurred in the mid 90s where technology was not moving as comparatively fast as it is today.

In addition, Weinger (1991), in his letter to the Journal of Anesthesiology highlighted that a holistic human factors examination of the operating room environment needed to take place, and that imposing alarm guidelines on individual device manufacturers would likely exacerbate an increasing concern of noise and stress in the Operating Room. He specifically points out that the Patterson alarms were too strictly defined and their approach could hinder innovation in alerting technology. This would technically also apply to any standard guideline proposed and not just the Patterson alarms; at least until operating room equipment from various manufacturers is able to be integrated with logic prioritisation of alarms and central displays. The Patterson alarms were not formally tested for learnability and usability prior to their proposed inclusion.

Building upon the understanding that people learn and retain information faster when it is already familiar to them, Block (1992) proposed a set of six alarm sounds based on popular themes and songs. The labels for these alarms were readily identified by a sample set of North American anaesthetists but would have been culturally specific and unlikely to be as easily adopted by practitioners internationally. Manufacturers were also reluctant to use these types of alarms because of uncertainty of what ‘tunes” could be reasonably used, as well as the fact that the melodies used in Block’s study were copyrighted and did not fall under blanket permissions (such as for radio use). A few years later, (Block et al., 2000) composed proprietary sounds. Although these sounds were incorporated into the Datex-Ohmeda AS/3 monitors, no formal testing or operator acceptance data was collected prior to 2003 when (Mondor & Finley, 2003) conducted a study in order to examine the perceived urgency of 13 auditory warning alarms commonly occurring in the operating room. The Datex-Ohmeda AS/3 monitor was one of the machine alarms used. The other machines were the Hewlett-Packard (HP) model 66S monitor and the North American Dräger Vitalert 3200. A number of studies have found that the perceived urgency of many alarms in the patient monitoring environment was not consistent with the actual urgency of the situation which triggers the alarm (Momtahan et al., 1993; Mondor & Finley, 2003). Participants also did not indicate that they perceived any differences in urgency between low, medium and high priority alarms in the Datex-Ohmeda and the HP machines.

An ISO group worked on an update to the ISO 9703 alarms prior to 2000, but in 2000 it was superseded by an IEC-convened Joint Working Group (JWG) focusing on alarm signals. The work of this group eventually led to IEC 60601-1-8, *General requirements, tests and guidance for alarm systems in medical electrical equipment and medical electrical systems*. In IEC 60601-1-8 the alarm sounds were encoded by organ system as proposed by (Patterson et al., 1986) and were designed with a different melody for each alarm but conformed to the same rhythmic pattern designed by Block and Rouse for Datex-Ohmeda in the mid 1990s. The general alarm sound was already standard, but a further sound for “power failure” was added, bringing the total to eight categories of alarms (Table 3-1). These categories were further separated into high and medium priority alarms bringing the grand total to 16 alarms in the proposed set to satisfy the ISO, CEN and ASTM standards.

Alarm	Melody*† and mnemonic		Other information in support of mnemonic
	Medium priority	High priority	
General (GE)	C4-C4-C4	C4-C4-C4—C4-C4 (repeated)	Fixed pitch, traditional (usual) 9703 sound
Oxygen (OX)	C5-B4-A4 'OX-Y-GEN'	C5-B4-A4—G4-F4 (repeated) 'OX-Y-GEN A-LARM'	Slowly falling pitches; top of a major scale; falling pitch of an oximeter
Ventilation (VN)	C4-A4-F4 'VEN-TI-LATE' 'RISE-AND-FALL'	C4-A4-F4—A4-F4 (repeated) 'VEN-TI-LATE A-LARM' 'RISE-AND-FALL AND-FALL'	Old 'NBC chime'; inverted major chord; rise and fall of the lungs
Cardio-vascular (CV)	C4-E4-G4 'CAR-DI-AC'	C4-E4-G4—G4-C5 (repeated) 'CAR-DI-AC A-LARM'	Trumpet call; call to arms; major chord
Temperature (TE)	C4-D4-E4 'TEM-P'RA-TURE'	C4-D4-E4—F4-G4 (repeated) 'TEM-P'RA-TURE A-LARM'	Slowly rising pitches; bottom of a major scale; related to slow increase in energy or (usually) temperature
Infusion (IN)	C5-D4-G4 'IN-FU-SION'	C5-D4-G4—C5-D4 (repeated) 'IN-FU-SION A-LARM'	Jazz chord (inverted 9th); drops of an infusion falling and 'splashing' back up
Perfusion (PE)	C4-F#4-C4 'PER-FU-SION'	C4-F#4-C4—C4-F#4 (repeated) 'PER-FU-SION A-LARM'	Artificial sound; tri-tone; similar to 'yo-ee-oh' of the Munchkins in 'The Wizard of Oz'
Power failure (PF)	C5-C4-C4 'POW-ER FAIL' 'GO-ING DOWN'	C5-C4-C4—C5-C4 (repeated) 'POW-ER GO-ING DOWN'	Falling pitch as when the power has run down on an old Victrola

Table 3-1– IEC 60601-1-8 Alarms with Melody, Mnemonic Lyric and the Rationale for the Mapping of the Melody to the Alarm

3.1 The problem with the latest standards

Experiments by Patterson and Milroy (1980) have indicated that participants can readily learn to recognize four to six audible alarms and that a week after learning 10 alarms, naïve participants could reliably recognise only six. Although the 16 alarm sounds in Block et al. (2000) appear to be easily broken down into eight different organ system (each organ system with a medium and high priority alarm) the sounds might be too difficult to learn and retain.

The potential for confusion is another area of examination for this research program. Patterson (1990) indicates that warnings with similar rhythms and repetition rates were likely to be confused even when there were gross spectral differences between the warning sounds. The warning sounds in the most recent IEC standard all share a similar rhythm developed and published in the ISO 9703 series of standards. Although the melody of each warning sound was designed to be different by standards of pitch variation alone, it might lower potential confusions and lessen the learning time needed if the warning sounds were rhythmically distinct. Studies into absolute and relative pitch discrimination have indicated that without formal music training to increase discrimination sensitivity, individuals are generally not very good at making accurate judgement of pitch intervals (Hyde et al., 2009; Zwicker & Fastl, 1999). This may pose a significant problem if different alarms have similar melodic shape but subtle differences in pitch intervals, such as the alarms presented in the IEC standard do.

The IEC melodic alarms distinguish urgency by increasing the length of the alarm, and repeating it at a faster speed. Block et al. (2000) mention a concern that conveyance of urgency by this acoustical feature may not convey a proper sense of urgency (a concern that Mondor and Finley (2003) put to the test), but they (and others) contend that levels of urgency can be learned. Mondor and Finley (2003) have counterarguments to such an untested assertion. Moreover, Wright Mackenzie, Buchan, Cairns and Price (1991, in (Walsh & Beatty, 2002) show that 80% of preventable mishaps were attributed to human error contributed by mainly inexperienced staff. We cannot, therefore, assume that staff will all have the necessary training and time to get used to working with alarms that are not intuitive.

The alarms proposed by (Block et al., 2000) and released by the IEC as part of the international standard for alarm systems in medical electrical equipment and medical electrical systems (IEC 60601-1-8) were not formally tested for learnability, memorability and discriminability in a way that informed the standard in either its draft release in 2003 or its full release in 2005. Independent tests only emerged in 2005 as part of this thesis (Sanderson, Wee, & Lacherez, 2006; Sanderson, Wee, Seah, & Lacherez, 2006; Wee & Sanderson, 2005, 2006) and in other independent research (Williams & Beatty, 2005) .

4 INTRODUCTION TO THE RESEARCH PROGRAM

4.1 Motivation

The first aim of the experimental program was to ensure that the set of stimuli used was indicative of that which would be used in medical institutions. In 2005, an international standard, IEC 60601-1-8 was published which describes guidelines to the design of visual and auditory alarm systems in medical electrical equipment which is recognised as a collateral standard in many countries. The IEC 60601-1-8 is a comprehensive international standard that specifies basic safety and essential performance requirements and tests for alarm systems in medical equipment. It’s 71 pages including four annexes covers both visual and audible alarms, but a vast majority of the specification is devoted to audible alarm issues for medical equipment and applications. The standard covers a broad and comprehensive range of topics from what kind of medical condition should trigger an audible warning sound to the specific frequency and shape of the audible sound waveform. For the purposes of developing the stimuli to be used in this program of research the following areas were given particular focus:

4.1.1 Priority condition

IEC 60601-1-8 gives guidance on whether a condition should be assigned a high, medium or low priority. This guidance is based on the potential result of a failure to respond to the cause of the alarm condition and how fast the harm could happen to the patient. For example, a high priority alert would be assigned if death or irreversible injury could happen quickly; however, a low priority alert would be assigned if minor injury or discomfort may happen after a period non-interference. It has been left to the equipment designer to decide which specific conditions merit which priority. A heart rate (HR) of 170 on a treadmill test may result in a low condition priority alert whereas the same HR at an intensive care patient monitoring station may be assigned a high priority alert.

4.1.2 Audible Alarm Pulse

The audible alarms guidelines in IEC 60601-1-8 requires that three separate pulse trains (or bursts patterns) of sounds are used. These three pulse trains correspond to the three priority conditions discussed in 4.1.1. A high priority condition burst has 10 fast pulses that repeat, a medium priority condition has 3 slightly slower pulses, and a low priority condition has 1 or 2 even slower pulses that may repeat (optionally). The shape of the pulse trains are the spacing restrictions of each pulse for the priority conditions are presented in the standard in great detail.

The characteristics of the individual pulses are also described in the standard. Each individual sound pulse must have a fundamental frequency (pitch) somewhere between 150 to 1000 HZ and there must be at least four harmonics from 300-4000 Hz. To the human ear, the resulting blended sound would be similar to hearing a single harmonically complex sound. A final requirement for the sound pulse characteristic is that the sound level (measured in dB) of the four harmonic tones must be within +15dB of the fundamental frequency tone.

The spaces and timing of the pulses used in this program will be described in further detail in Section 4.2.

4.1.3 Melodic Physiological/Technical Variations

The IEC 60601-1-8 specification indicates that any additional alarm melodies included in piece of equipment being manufactured should be identifiable and not be confusable with the specified melodies in Annex EEE of the standard (the 8 alarm categories e.g. General, Cardiovascular etc). This means that the equipment would be able to play another musical melody to convey unacceptable states other than those specifically stated in the standard as long as it precluded the possibility of confusion with the specified Annex EEE melodies..

The melodies in the IEC 60601-1-8 were designed under the challenging constraint of conforming to fixed rhythmic patterns already established for medium and high priority alarms in ISO 9703-2. Unfortunately, there was no formal test of the effectiveness of the melodic alarm with representative users before they were included in the IEC standard. The sequence of musical notes for each melody is spelled out in detail in IEC 60601-1-8 for several different kinds of medical applications. For example, the cardiac alarm is assigned a unique melody that is musically different from the one assigned to the infusion alarm. By assigning a unique melody for each kind of medical physiological system, the hope is that medical personnel can become familiar with each different kind of melody and more quickly identify the source of an audible warning sound.

Six physiological systems were identified to have their own unique alarm melodies. These are Oxygen, Ventilation, Cardiovascular, Temperature, Infusion and Perfusion. Two more sounds - one for General and another for Power Failure – were added bringing the total to 8 categories. These categories were further separated into high and medium priority alarms bringing the grand total to 16 alarms in the proposed set. Figure 4-1 presents the musical melodies of the 16 alarms in medium and high priority along with the way the alarm name could be sung to the tune. The standard describes the high priority alarm as repeated once even though Figure 4-1 displays it only once.

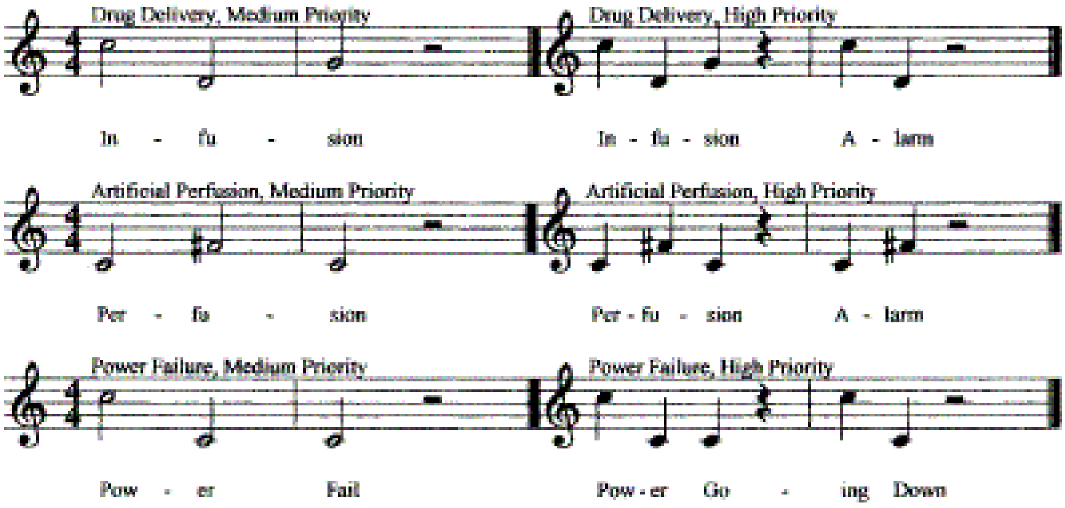


Figure 4-1: Musical melody of each alarm by physiological system with General and PowerFail. The melody for the high priority alarm is played twice.

4.1.4 Sound level Requirements

The IEC standard contains only basic requirements for how low the audible warning sounds need to be. The only requirement listed is that the high priority alarm condition warning sound must be louder than the medium priority condition warning sound which must be louder than the low priority condition warning sound. IEC 60601-1-8 does not specify how loud the sounds need to be, choosing instead to rely on the experience of medical equipment designers, and equipment users to determine ultimately how loud the warning sounds should be based on usage, environment and individual practitioner. This lack of guidance has the potential to cause issues as equipment manufacturers, prudently choosing to avoid the potential of litigation, may make a warning sound loud enough to be heard in all situations which may cause those who are near to the warning to be annoyed or distracted.

A practical recommendation is that there should be a 3 to 6 dB difference between each of the three priority condition warning sounds. It takes at least a 3 dB difference between two different sound levels before a healthy human ear can tell one sound is louder or softer than another sound, therefore anything less than 3 dB would not make sense. However, it is important to ensure that the high priority alert not be too much louder than the low priority condition or the lower priority conditions may be masked when high priority alarms are being sounded. This may also cause distractions or annoyances to patients and other medical staff in the nearby vicinity.

4.2 Structure of Research Goals

As mentioned in Section 1.2 this thesis has been developed using Goal Structured Notation (GSN) as a formatting means to communicate how a particular line of inquiry has been developed and addressed by means of experiment or analysis. The complete experimental GSN is presented in Appendix A.

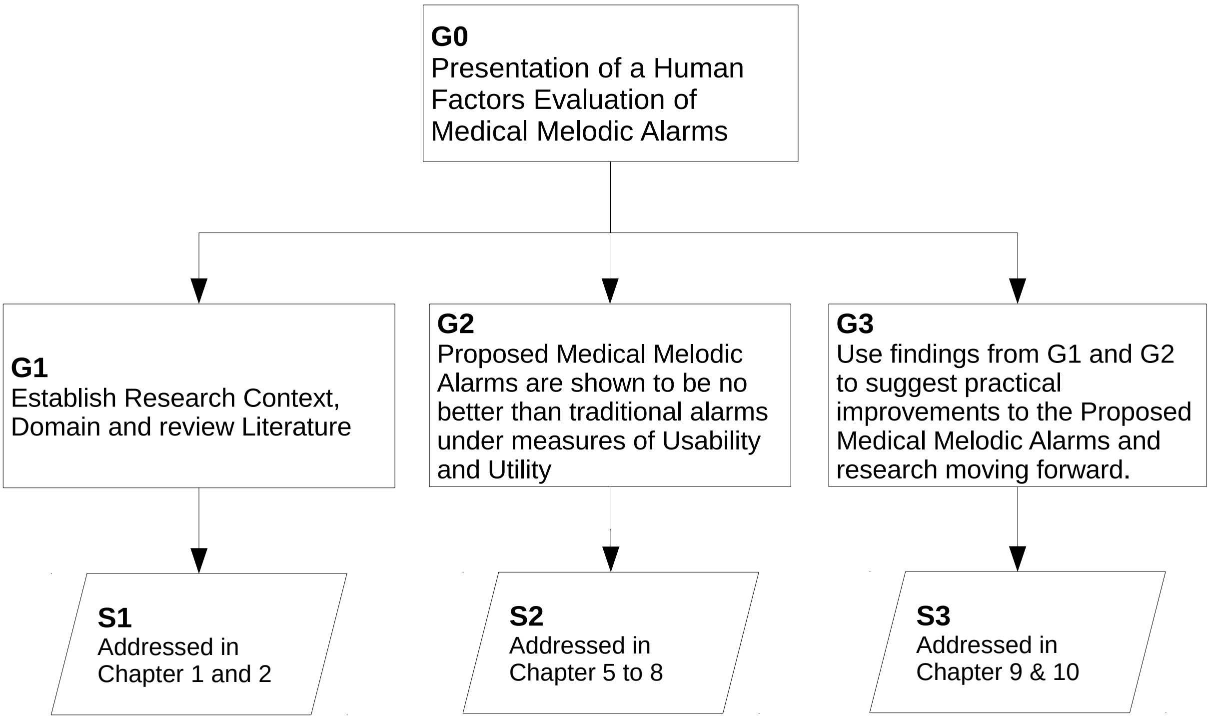


Figure 4-2: Top Level Goals

This high level research goal is addressed by three subgoals. These are:

- Goal 1- The establishment of the Research Context, Domain and Literature
- Goal 2: Evaluation of the Medical Melodic Alarms under measures of Usability and Utility
- Goal 3: Using theoretical references and experimental data to suggest practical improvements to the proposed medical melodic alarms.

In order to quantify the usefulness of melodic alarms I interpret that the measure of the usability of the alarms can be equated to how easily, quickly and well a user can be trained to understand the alarm presentations and their associated meanings. These are presented in Goal 2 as assessment of the alarms by measures of:

- Goal 2.1.1 Learnability – Accuracy of responses to presented stimuli;
- Goal 2.1.2 Memorability – Ability to recall the alarm meanings in short and long term;
- Goal 2.1.3 Speed – Ability to respond to the alarms in a timely manner;
- Goal 2.1.4 Efficiency – Ability to respond to the alarms while performing another task;
- Goal 2.1.5 Analysis of Errors – Review the errors that are being made; and
- Goal 2.1.6 User Feedback – Review of user subjective feedback.

4.3 Methodology

Four experiments were conducted in this thesis. In all experiments the IEC 60601-1-8 melodic alarms were used to observe how different learning conditions, durations and distractions affect a participant's ability to learn the alarms.

Experiment 1 addresses the problem of the latest set of IEC standards on auditory alarms not having been formally tested for learnability, memorability, response times, efficiency, error analysis and user feedback. This experiment was conducted to establish if the alarms proposed by the standard were coherent and to identify any design flaws that need to be addressed. The results indicate that with non-medically trained participants there is still a significant amount of confusion between alarms (Wee and Sanderson, 2005). This indicates that improvements still need to be made to the alarm design to improve patient safety.

In Experiment 1, the independent variable (learning with and without mnemonics) was manipulated between participants with half the students being introduced to the stimuli and their associated mnemonics, and the other half being left to utilise their own memory aids in order to accomplish the identification task in 3 learning phases (short term, long term, distracted). Chapter 5 reports on the details and findings of this study.

Upon publishing and presenting the results of Experiment 1 as a journal article (Sanderson, Wee, & Lacherez, 2006) conference presentations (Wee & Sanderson, 2005, 2006), some concerns were voiced by members of the scientific community that the reason that the effect of mnemonics (learning condition) was not evident was due to the fact that the mnemonics were designed for medical practitioners, and that one trained in medical knowledge would more readily be able to relate to the proposed mnemonics. Block et al. (2000) pointed out that the alarm sounds were designed to convey related contour patterns of physiological emergencies (e.g. Rising temperature = bad, therefore temp alarm is increasing pitch, Ventilation indicates breathing which is a up and down motion, therefore the ventilation alarm is a varying pitch alarm, etc). This may have suggested to some that medical professionals may have developed a greater ability to make these associations than people who don't operate in the same domain.

As a result, Experiment 2 extends Experiment 1 by running a similar test with trained healthcare staff (registered nurses). This is to test whether 22 experienced nurses currently employed at 2 major hospitals located in Brisbane, Queensland, showed the same confusions that non-medical participants (students) showed in Experiment 1. Participants in Experiment 2 were also required to answer a quick survey to determine if the selected associations of melodies with the organ systems are the ones that they would naturally make, and to determine the perceived urgency of the alarms. Chapter 6 provides further detail and discussion about the findings of this experiment and the cumulative knowledge gained from a comparison of Experiment 1 and 2.

Experiment 3 was designed to further understand how the "persistent confusions" between groups of alarms affected the participants' ability to master the alarms and achieve better and faster accuracy scores. To achieve this, one of the alarms in a "persistent confusion" pair (Cardiovascular), was removed to observe if the information transmission from presentation to response showed any observable improvement. The details and results of this experiment are presented in Chapter 7.

With Experiment 4, an alarm in a "Persistent Confusion" pair (Temperature) was modified to change it enough to make it more musically dissimilar from the alarm it was repeatedly being confused with (Cardiovascular). The confusion was observed repeatedly in Experiment 1 and 2. This was theorised to be due to the fact that both alarms shared a similar musical pattern and shape and was therefore more prone to being confused with the other. The modified Temperature alarm was designed to keep within the requirements set out within IEC 60601-1-8 and conformed to the same pattern and frequency requirements. Only the melody was manipulated.

4.4 Stimuli

All stimuli used in this experimental program were created using the Csound sound generating software (WinSound version 4.08). To generate sound files, Csound requires an orchestra file and a score file, which contain the parameters of the sound. These files are executed in order to render the sound, which is then stored as a wav file.

For Experiment 1, the eight high priority alarms were a series of five pulses (notes) consisting of three initial pulses separated by a slight pause before a further two pulses. After a longer pause the pattern was repeated, to make a ‘3-2, 3-2’ pattern overall. Pulse rise time (time from zero to maximum amplitude) and pulse fall time were each 10 milliseconds. Pulse sustain time (time at maximum amplitude) was 100 milliseconds and pulse width (pulse sustain time + 10% pulse rise time + 10% fall time) was 102 milliseconds. The silence between each adjacent pulse in the group of three pulses and the group of two pulses was 30 milliseconds. The interval between the beginning of the fall time for one pulse and the end of the rise time for the next pulse was 50 milliseconds. Spacing between the third and fourth pulses was 498 milliseconds. For the eight medium priority alarms, all values were twice the length of those used for the high priority alarms (spacing between pulses of the medium priority alarms was inadvertently 25 ms shorter than recommended by the standard. However, the intended perceptual difference between the medium and high priority alarms was preserved). The melodies in Table 1 were superimposed on these patterns of pulses. The lowest note was middle C (C4: 278.4 Hz) and the highest note was the octave above middle C (C5: 556.8 Hz)

From Experiment 2 onwards, the high priority alarm was created with the IEC 60601-1-8 “3-2, 3-2” pattern of tone pulses (Figure 4-3) with a total duration of 2924 milliseconds. Each medium priority alarm was created with the “3” pattern (Figure 4-4) with a total duration of 1044 milliseconds.

Pulse rise times, duration and spacing were twice as short for the high priority alarms than the medium priority alarms and consistent with ranges specified in the standard.

High Priority Alarm

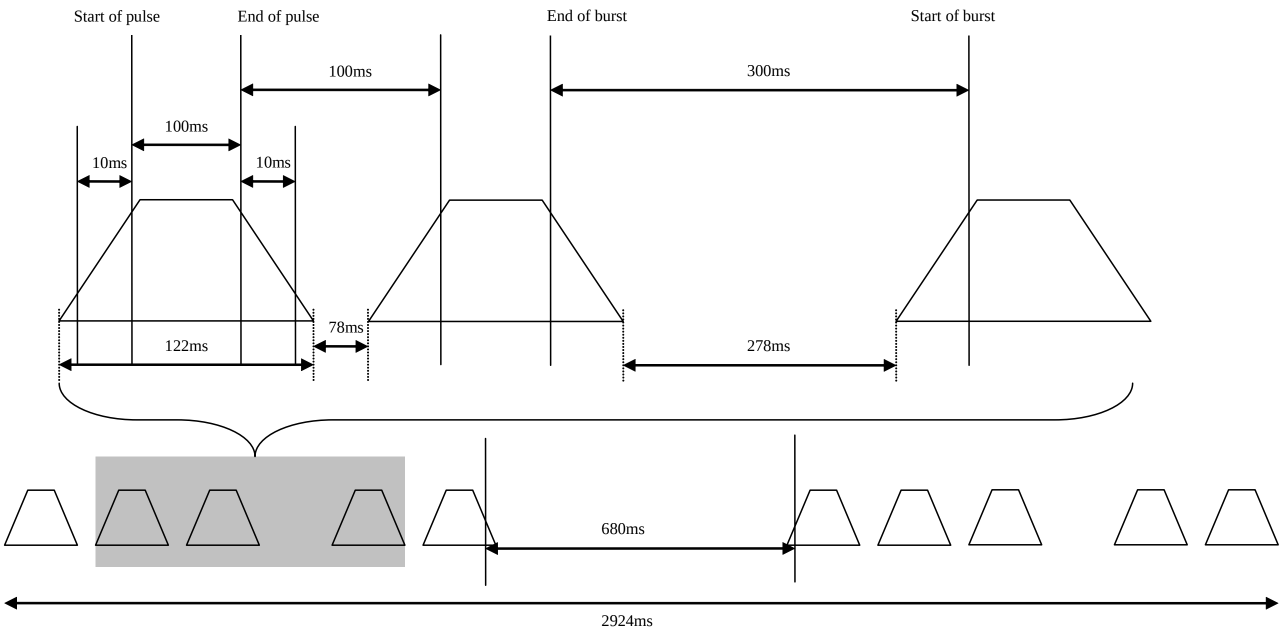


Figure 4-3: Temporal profile for the high priority alarm

Med Priority Alarm

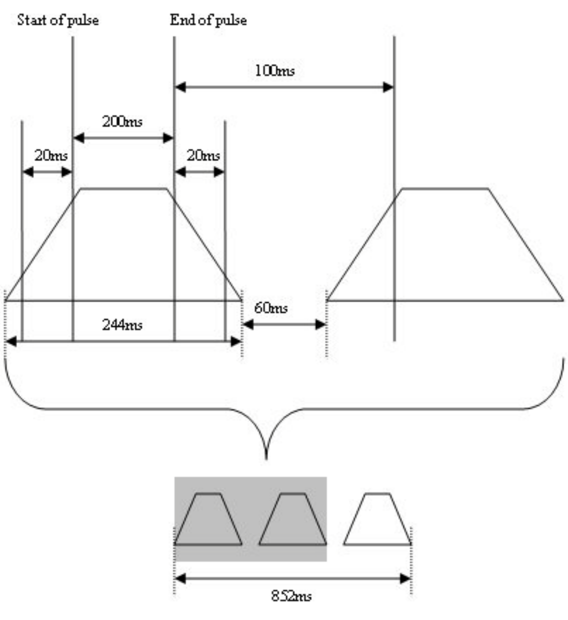


Figure 4-4: Temporal profile for the medium priority alarm

Pulse durations and spacing followed constraints specified in IEC 60601-1-8 as follows:

- Pulse rise time (time for amplitude to rise from 10% to 90% of maximum pulse amplitude) and pulse fall time (time for amplitude to fall from 90% to 10% of maximum pulse amplitude for each tone was 10ms;
- For high priority alarms the pulse duration was 100 milliseconds, measured from the end of one pulse (defined as when its amplitude fell to 90% of maximum) to the start of the next pulse (defined as when its amplitude rose to 90% of maximum);
- For medium priority alarms the pulse duration was 200 milliseconds, measured from the end of one pulse (defined as when its amplitude fell to 90% of maximum) to the start of the next pulse (defined as when its amplitude rose to 90% of maximum);
- For high priority alarms the spacing between the end of burst (end of third pulse) and start of next burst (start of forth pulse) was 300 milliseconds (pulse start and end defined as above);
- For high priority alarms the spacing between the fifth and sixth pulses was 680 milliseconds (pulse start and end defined as above).

The lowest alarm note was the middle C (C4: 278.4 Hz) and the highest was an octave higher (C5: 556.8 Hz). The experimenter checked that the sound level was clear and comfortable for the participant before starting each experiment. No participant reported discomfort and difficulty hearing the alarms and no changes were made to the volume of the alarms which were played at the same volume for each participant.

4.5 Analysis

Three types of analysis were performed on the data in the following experiments:

For multi-variable analysis (such as between experiments and experiments involving several independent variables), ANOVAs (with regression, due to varying sample sizes) were used to measure the significance within and between groups.

For single variable analysis (i.e. after mnemonics were removed from later experiments), t-tests were performed to measure significance in some aspects of the experiments.

For learnability (with the exception of a series of ANOVAs used to measure a potential confound between experiment 1 and 2), a simple linear regression model was used (as below) with total number of correct responses in each block as x (e.g. 11 correct) and the block number as y (e.g. 22).

The standard error of the slope β = $s_\beta^2 = \dfrac{n s_\varepsilon^2}{n S_{xx} - S_x^2}$

Where: $s_\varepsilon^2 = \frac{1}{n(n-2)}\left(n S_{yy} - S_y^2 - \hat{\beta}^2 (n S_{xx} - S_x^2)\right)$

Where: $\hat{\beta} = \dfrac{n S_{xy} - S_x S_y}{n S_{xx} - S_x^2}$

Learnability in all such tests is the slope of the linear regression plot.

Alpha values for all t tests and ANOVAs were set at 0.05.

4.6 A priori prediction of results

From the beginning of Chapter 2 to this point, the following have been presented:

- The broader research context (Chapter 2);
- The focus on the research question (Chapter 2);
- The present situation of the industry (Chapter 3); and
- The introduction to the research program (Chapter 4).

These details provide us with the foundation upon which the following experiments were conducted, the fundamental question of which is:

"Do the IEC 60601-1-8 alarms meet the limitations and requirements of the User in order to achieve the intended purpose?"

Before discussing the experiments a few predictive observations can be made concerning the expected results of the experiments based on the reviewed literature presented in Chapter 2.

4.6.1 Musically trained individuals will perform better than non-musically trained individuals

As described in Chapter 3, the alarms to be tested are melodic alarms which fall into the category of varying tone alarms (see discussion in section 2.5.1.1.1and 2.5.1.2). All alarms in this category are distinguished from each other by modulations in pitch. No variability of timbre, rhythm or tonal length was permitted by the standard and, as discussed in Section 2.2, the literature on psychoacoustics presents a strong indication that musically trained individuals have a greater sensitivity for relative pitch discrimination (RP), and will therefore be much better at distinguishing between subtle pitch differences between alarms with similar pitch characteristics.

4.6.2 The number of different alarms will make recall of all the alarms and their meanings difficult for most individuals

There are 16 alarms presented in total grouped into 2 priorities and 8 physiological systems. Studies into attention and limitations to working memory have suggested that individuals do not readily retain much more than 7 pieces of different information within one context (see discussion in section 2.3.3.4) without being able to cluster the information in a way that is intuitive to the individual. As experiment participants are not going to have significant previous experience with the alarms to be able to draw upon long term memory, short term and working memory are the only means by which the participants in the experiments are going to be able to learn the alarms. It is expected that the participants will not be able to learn all the alarms in any reasonable amount of time in the experimental context without being exposed to them on a day to day basis.

4.6.3 The presence of supplied mnemonics is not going to have any effect on an individual's ability to learn the alarm and their associated labels.

The IEC 60601-1-8 melodic alarms are each presented with 2 supplied mnemonics. One has been designed with the physiological system to which it refers to and the other is based on North American cultural references.

Studies into interface design have recommended the use of memory references to aid in the recall of interface meanings (see discussion in section 2.3.3.2) and recent studies have indicated that population and cultural stereotypes (Norman, 1988) and designs based on environmental aspects (Leung et al., 1997) can greatly aid learnability of alarm meanings.

However, the supplied cultural mnemonics were not based on widespread international stereotypes and the medical mnemonics do not distinguish themselves from each other enough (e.g. Ventilation – Rise and fall of lungs when the Perfusion alarm could also represent the same). Studies have also shown that mnemonics are most useful when the individual naturally comes up with the association rather than having the association provided to them (Edworthy & Hellier, 2005).

4.7 Summary and Conclusions

The auditory information channel in many ways is similar to the visual information channel. There is a richness and variety in sound as well as in colour and shapes with which we can take advantage of from a communication standpoint. Additionally, like the visual information stream, allowances have to be made for limitations in ability, such as impaired vision, colour blindness or tone deafness. For example, traffic lights rely not only on colour to provide information but also position (Red is always on top). This less obvious but useful standardisation helps colour blind individuals differentiate between “Stop” and “Go” signals. With approximately 8% of males and 0.5% of females in the human population are in some way colour blind (see Table 1-5 from Sharpe, Stockman, and Jägle (1999)), most designers take this into consideration when designing visual displays on safety critical equipment.

Similarly, the melodic alarms are heavily reliant on the user’s ability to discriminate between the relative pitches of the different alarm melodies. Relative pitch differential sensitivity has been show in Neurobiology (see section 2.1.1) and studies into neuroplasticity to be a largely affected by training and musical ability (see section 2.1.3), and exposure is not a universal trait that can be depended upon in the general population.

The purpose of this study was to highlight a potential weakness in the development of informational auditory alerts. The focus of which has been specifically directed to the IEC proposed set for informational melodic alarms, yet the findings can be applied towards the development of complex auditory alerts in any domain.

The primary point of focus here is to determine if alerts with varying pitch alone, described as melodic alarms can satisfy the requirement of providing information as well as the provision of directing attention to events of significance to a professional user. The premise of providing additional information onto the all too familiar alarm is a natural evolution of the media. Auditory alarms in the medical domain provide significant advantages in directing focus to abnormal events requiring attention. It would be natural to assume that the next logical step would be to take advantage of this highly useful channel and encode additional information. Medical professionals' attention must be directed to the appropriate information at the appropriate time in order to maintain situational awareness. However, this noble endeavour cannot be impulsively undertaken.

In the following chapters, I have presented a series of experiments which highlight the need for a more holistic and human factors approach to the development of auditory alarms in safety critical systems, focusing on the medical domain as an example of such a system. The results of the studies provide evidence that in the case of the IEC melodic alarms, design assumptions have been made which do not reflect the ability and possible mental fatigue of the end user.

In summary, the goal of informational alarms should be to be able to provide better attention control in a safety critical and cognitively complex domain by directing attention to the correct area. In order to do this, the design of such alarms must take into account the limitations and abilities of the eventual user and result in an alarm which enhances the capabilities of the user over current non-informative alarms; yet, not contribute significantly to the cognitive load already inherent in the working environment.

5 EXPERIMENT 1 - EFFECT OF MNEMONICS (STUDENTS)

Prior to the release of the IEC standard and the commencement of this study (described in Wee and Sanderson (2005)) there was nothing published to indicate that the IEC 60601-1-8 Melodic alarms has been formally tested for learnability, discernability and memorability.

In addition, it was unclear whether the proposed mnemonics would be helpful to people when learning to recognise the alarm sounds and the organ systems which they are meant to indicate. In this experiment, the implications of mnemonics were considered. Specifically, whether participants who were supplied with written mnemonics (M) while learning label-alarm associations were able to more accurately recall what they learnt better than those not supplied with a mnemonic (NM) in short and long term (2 weeks later) testing modes. We also considered the ability of participants to accurately identify alarms which occurred at random intervals while the participants performed a timeshared task.

The Mnemonic (M) and Non-Mnemonic (NM) conditions of the experiment were run in a between-subjects design with approximately half of the participants being randomly assigned to the Mnemonic Condition and half randomly assigned to the Non-Mnemonic condition. Results presented were based on a comparison of participants in the Mnemonic and Non-Mnemonic conditions which have been completed. Some of the findings of this study have been presented in Wee and Sanderson (2005) and Sanderson, Wee, and Lacherez (2006).

5.1 Methodology

Participants were required to learn the 16 different alarm sounds (eight medium priority and eight high priority) proposed by Block et al. (2000) for auditory alarms in medical machines. Familiarisation and introduction were provided, after which participants were randomly assigned to two groups. One group learnt with the help of mnemonics given in Block et al. and the IEC 60601-1-8 standard. The second group were not provided with mnemonics and the sound’s associations on the screen:

- Non Mnemonics (NM): Participants were only provided with the alarm sound and label.

- Mnemonics (M): In addition to the above, participants were provided with the written mnemonic which as well as given an initial demonstration of the sung mnemonic by the experimenter.

Two conditions are being examined with 18 participants in the Mnemonic condition and 15 in the Non-Mnemonic condition participants in each condition. The experiment is broken up into two sessions in order to test for long and short-term memory retention (Figure 5-1).

In the first session, participants were introduced to the alarm sounds and asked to learn them using an active learning technique. Participants chose the order in which they are learning the alarms as well as whether they wanted to learn the sounds (e.g. that is, a participant who has already mastered the general alarm sound may elect not to learn it during the next learning session). Participants are allowed to listen to each alarm only once per learning trial.

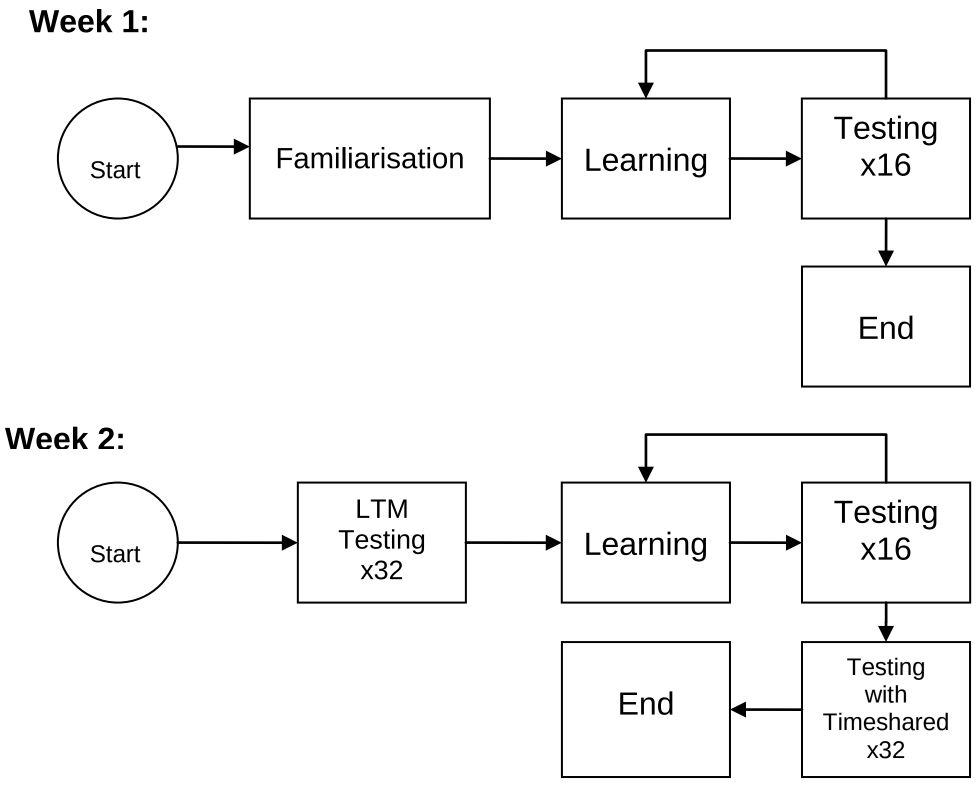


Figure 5-1 - Experiment procedure

Participants were then tested on a randomised set of alarm sounds. A week later the participants were required to return for a second session in which they were retested (without learning) in order to observe the ability of participants to retain the alarm associations over the long term.

Participants are then given a quick revision of the alarm sounds before being asked to recall the sounds which are randomly presented to them while they are occupied with the primary arithmetic task (timeshared task).

5.2 Participants

Participants were 38 first year undergraduate students from The University of Queensland, recruited through the School of Psychology research sign-up system. Although no formal hearing tests were administered, participants were asked prior to the experiment whether they had any known hearing impairments. Participants were asked if they had any formal musical training and for how many years (Musically trained = 24 [63.2%], Non-musically trained = 14 [36.8%]).Of these participants only 33 returned for the second day of testing. (Musically trained = 21 [63.6%], Non-musically trained = 12 [36.4%]).

The study was conducted in accordance with the ethical review processes of the University of Queensland and within the guidelines of the National Health and Medical Research Council.

5.3 Apparatus and Stimuli

5.3.1 Alarm settings

Sixteen alarm sounds were created using Csound based on criteria outlined in the recently published IEC standard for medical equipment (IEC 60601-1-8). These alarms were played on a Pentium® 4 1.9 GHz PC compatible with Soundblaster live 5.1 digital soundcard. The sounds were presented via Harmon/Kardon HK695 speakers at a mean amplitude of 60dB SPL(A), located about one meter from the participant. Alarm labels and mnemonics were displayed on a 19 inch touch screen monitor (Sanderson, Wee, & Lacherez, 2006; Wee & Sanderson, 2005) . A custom-written Java program was used to display the sound labels and record the participants' responses.

5.3.2 Timeshared Task

Wilcox (2011) recommended that when designing alarm signals that testing of the alarms be done in an environment that re-creates the essential attributes of the intended use environment. Seagull et al. (2001) also utilised a similar method to measure the performance of an auditory sonification set with a simplified distracter task.

In this experiment, a simple arithmetic task was used to simulate work that a clinician might be doing while having to monitor the equipment located beside or behind him. The arithmetic task consists of a series of simple addition and subtraction equations displayed on a computer screen, to which the participant is required to answer TRUE or FALSE. A graph on the same screen plots the participants' speed and accuracy over the length of the trial. Participants are told at the beginning of the test that they must try to keep the indicator on the graph as close to the top right hand corner of the graph as possible (i.e. high speed and high accuracy) which would require them to consistently make very quick and correct responses.

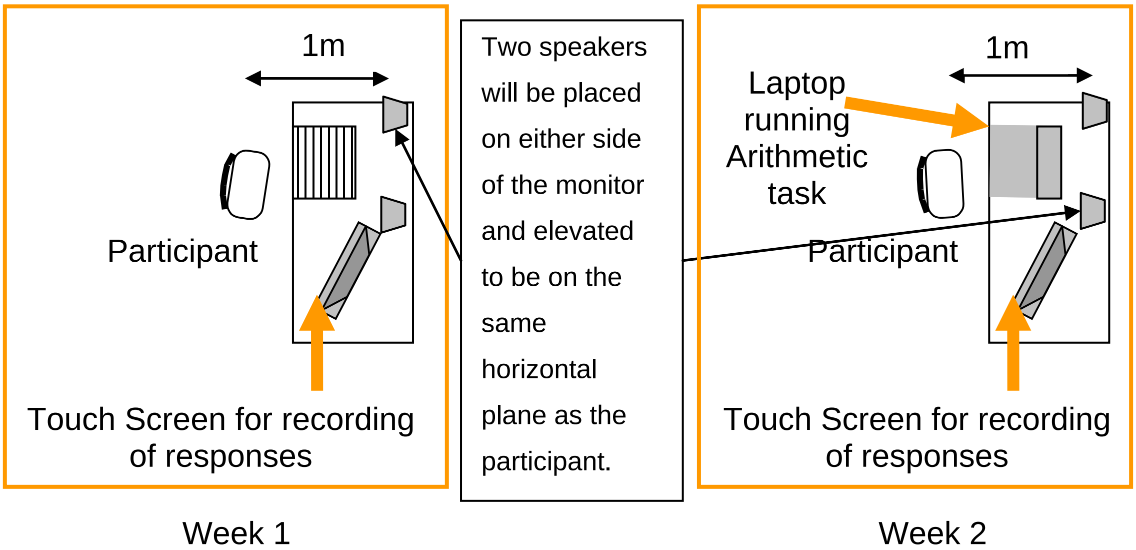


Figure 5-2: Room layout for the arithmetic task and the alarm recognition task

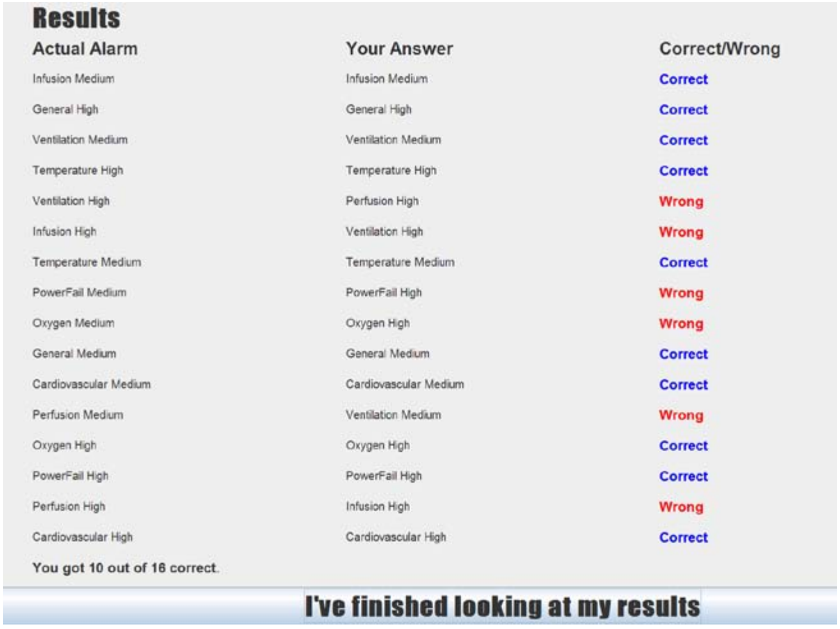
5.4 Procedure

Each participant was seated in a quiet room laid out as in (Figure 5-2). They were given a short introduction to the experiment and after written consent procedures and background questionnaires were completed they were given the necessary instructions. The introduction consists of a familiarisation period where the participants were introduced to the alarm sounds and their associated (physiological system) labels. The participants in the Mnemonics group were introduced to the mnemonics of each alarm sound whereas the Non Mnemonics group are not given this information.

Day 1: Following the familiarisation stage were the learning and testing stages. Participants learnt the 16 different alarm sounds (8 medium priority and 8 high priority) proposed by Block et al. (2000) for auditory alarms in medical machines. The participant proceeded on with a series of learn-test cycles until the participant had satisfied the learning criterion of achieving two consecutive sets of correct tests or until 35 minutes have passed—whichever occurred first.

In the learn phase of the learn-test cycle, the experimenter introduced the sound and meaning of each alarm. A screen was then presented that displays the alarm labels categorized by priority. Participants learnt the alarms by touching the label of the alarm they wish to hear. They could listen to each alarm only **once** and could listen to the alarms in any order. Participants who had already mastered certain alarm sounds could choose not to listen to them again during a subsequent learned phase of the learn-test cycle.

In the test phase of the learn-test cycle, alarms were played in random order. After each alarm, the participant was prompted to select which label, out of a list of 16 provided, was the correct one. At the end of the test (all 16 sounds being played) a results screen would display the alarms that the participant correctly and incorrectly identified (Figure 5-3). If the learning criterion was not reached during the test phase, the participant was returned to the learn phase of the next cycle to review any alarms they were unsure about before attempting another test.



Results

Actual Alarm	Your Answer	Correct/Wrong
Infusion Medium	Infusion Medium	Correct
General High	General High	Correct
Ventilation Medium	Ventilation Medium	Correct
Temperature High	Temperature High	Correct
Ventilation High	Perfusion High	Wrong
Infusion High	Ventilation High	Wrong
Temperature Medium	Temperature Medium	Correct
PowerFail Medium	PowerFail High	Wrong
Oxygen Medium	Oxygen High	Wrong
General Medium	General Medium	Correct
Cardiovascular Medium	Cardiovascular Medium	Correct
Perfusion Medium	Ventilation Medium	Wrong
Oxygen High	Oxygen High	Correct
PowerFail High	PowerFail High	Correct
Perfusion High	Infusion High	Wrong
Cardiovascular High	Cardiovascular High	Correct

You got 10 out of 16 correct.

I've finished looking at my results

Figure 5-3 Results screen which displayed a participants correct and incorrect answers to played alarms.

Day 2(Approx 1 week later): In Week 2 the experiment was broken into three sections.

In the first section, participants were tested on their ability to identify the alarms in two test phases of the 16 alarm sounds (total 32 alarms). The tests were similar to those in Week 1 except that there are no prior learn phases.

In the second section, participants returned to the learn-test cycles as in week 1 until they reach a certain level of expertise. If the participant achieves two sets of complete tests they proceed to stage three. However, if the participant did not achieve the criterion then they were allowed to proceed to the third section only after completing seven learn-test cycles.

In the third section, participants identified alarms while performing the arithmetic task on a laptop. The arithmetic task consisted of a series of simple addition or subtraction equations which the participant answered as “*TRUE*” or “*FALSE*”. At random intervals (between 20 and 70 seconds) a random alarm would be played. Participants had to identify the alarm by touching the appropriate button on the touch screen. 32 alarms (a non-replacement randomised set of the 16 alarms, twice over) were tested in this set.

5.5 Results

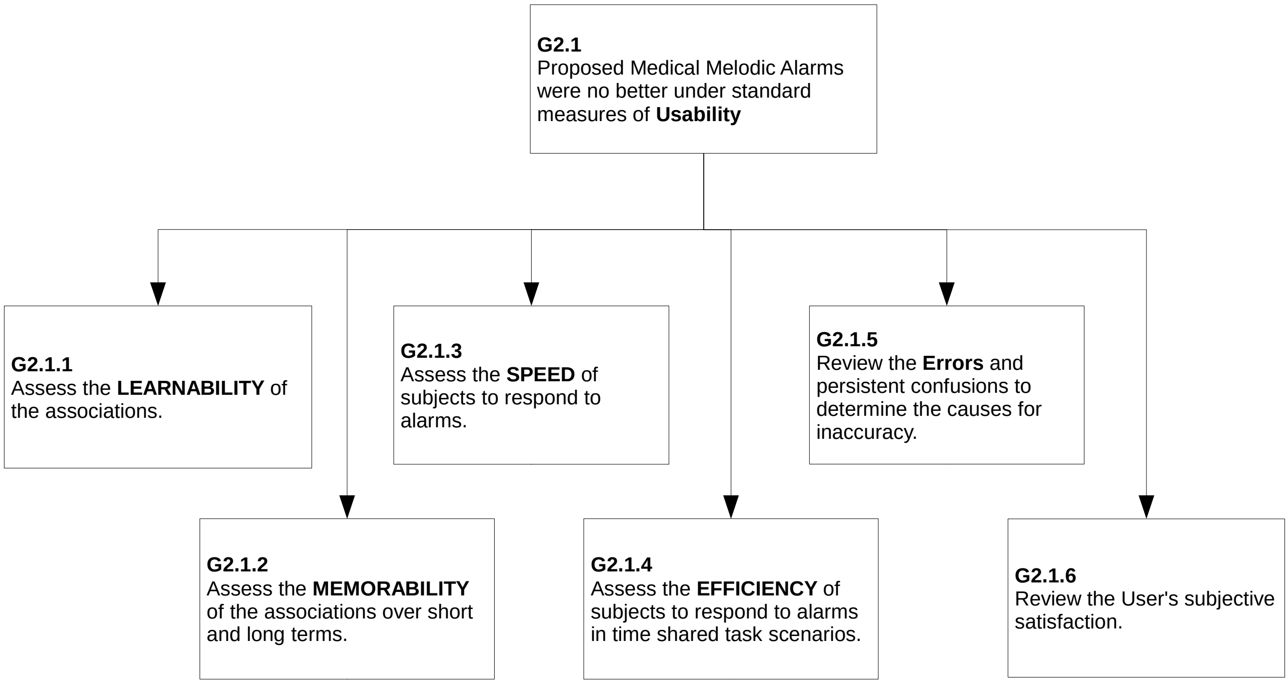


Figure 5-4: Methods of measurements for Experiments 1 and 2

The primary goal of this study was to discover if the supplied memory aids (Mnemonics) had any significant effect on a person's ability to:

- Learn the alarms and their associated labels(G2.1.1);
- Remember the alarms and their associated labels(G2.1.2);
- Identify the alarms in a timely manner (G2.1.3); and
- Correctly identify the alarms while being otherwise occupied (G2.1.4).

Secondary goals of this study included an analysis of:

- The effect of formal musical training on a person's performance on the above (G2.1.1.1.2, G2.1.2.1.2, G2.1.3.1.2 and G2.1.4.1.2);
- The participants self-reported satisfaction and confidence levels (G2.1.5); and
- The type of errors made by participants (G2.1.6).

5.5.1 Ability to learn the alarms

The learning criterion of two consecutive sets of correct tests was not achieved by all participants at the end of Day 1. 8/19 of participants on the Mnemonic condition and 4/19 of the participants in the Non-Mnemonic condition achieved the criterion. There was no significant effect of learning condition on accuracy. Participants in the Mnemonic group achieved a mean accuracy of 11.36 (71%) in each block and participants in the Non-Mnemonic group achieving a mean accuracy of 10.4(65%) in each block (p=0.35).

Experiment 1 - Day 1 (Accuracy) | ANOVA with regression

	SS	*Df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.645	1	0.645	24.065	2.28E-05	Yes
Mnemonics	0.024	1	0.023	0.889	0.352	No
Inter	0.015	1	0.015	0.566	0.456	No
Within	0.912	34	0.026			
Total	1.609	37	0.043			

<u>Mean</u>

	Mnemonic	Non-mnemonic	Mean
Musical	0.826	0.733	0.779
Non-Musical	0.515	0.504	0.509
Mean	0.670	0.618	0.644

<u>Variance</u>

	Mnemonic	Non-mnemonic	Mean
Musical	0.016	0.023	0.021
Non-Musical	0.033	0.044	0.036
Mean	0.045	0.042	0.043

Table 5-1: Experiment 1-Day 1 (Accuracy): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

However, analysis of the breakdown of musically trained and non-musically trained participants (Table 5-1) indicates that those with formal musical training were significantly more accurate than those without, with participants with at least a year of formal music training achieved a mean accuracy of 11.87 (78%) in each block and participants who had no formal music training achieved a mean accuracy of 8.02 (51%) in each block (p<0.0001).

A linear regression analysis of correct responses by participants in the mnemonic and non-mnemonic conditions in Day 1 (Figure 5-5) showed no significant variance in rate of learning.

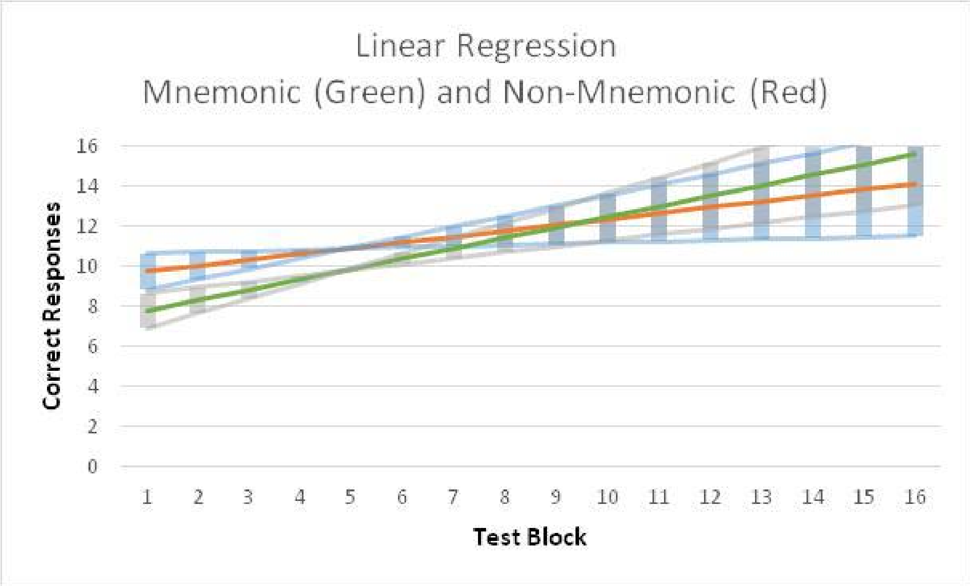


Figure 5-5: Experiment 1- Day 1 (Accuracy): Linear Regression of accuracy responses of mnemonic vs. non mnemonic participants during the first day of learning (Error bars represent the standard error in the slope)

A linear regression analysis of correct responses by musical and non musical participants in Day 1 (Figure 5-6) showed that although musical participants started and ended with better performance values than non-musically trained participants, the learning rate between the 2 populations was very similar (slope: Musical=0.53; Non musical=0.34).

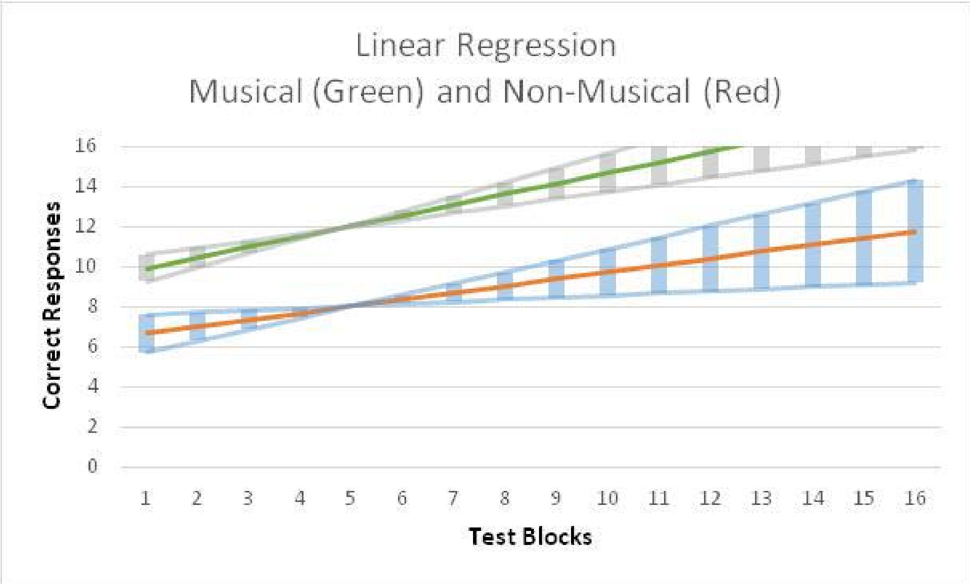


Figure 5-6: Experiment 1- Day 1 (Accuracy): Linear Regression of accuracy responses of musical vs. non musical participants during the first day of learning (Error bars represent the standard error in the slope)

On the second day of training, approximately a week after the first session, the participants were given a short refresher course to relearn the alarms to determine if their learning ability was improved. Only 5/18 of the Mnemonic participants and 4/15 of the non-mnemonic participants achieved the learning criterion before the end of the seven learn test cycles.

There was no significant effect of learning condition on accuracy after the second day of learning. Participants in the Mnemonic group achieved a mean accuracy of 13.44 (84%) in each block and participants in the Non-Mnemonic group achieving a mean accuracy of 12.32 (77%) in each block.

Experiment 1 - Day 2 (Accuracy) | ANOVA with regression

	SS	*Df*	*MS*	*F*	*p-value*	*Sig*
Musicality	0.433	1	0.433	14.164	0.001	Yes
Mnemonics	0.02	1	0.020	0.661	0.422	No
Inter	0.008	1	0.008	0.292	0.592	No
Within	0.887	29	0.03			
Total	1.352	32	0.042			

<u>Mean</u>

	Mnemonic	Non-mnemonic	Mean
Musical	0.867	0.850	0.859
Non-Musical	0.663	0.577	0.620
Mean	0.765	0.713	0.739

<u>Variance</u>

	Mnemonic	Non-mnemonic	Mean
Musical	0.007	0.024	0.013
Non-Musical	0.044	0.077	0.057
Mean	0.027	0.061	0.042

Table 5-2: Experiment 1-Day 2 (Accuracy): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

Analysis of the breakdown of musically trained and non-musically trained participants indicates that those with formal musical training were significantly more accurate than those without. Participants with at least a year of formal music training achieved a mean accuracy of 14.24 (89%) in each block and participants in who had no formal music training achieved a mean accuracy of 10.4 (65%) in each block ($p<0.0008$).

5.5.2 Ability to remember the alarms and their associated labels.

The ability to remember the alarms and their labels was measured by comparing the participants' ability to recall the alarms after being approximately a week away from their initial training. Average accuracy was taken for the initial two test sessions of the first day, for the last two test sessions of the first day, for the first 2 test sessions of the second day and for the last two test sessions of the second day.

The participants' data suggests that after initial accuracy improvement between the start of the first day and the end of the first day (measure of learning ability as discussed in Section 5.5.1) there is significant information loss between the end of Day 1 training and their return on the second day. This appears to occur regardless of whether the participant is in the Mnemonic or Non Mnemonic conditions (Mnemonic: Day 1 start = 57%, Day 1 end =81 %, Day 2 LTM = 70% and Day 2 end = 79%; Non-Mnemonic: Day 1 start = 44%, Day 1 end =78 %, Day 2 LTM = 65% and Day 2 end = 73%; see figures 5-7 and 5-8). This result suggests that the Mnemonics do not assist in aiding long term memory of the alarms.

There was significant improvement in accuracy within each day of testing (Day 1 start vs. Day 1 end, $p < 0.001$; Day 2 LTM vs. Day 2 end, $p < 0.001$). However, a Tukey HSD test of differences between the means for each learn-test phase showed that accuracy dipped significantly between the end of Day 1 and the beginning of Day 2 ($p < 0.001$). The end accuracy on both days was not significantly different ($p = 0.8$).

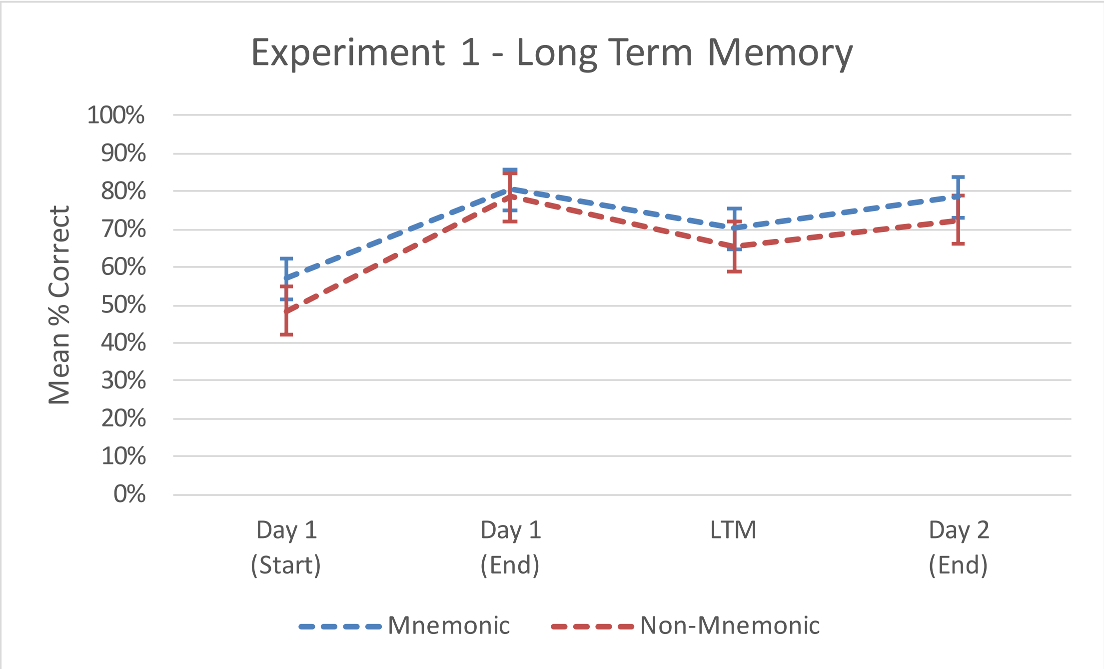


Figure 5-7 – Experiment 1 Results for learning trend at the beginning and end of each of two days of testing by learning condition (error bars represent the standard deviation)

Analysis of the same data but differentiated for musical training shows that musically trained participants achieved greater accuracy than non-musically trained participants in all phases of the experiment (Musical: Day 1 start = 61%, Day 1 end =91 %, Day 2 LTM = 78 % and Day 2 end = 85%; Non-Musical: Day 1 start = 38%, Day 1 end =59 %, Day 2 LTM = 52 % and Day 2 end =60%). However the rate of learning and retention does not significantly differ.

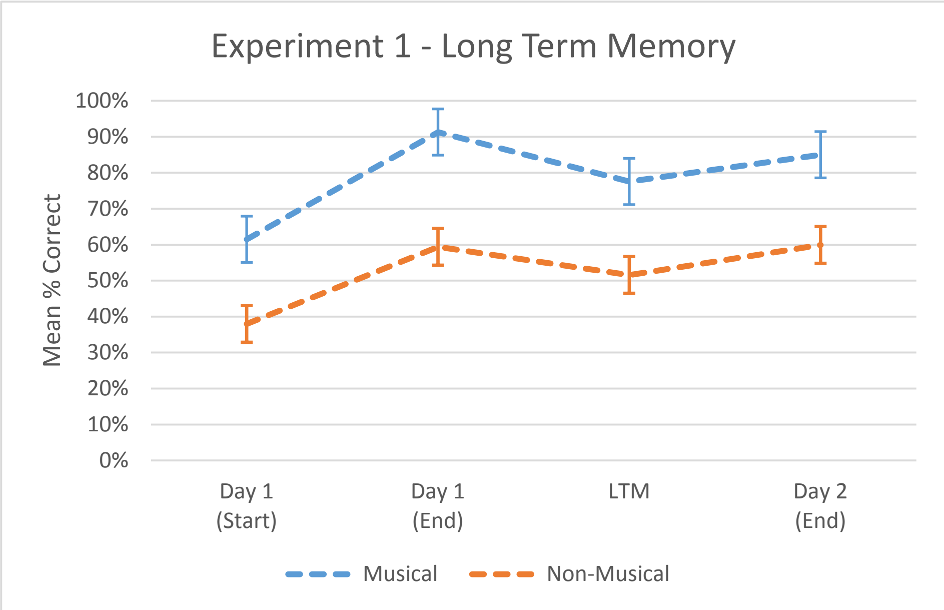


Figure 5-8 Experiment 1 Results for learning trend at the beginning and end of each of the two days of testing by musical training (error bars represent the standard deviation)

5.5.3 Ability to identify the alarms in a timely manner

One of the goals of any alarm design is to improve the time taken for an individual to react to the alerting situation. The included learning aids were designed to provide a prompt to allow the individual to make a quick cognitive association between the melodic tune and the alarm identity in order to apply any necessary corrective action (Block et al, 2000).

There was no significant effect of learning condition on response speed (Table 5-3). Participants in the Mnemonic group achieved a mean response time of 5.92s and participants in the Non-Mnemonic group achieved a mean response time of 6.58s.

Experiment 1 - Day 2 (Response Time) | ANOVA with regression						
	SS	*df*	*MS*	*F*	*p-value*	*Sig*
Musicality	15.051	1	15.051	3.600	0.066	No
Mnemonics	8.949	1	8.949	2.140	0.152	No
Inter	15.607	1	15.607	3.733	0.061	No
Within	142.139	34	4.180			
Total	176.892	37	4.780			

<u>Mean</u>

	Mnemonic	Non-mnemonic	Mean
Musical	5.930	5.607	5.769
Non-Musical	5.906	8.241	7.073
Mean	5.918	6.924	6.421

<u>Variance</u>

	Mnemonic	Non-mnemonic	Mean
Musical	6.739	1.745	4.085
Non-Musical	1.362	6.771	5.221
Mean	4.572	5.027	4.780

Table 5-3: Experiment 1-Day 2 (Response Times): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

Analysis of the breakdown of musically trained and non-musically trained participants indicates that those with formal musical training were quicker to respond than those without. However this was not found to be significant. Participants with at least a year of formal music training achieved a mean response time of 5.77s in each block and participants in who had no formal music training achieved a mean response time of 7.07s in each block (p=0.066).

Response time was smaller for medium priority alarms (average 5.56s) than for high priority alarms (average 6.94s) (p = 0.009). However, this result is not wholly unexpected as medium priority alarms are faster to play (and therefore listen to and subsequently process) than high priority alarms as well as being simpler to interpret.

Data Set	Participants	Mean Time Taken by Trial	p Value	t Stat	StDev	dF
High Priority	38	6.94	9.38E-03	2.289	2.446	36
Medium Priority	38	5.56		2.289	1.959	36

Table 5-4: Experiment 1 Response time analysis of responses of high priority and Medium priority alarms during the first day of learning.

Interactions of alarm priority, learning phase, condition and musical training indicates that non musically trained participants in the non-mnemonic condition showed much less improvement in speed of responses from the start to the end of Day 1 compared with participants in the mnemonic condition and all other musical participants (p = 0.017).

Figure 5-9 shows a significant speed up across all phases except from the end of Day 1 to the start of Day 2 (p = 0.91).

a) Mnemonic b) Non mnemonic

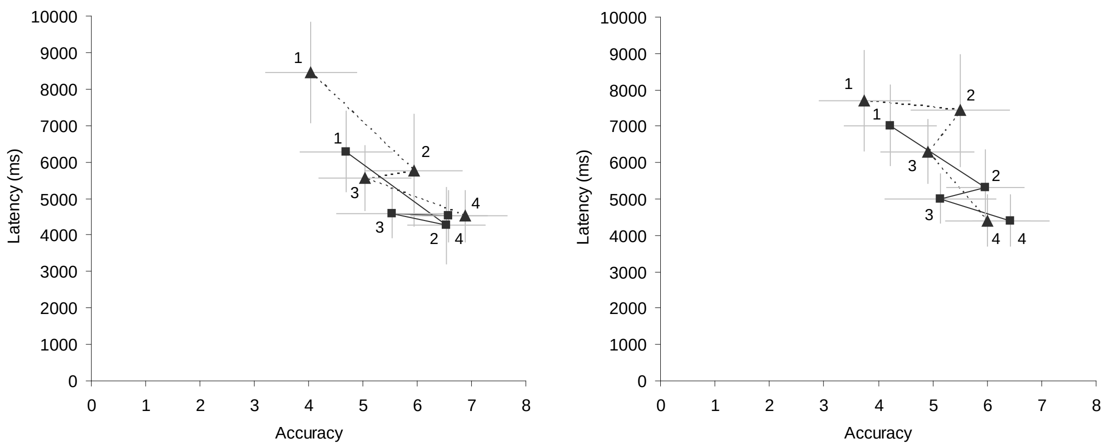


Figure 5-9 - Response latency against accuracy over phases. Accuracy is out of a maximum possible eight correct for high (--▲--) and medium (—■—) priority alarms. 1 = Day-1 start, 2 = Day-1 end, 3 = Day-2 start, and 4 = Day-2 end. Error bars (in grey) are 95% CI. The bottom right of each figure represents fast, accurate performance.

5.5.4 Ability to correctly identify the alarms while being otherwise occupied

In practical application the ability of the alarms to direct attention requires that it be measured against a situation where the individual is otherwise occupied and that attention can be directed towards the correct physiological system. Therefore, in the final part of this experiment, the participant was required to perform a simple math task while a random alarm was played at random intervals. These results were compared by learning condition and prior musical training.

Once again there was no observable effect of learning condition (Table 5-5). Participants in the Mnemoninc group achieved a mean accuracy of 5.28 (33%) in each block and participants in the Non-Mnemonic group achieved a mean accuracy of 5.6 (35%) in each block.

Experiment 1 – Distractor (Accuracy) | ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.108	1	0.108	4.541	0.041	yes
Mnemonics	0.001	1	0.001	0.077	0.783	no
Inter	0.007	1	0.007	0.309	0.582	no
Within	0.694	29	0.023			
Total	0.810	32	0.025			

Mean

	Mnemonic	Non-mnemonic	Mean
Musical	0.359375	0.40625	0.382813
Non-Musical	0.270833	0.255208	0.263021
Mean	0.315104	0.330729	0.322917

Variance

	Mnemonic	Non-mnemonic	Mean
Musical	0.022283	0.021729	0.021512
Non-Musical	0.022917	0.032194	0.025117
Mean	0.023003	0.029781	0.025315

Table 5-5: Experiment 1 – Distractor (Accuracy): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

Musically trained participants were slightly more accurate (38% correct) than non-musically trained participants (26%) but the calculated statistic was marginal at p=0.041.

Experiment 1 - Distractor (Response Time) | ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*Sig*
Musicality	11.457	1	11.457	1.580	0.218	no
Mnemonics	30.476	1	30.476	4.203	0.049	yes
Inter	7.181	1	7.181	0.990	0.327	no
Within	210.278	29	7.250			
Total	252.574	32	7.892			

Mean

	Mnemonic	Non-mnemonic	Mean
Musical	5.338	4.306	4.822
Non-Musical	7.541	4.562	6.051
Mean	6.439	4.434	5.437

Variance

	Mnemonic	Non-mnemonic	Mean
Musical	2.294	4.471	3.324
Non-Musical	21.024	8.829	15.989
Mean	8.810	5.725	7.893

Table 5-6: Experiment 1 – Distractor (Response Time): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

A marginally significant effect of learning condition was observed on response latency during the timeshared task (Mnemonic = 5.5s, Non-Mnemonic = 4.4s; p = 0.049) No significant effect was observed in musical training (musically trained = 4.9s, not musically trained = 6.1s, $p = 0.36$). Response latency was also observed to have changed between absolute identification testing and timeshared task testing ($p < 0.001$).

5.5.5 Self-reported satisfaction and confidence levels

Participants were asked to rate the relative urgency of the alarm sounds on 7 point Likert scales and rated the high priority alarms as sounding more urgent than the medium priority alarms (High priority = 4.8, Medium priority =3.1, $p <0.001$). Musically trained participants reported that it was easier to distinguish the alarms (Musically trained = 4.8, not musically trained = 2.9, $p < 0.001$).

At the end of the second day, participants were asked to rank the alarms according how easy they thought they were to learn. There was a consensus that the General alarm was the easiest to learn with nearly all the participants assigning it a value of 1. The two decreasing tones of Powerfail and Oxygen were ranked the next easiest followed by Cardiovascular, Temperature and perfusion alarms. The final group of alarms were Ventilation and Infusion which were ranked as the hardest alarms to learn.

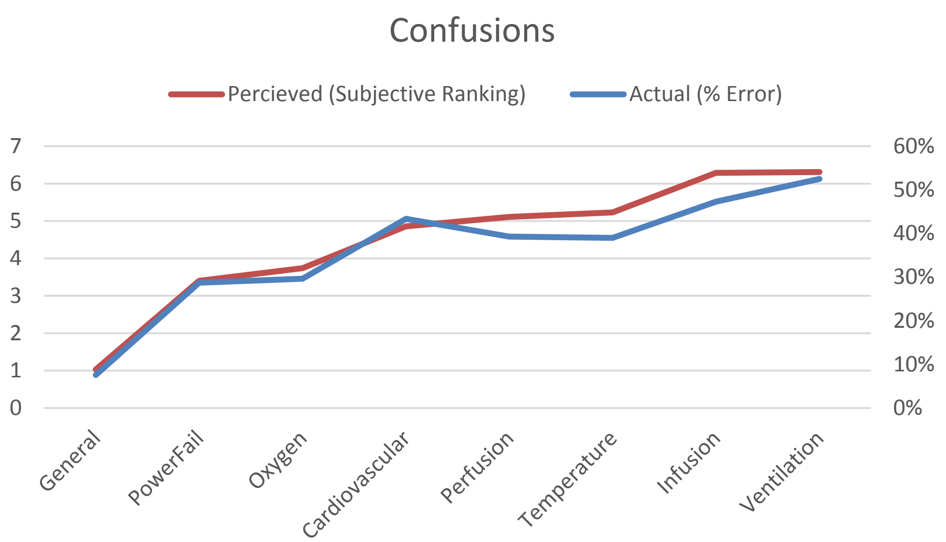


Figure 5-10: Experiment 1 Participants ranking of each alarm on the ease with which they found the alarm to learn (1-Easiest 8 hardest); paired against the actual difficulty responding to each alarm

Participants were also asked to report if they had utilised any learning aids to help them remember the alarms and their meanings. In addition, participants in the mnemonic condition were asked to report of they found the recommended mnemonics useful in helping them remember the alarm sounds. Nearly all participants in both learning conditions (35 out of 38) reported that they designed their own learning aids based on familiar associations such as “being similar to their mobile phone ringtone” and “similar melody to a song”. None of the mnemonic participants reported using all the supplied mnemonics to aid their learning but some mnemonics were found to be more useful than others.

Alarm Name	Participants who found the supplied Mnemonic useful (out of 19 participants)
General	0
Oxygen	5
Powerfail	9
Cardio	6
Temp	6
Perfusion	3
Infusion	6
Ventilation	8

Table 5-7: Number of participants in the Mnemonic learning group who reported using the supplied mnemonic for each alarm. Total participants in Mnemonic learning condition = 19.

5.5.6 Analysis of Errors and Confusions

Participants showed very little confusion between high and medium priority alarms (less than 3%) indicating that the difference in patterns effectively distinguishes one level of urgency from another. There was less confusion between alarms, in both conditions on Day 2 than on Day 1 but many of the same patterns of confusions remained. In an analysis of the confused alarms made on more than 25% on a participant’s trials we see clear indications that the Cardiovascular (CV) alarm and the Temperature (TE) alarm are mutually confused by many participants as are the Infusion (IN) alarm and the Ventilation (VN) alarm. However, there was no significant difference between the numbers of incorrect responses between conditions. Figure 5-11 shows the confusions made on Day 2 learn test phases by participants in the Mnemonic and Non-Mnemonic conditions. Numbers on the arrows indicate the percentage of participants who on more than 25% of the Day 2 test trials confused the alarm at the start of the arrow with the alarm at the end of the arrow. Participants in the Mnemonic condition showed a narrower range of confusions than those learning under the non-Mnemonic condition. In the Mnemonic condition 80% of the participants’ confusions were shared by at least one other participants whereas in the NM condition only 29% of participants' confusions were shared by at least one other participant (p = 0.011).

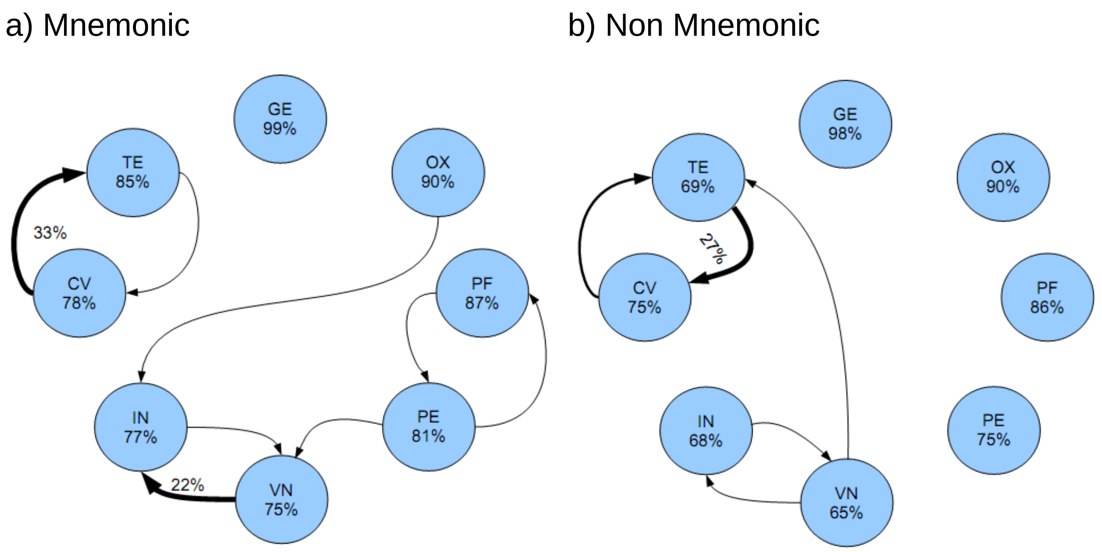


Figure 5-11 - Pattern of confusions between alarms from Mnemonic and Non-Mnemonic participants. Participants misidentified the sound at the origin of the arrow as the alarm at the end of the arrow on more than 25% of the trials. Three arrow weights are shown. The thinnest indicates confusions made by at least 2 participants followed by confusions made by at least 3 particiapnts. Confusions made by 4 or more participants are presented as the thickest arrows with the precentage of of participants who misidentified the alarms presented on the links. Values inside the circle are the proportion correctly answered per alarm.

The pattern of confusions was also plotted for musical and non-musical participants. Musical participants were observed to have a narrower range of repeated confusions as compared to non-musically trained participants. Accurate identification of each alarm was also significantly better in musical participants than non-musical participants (Figure 5-6).

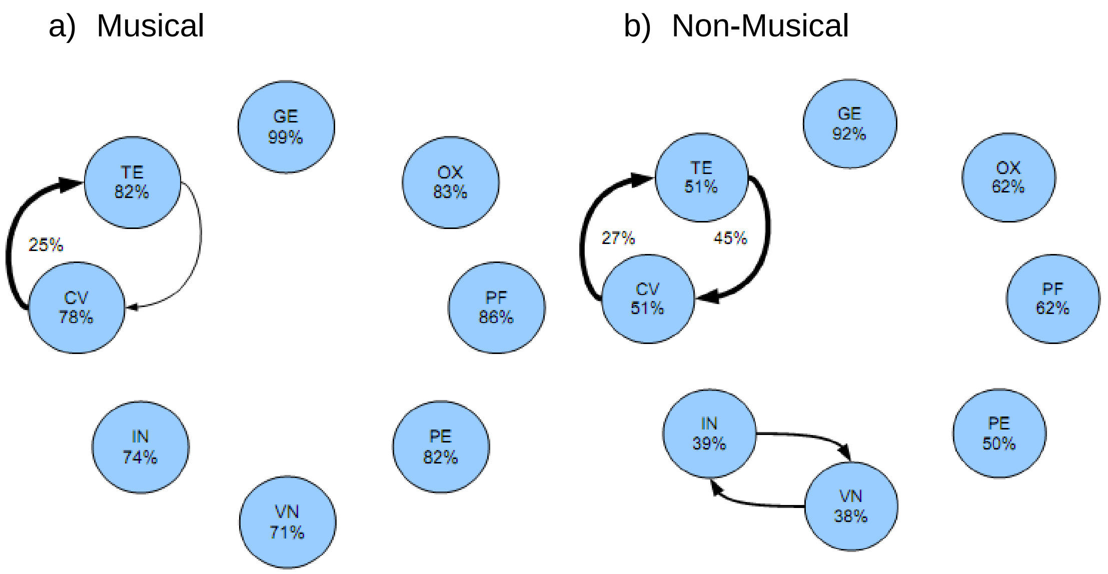


Figure 5-12: Patterns of confusions between alarms from Musical and Non Musical Participants in Experiment 1. Values inside the circle are the proportion correctly answered per alarm.

5.6 Conclusion

At the end of two learning sessions fewer than 30% of the students could meet our criteria of 100% accuracy in two successive tests. The mnemonics did not seem to have any effect on the accuracy and speed of participants responding to the alarms. This could be attributed to the mnemonics being culturally and demographically specific. For example, the NBC chime and the sound of a Victrola are more familiar to older North Americans than most other demographics. In fact, most of the participants in the mnemonic condition had to ask the experimenter to explain one or more of the mnemonics when they were displayed on the screen. The only difference observed between the Mnemonic and Non-mnemonic conditions was that participants in the Mnemonic condition showed a narrower range of confusions between alarms even though the numbers of confusions were about equal between the conditions. Some confusions persisted over the 2 days of testing, in particular confusions between the Cardiovascular and Temperature alarms and the Infusion and Ventilation alarms. We attribute these confusions to the similar musical contours of each pair of tunes. If occurring in an operating room, these confusions could have real consequences ranging from a delayed response to a dangerous situation to something more serious such as the administration of the wrong drugs or corrective treatment. It is therefore necessary to see if these confusions occur in tests with healthcare staff to determine if changes should be made to the alarms involved before they can be safely differentiated.

Participants managed to easily distinguish between high and medium priority alarms with confusions between the priorities occurring in only 3% of the trails. The IEC melodic alarms have managed to do this without variation in tonal features such as loudness or timbre which is what some researchers argue to be the way to convey varying levels of urgency (Mondor & Finley, 2003). However, there is concern that our participants were faster and more accurate in recognising the medium priority alarms than the high priority alarms. The high priority alarms should be more quickly and accurately identified because they signify more dangerous events. The reason for this could be because the high priority alarms contain more notes, including 2 pulses at the end of each sequence which do not map to the first three pulses in the medium priority alarm. These are played at a faster pace and thus require more time and cognitive effort to process before participants can retrieve the correct label from memory.

An interesting, though not altogether unexpected finding (based on known research on relative pitch sensitivity (McDermott et al., 2008)) was that participants with formal musical training identified the melodic alarms almost 50% more accurately and found it easier to learn than those who no formal musical training. This is due to the fact that musical training enables a person to listen for and distinguish more subtle changes in tone and remember them (Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). As we cannot depend or expect clinicians to have had musical training, the effectiveness of these alarms should not depend on clinicians having had it.

These results suggest a cause for concern that the IEC 60601-1-8 melodic alarms might cause confusions when used in the clinical context. Our results show persistent confusions remain even after extensive training which are not lessened by the use of mnemonics. Moreover the data on participants' ability to recognise the alarms while performing a timeshared task indicate that accuracy drops when the amount of workload increases.

The purpose of this study was to make a preliminary determination if there was a cause for concern with the IEC 60601-1-8 melodic alarms. The results have shown that in general there was a need for more thorough user studies and proper human factors evaluations performed *before* novel interface designs such as the IEC 60601-1-8 melodic alarms are presented as an international medical equipment standard. As the participants in this study are non-clinicians, we have yet to see if the medically related mnemonics aid medically trained clinicians in identifying and discriminating the alarms. This point was subsequently explored and is presented further in the next chapter of this dissertation.

6 EXPERIMENT 2 – EFFECT OF MNEMONICS (NURSES)

The design of Experiment 2 is similar to that of Experiment 1. However, in place of first year psychology students, nurses from two major hospitals in the Brisbane metropolitan area were the participants in the study. Slight modifications were made to the protocol to take into account the medical expertise of the participants.

The purpose of conducting this experiment was two-fold. The first purpose was to determine whether there was an effect of expertise in identifying alarms which have been designed to represent specific organ systems. The second purpose is to see if subtle effects of learnability and memorability observed in the first study are significant with greater statistical power.

By running a sample of professional healthcare workers in these experiments we will be able to tell if some of the problems with melodic alarms revealed in the previous study are relevant to the healthcare population.

6.1 Method

The method used was similar to the experimental method used in Experiment 1 with some modifications:

- Participants were asked to provide information about their nursing experience;
- Following the initial familiarisation phase, participants were allowed to listen to the alarms as many times as they wished during the learning phase on the learning screen (Figure 6-1 and **Error! Reference source not found.**) until they indicated that they were comfortable to proceed to the testing phase by hitting the button on the bottom of the screen; and
- The software program was modified to allow the timeshared cognitive task and the alarm response frame to appear in the same screen. Thereby reducing the need for participants to turn their head. This combination also allowed the experimenter to collect accurate timestamp data for participant's responses to the math task and the responses to the random alarms.

Participants

Participants were 22 registered nurses from hospitals in the Brisbane metropolitan area, recruited through contacting nursing directors of the various hospitals and through flyers and announcements. The nurses were between the ages of 21 and 58 years of age (Average age 35). Six were male and sixteen were female with between 1 and 36 years of nursing experience (average 13 years). All nurses were involved in critical care fields such as Intensive Care (ICU), Cardiac Care (CCU) and Operating Rooms. Participants were asked if they had any formal musical training and for how many years (Musically trained = 8 [36.4%], Non-musically trained = 14 [63.6%]). No participants had previous experience with the IEC 60601-1-8 melodic alarm sounds.

A power analysis based on data from Experiment 1 indicated that with 22 participants a 25% improvement in performance with mnemonics would be detected with statistical power of 0.71 and a 25% improvement in performance with musical trained participants would be detected with statistical power of more than 0.9.

Participation was voluntary and participants were compensated for their time with a token payment of AUD$40 for full participation of 3 hours over 2 days or AUD$10 for participation on the first day (1 hour) only. All 22 participants included in this study participated in both sessions. One recruited participant participated only in the first day's session but decided that she was unable to understand the mnemonics and instructions due to her command in English not being strong enough. She was compensated for her time but her data was excluded from the sample.

The study was conducted in accordance with the ethical review processes of the University of Queensland, cleared with the ethical board of the Princess Alexandra Hospital and within the guidelines of the National Health and Medical Research Council.

6.2 Apparatus and Stimuli

Due to a slight correction in timing of the alarm rise and fall, and to be completely consistent with the standard, the alarms used in Experiment 1 were not used in this study. Instead, the "3-2, 3-2" pattern for high priority alarms, and the "3" pattern for medium alarms were created within the constraints specified in the IEC 60601-1-8. Pulse rise time

(time for 10% pulse amplitude to 90% pulse amplitude) and pulse fall time for each tone within each high priority alarm was 10ms and 20ms for each medium priority alarm. For high priority alarms, the pulse sustain time (time as maximum amplitude) was 98 ms and pulse width ([pulse sustain time] + [10% pulse rise time] + [10% fall time]) was 100ms. The gap of silence between pulses within each group of notes was 78 ms. The interval between the start of the fall time for each pulse and the end of the rise time for the next pulse (90% amplitude of pulse 1 to 90% amplitude of pulse 2) was 100ms. Spacing between the 3rd and 4th pulse was 278 ms. Spacing between the 5th and 6th pulse was 680 ms. The total duration of high priority alarm was 2924 ms. For the medium priority alarms, the pulse rise times and pulse sustain times were twice as long as for the high priority alarms. Total duration of the medium priority alarms was 1044ms.

The sounds described above were created in Csound using middle C as the lowest note (C4: 278.4Hz) and the highest note was an octave higher (C5: 556.8Hz). The alarms were processed on the integrated soundcard of a Pentium 4, 1.9 GHZ Acer Laptop and presented on AKG K 240DF studio Monitor earphones attached to the laptop. During the introduction and familiarization phase of the experiment, the experimenter checked that the sound level was clear and comfortable for the participant in order to screen out potential unreported hearing difficulties. No participant reported discomfort and difficulty hearing the alarms and no changes were made to the volume of the alarms which were played at the same volume for each participant.

Alarm labels and mnemonics were displayed on the laptops 17 inch flat screen display at 1280 x 1024 screen resolution by a custom written Java programming which presented the learning (Figure 6-1 and **Error! Reference source not found.**) and test screens (**Error! Reference source not found.**) as well as recording the participants responses. Responses and response times for each participant were recorded as individual comma separated values (.csv) files.

Figure 6-1: Learning Screen from Experiment 2 (No Mnemonic Condition)

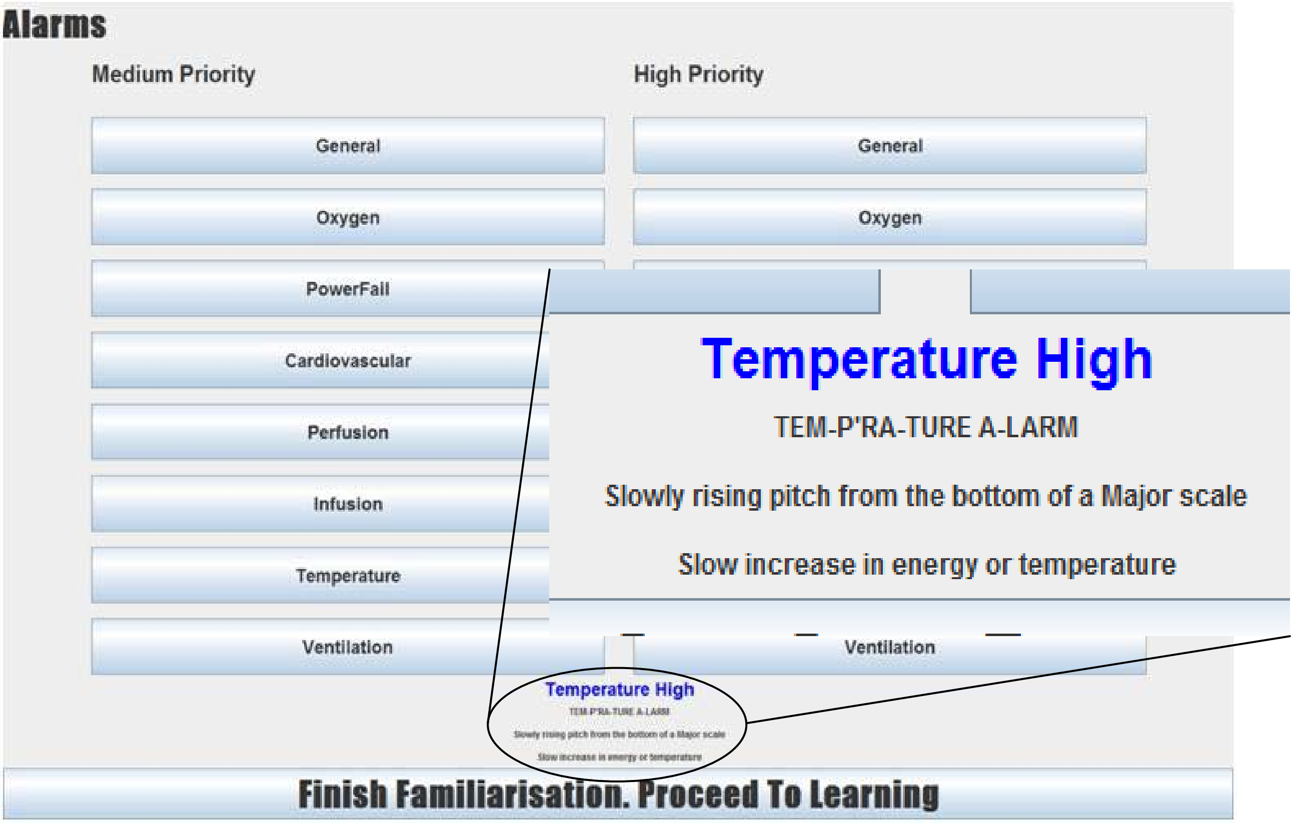


Figure 6-3: Learning Screen from Experiment 2 (Mnemonic Condition)

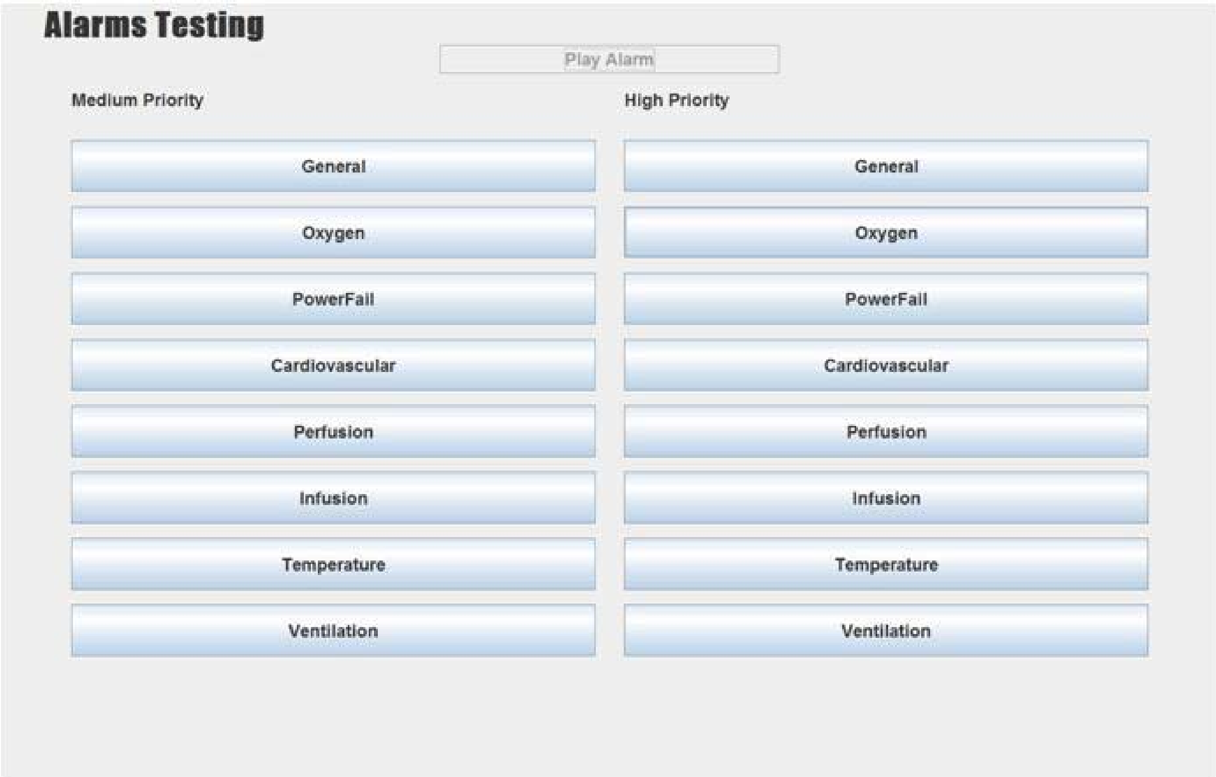


Figure 6-2: Subjects self testing screen for Experiment 2

6.3 Procedure

The procedure is very similar to that described in Experiment 1. However, several subtle differences have been implemented that warrant detailed description.

Prior to the start of the experiment, participants were asked to complete a short questionnaire which asked about their age, any known hearing impediments, experience in the nursing profession, current nursing domain and whether they had had any formal musical training and if so, what instrument and for how many years.

The nurses were pseudo randomly allocated into either the Mnemonic group (n =11) or the Non-Mnemonic group (n=11) to keep years of experience and musical training relatively equal. The nurses were trained and tested with the melodic alarms in 2 separate sessions on different days spaced 6-11 days apart as described below.

Day 1: After being seated in a quiet room and written informed consent was obtained, participants were introduced to the 16 alarm sounds (8 medium priority and 8 high priority). They were told the identity of each alarm (e.g." High priority oxygenation" or "Medium priority cardiovascular"). The Nurses in the M group are also familiarised with the mnemonics as described in Table 3-1. The experimenter also played each alarm melody for the M participants and explained the musical and medical mappings to each alarm sound. NM participants simply learnt the alarm label and were never introduced to the mnemonics.

The learning-test phase of the experiment were exactly as detailed in Experiment 1, with the exception that unlike the students, the nurses were able to listen to the alarms as many times as they wanted until they felt comfortable with proceeding to the test phase.

This cycle of learning and testing continued until either the participant reached the learning criterion of two consecutive sets of perfect test scores or 45 mins of the experiment time had elapsed. No learn-test cycle was left incomplete.

Day 2 (Approx 1 week later): Within 6-11 days the nurses would return for a second session. As with Experiment 1, the first phase of Day 2 was a long-term memory (LTM) test in which the participants would go through 2 sets of tests phases as experienced during the first day of the experiment. However, unlike the first day they were not given the opportunity to relearn the alarms.

The second phase of Day 2 was similar to the learn-test cycles in Day 1. These learn-test cycles continued until the participants achieved 2 consecutive perfect scores or if 35 mins of the learn-test cycle has passed or the learn-test cycle has been completed eight times.

The final phase of the experiment on Day 2 required the nurses to identify the alarms while performing a timeshared arithmetic true/false task (Figure 6-4). The participant saw a simple equation such as “2+2=4” or “4-2=12’ and responded by pressing the “P” key if they thought that the equation was true and the “Q” key if they thought that the equation was false. Both keys and their mappings were clearly displayed on the screen for the participants’ references at all parts of this stage. At quasi-random times during this task, between 21 and 55 seconds apart, a melodic alarm would sound. The participant clicked the button on the screen representing the alarm that they had thought they heard. The alarms were played in random order with each alarm occurring only once in each of the three trials.

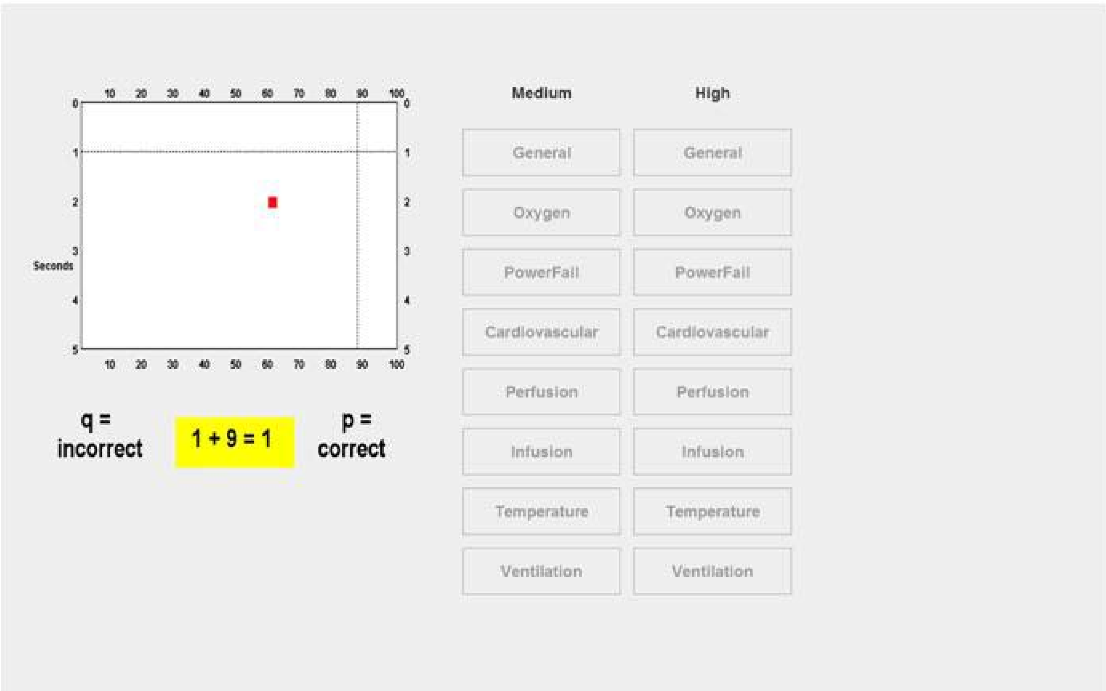


Figure 6-4: Timeshared Arithmetic Task Screen for Experiment 2

Questionnaires were given to the participant at the end of Day 1, the end of phase two on Day 2 (learn-test cycles), and at the end of Day 2. The nurses were asked to respond to questions such as the ease of learning the alarms, the ease of associating alarms meanings with their sounds, and the relative urgency of the medium and high priority alarms, by rating them on a 7-point Likert scale. The questionnaires were paper based and the experimenter later manually entered participant's responses for further analysis.

6.4 Results

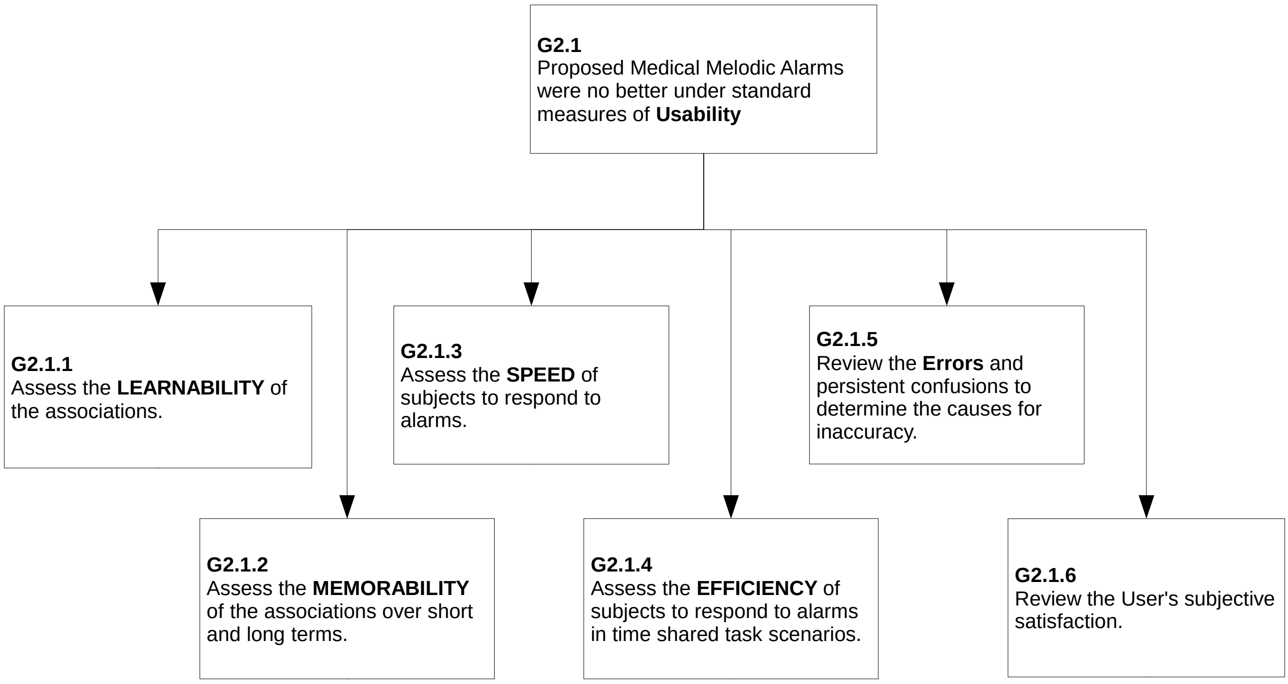


Figure 6-5: Methods of measurements for Experiments 1 and 2

The primary goal of this study was to discover if the supplied memory aids (Mnemonics) had any significant effect on a skilled professional's ability to:

- Learn the alarms meanings(G2.1.1);
- Remember the alarm meanings (G2.1.2);
- Identify the alarms in a timely manner (G2.1.3); and
- Correctly identify the alarms while being otherwise occupied (G2.1.4).

Secondary goals of this study included an analysis of:

- The effect of formal musical training on a person's performance on the above (G2.1.1.2.2, G2.1.2.2.2, G2.1.3.2.2 and G2.1.4.2.2);
- The participants self-reported satisfaction and confidence levels (G2.1.5);
- The type of errors made by participants (G2.1.6);
- The responses of the general population participants in Experiment 1 against the skilled professional participants reported in this experiment (G2.1.1.3.1, G2.1.2.3.1, G2.1.3.3.1 and G2.1.4.3.1); and
- The responses of musically trained participants in both populations against the non-musically trained participants (G2.1.1.3.2, G2.1.2.3.2, G2.1.3.3.2 and G2.1.4.3.2).

6.4.1 Ability to learn the alarms

No participants achieved the learning criterion on the Day 1 and only 2 participants (9%) achieved 100% correct identification on one test cycle. We found that the average accuracy across all the alarms other than the general alarm was poor (Accuracy range of correct answers: 37% to 82%; Average 57%). Accuracy performance of the general alarm notably better in the mnemonic group than the non mnemonic group (Mnemonic =95%; Non-Mnemonic =85%; Average 89%).

As with Experiment 1 we found no significant effect of learning condition on accuracy (Mnemonic = 50%, Non-Mnemonic = 47%, p = 0.59) even though the mnemonics were designed to aid medically trained clinicians. Participants with formal musical training were more accurate at identifying the alarms than those without (Musically-trained = 68%, non-musically trained =38%, p < 0.001).

Experiment 2 -Day 1 (Accuracy) - ANOVA with regression

	SS	*Df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.463	1	0.463	24.676	9.96E-05	yes
Mnemonics	0.005	1	0.005	0.305	0.587	no
Inter	7.64E-05	1	7.64E-05	0.004	0.949	no
Within	0.338	18	0.019			
Total	0.808	21	0.038			

Mean	Mnemonic	Non-mnemonic	Mean
Musical	0.693	0.663	0.678
Non-Musical	0.395	0.358	0.377
Mean	0.544	0.510	0.528

Variance	Mnemonic	Non-mnemonic	Mean
Musical	0.003	0.031	0.014
Non-Musical	0.025	0.014	0.019
Mean	0.038	0.041	0.039

Table 6-1: Experiment 2 – Day 1 (Accuracy): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

A linear regression analysis of correct responses by participants in the mnemonic and non-mnemonic conditions in Day 1 (Figure 6-6) showed no significant variance in rate of learning.

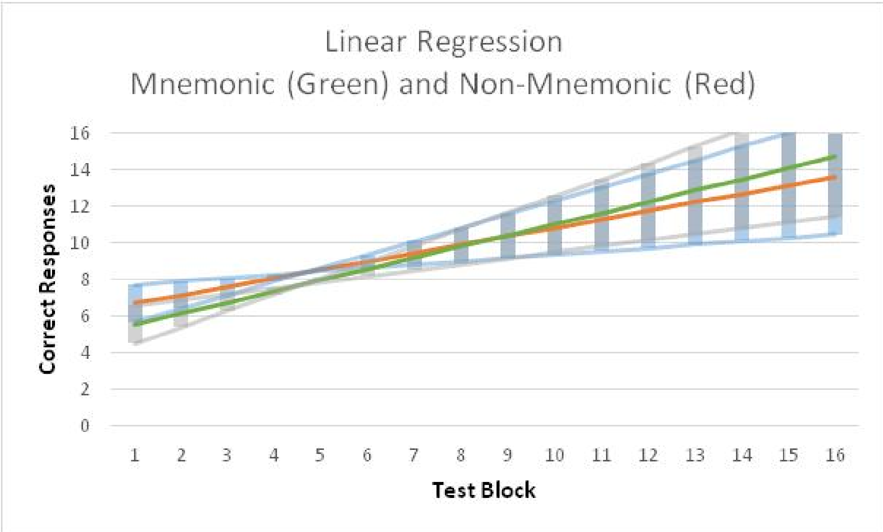


Figure 6-6: Experiment 2- Day 1 (Accuracy): Linear Regression of accuracy responses of mnemonic vs. non mnemonic participants during the first day of learning (Error bars represent the standard error in the slope)

A linear regression analysis showed that although musical participants started and ended with better performance values than non-musically trained participants, the learning rate between the 2 populations was very similar.

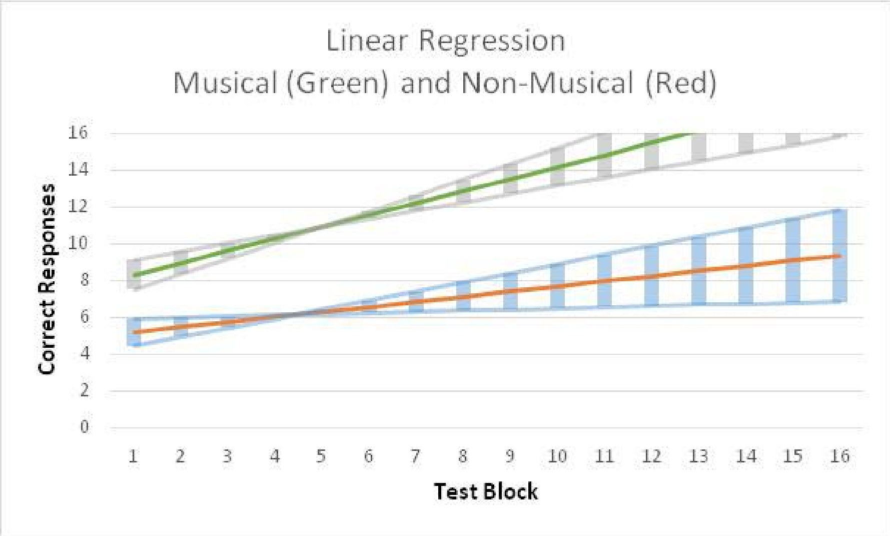


Figure 6-7: Experiment 2-Day 1 (Accuracy): Linear Regression of accuracy responses of musical vs. non musical participants during the first day of learning (Error bars represent the standard error in the slope)

At the end of the learning sessions of Day 2 only one participant had managed to reach our learning criterion of 100% correct identification of the 16 alarms in two consecutive test cycles and another 5 participants managed to achieve at least one 100% correct identified test trial. There was no observable effect of learning condition after the second day of learning. But once again musically trained participants were significantly more accurate than those without ($p<0.0001$).

Experiment 2 –Day 2 (Accuracy) - ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.752	1	0.752	28.152	4.81E-05	yes
Mnemonics	0.009	1	0.009	0.345	0.564	No
Inter	0.012	1	0.011	0.446	0.512	No
Within	0.481	18	0.026			
Total	1.262	21	0.06			

<u>Mean</u>

	Mnemonic	Non-mnemonic	Mean
Musical	0.834	0.84	0.837
Non-Musical	0.498	0.407	0.453
Mean	0.666	0.623	0.645

<u>Variance</u>

	Mnemonic	Non-mnemonic	Mean
Musical	0.007	0.005	0.005
Non-Musical	0.036	0.038	0.036
Mean	0.052	0.072	0.06

Table 6-2: Experiment 2 – Day 2 (Accuracy): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

6.4.2 Ability to remember the alarms and their associated labels

There was a difference in accuracy over each of the phases (Day 1 start =42%, Day 1 end = 59%, Day 2 LTM = 52% and Day 2 end = 68%, $p < 0.001$). There was an improvement in accuracy within each day and between the start of Day 1 and the end of day 2 (comparisons were significant at $p < 0.05$). Medium priority alarms were more readily identified than high priority alarms (medium = 58%, high = 53%, $p = 0.032$). There was no significant difference between the mnemonic and non-mnemonic conditions (Mnemonic: Day 1 start = 39%, Day 1 end =53 %, Day 2 LTM = 45% and Day 2 end = 55%; Non-Mnemonic: Day 1 start = 35%, Day 1 end =54 %, Day 2 LTM = 47% and Day 2 end = 53%)

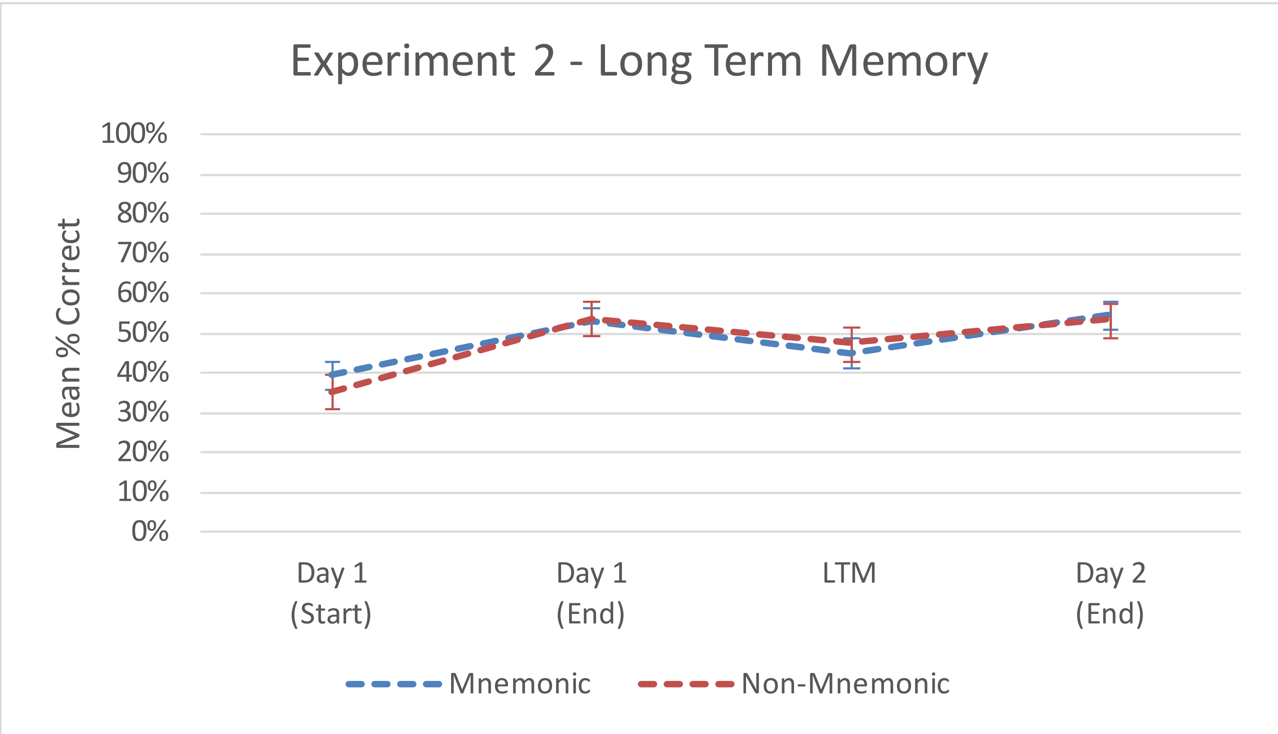


Figure 6-8: Experiment 2 Results for learning trend at the beginning and end of each of two days of testing by learning condition

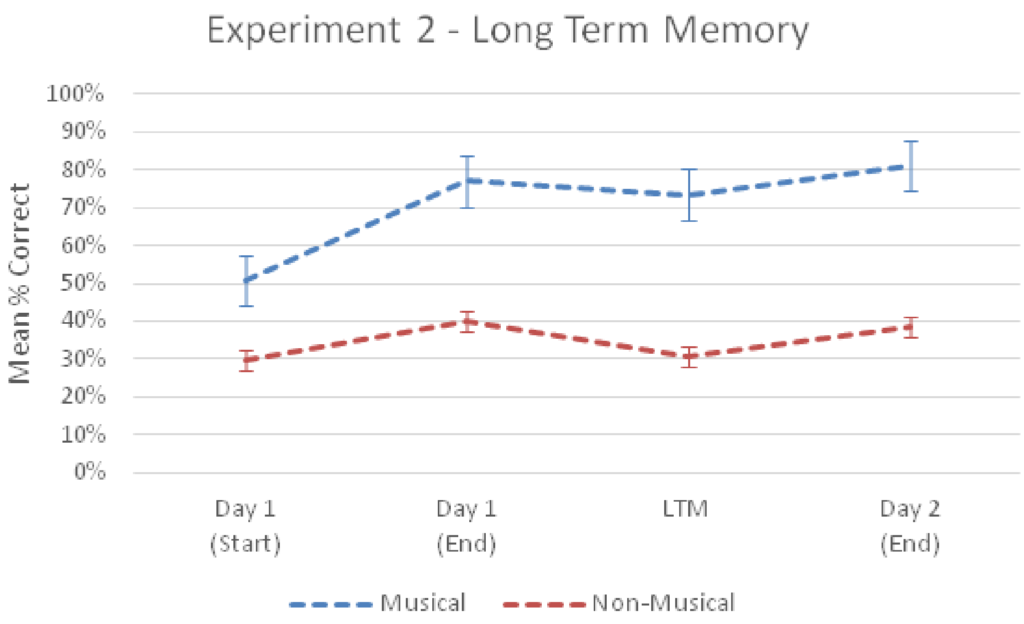


Figure 6-9: Experiment 2 Results for learning trend at the beginning and end of each of the two days of testing by musical training

Analysis of the same data but differentiated for musical training shows that musically trained participants are significantly better than non-musically trained participants in all phases of the experiment (Musical: Day 1 start = 51%, Day 1 end =77 %, Day 2 LTM = 73 % and Day 2 end = 81%; Non-Musical: Day 1 start = 30%, Day 1 end =40 %, Day 2 LTM = 31 % and Day 2 end =39%).

There was no significant effect of learning condition on response latency (M =6.0s, NM = 6.1s, $p = 0.9$) or musical training (musically trained = 5.6s, not musically trained = 6.5s, $p = 0.23$). Latency changed over phases (Day 1 start = 7.3s, Day 1 end = 5.7s, Day 2 LTM = 6.1s, Day 2 end = 5.1s, $p < 0.001$). Figure 6-10 shows significant speed differences only between the start of Day 1 and each of the other phases.

a) Mnemonic b) Non mnemonic

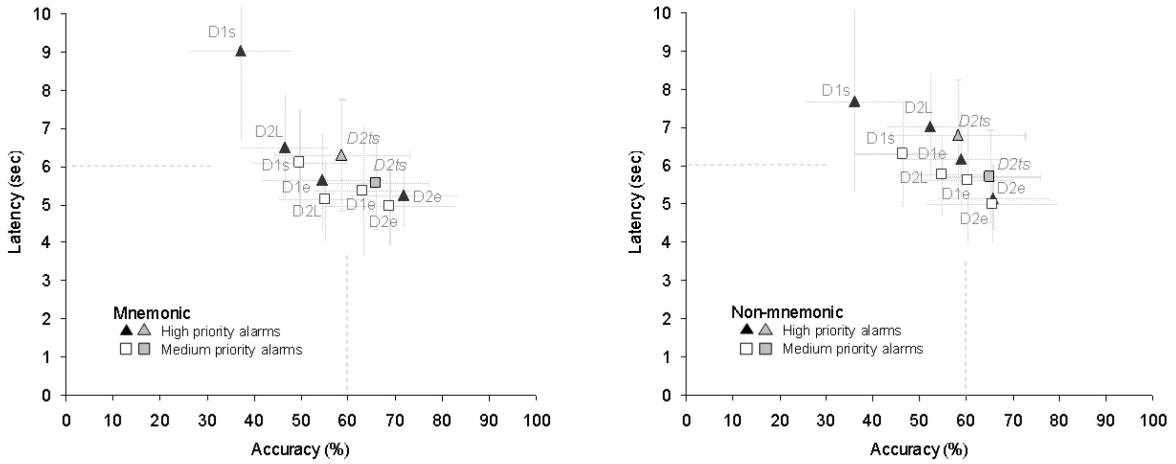


Figure 6-10 - Average accuracy (percentage correct) and latency (seconds) under M and nM learning conditions for Day 1 start (D1s), Day 1 end (D1e), Day 2 LTM (D2L), Day 2 end (D2e), and Day 2 timeshared task (D2ts). Light gray bars are 95% confidence intervals.

A combined analysis of Experiment 1 and 2 data mirrored Experiment 2 accuracy analysis over the four phases of the experiment (Figure 6-11 and Figure 6-12). No effect of mnemonic was observed however Musical participants in the combined experiment dataset were significantly better than their non-musically trained counterparts in all phases of the experiment. (Musical: Day 1 start = 58%, Day 1 end =85 %, Day 2 LTM = 76 % and Day 2 end = 84%; Non-Musical: Day 1 start = 37%, Day 1 end =57 %, Day 2 LTM = 40 % and Day 2 end =48%).

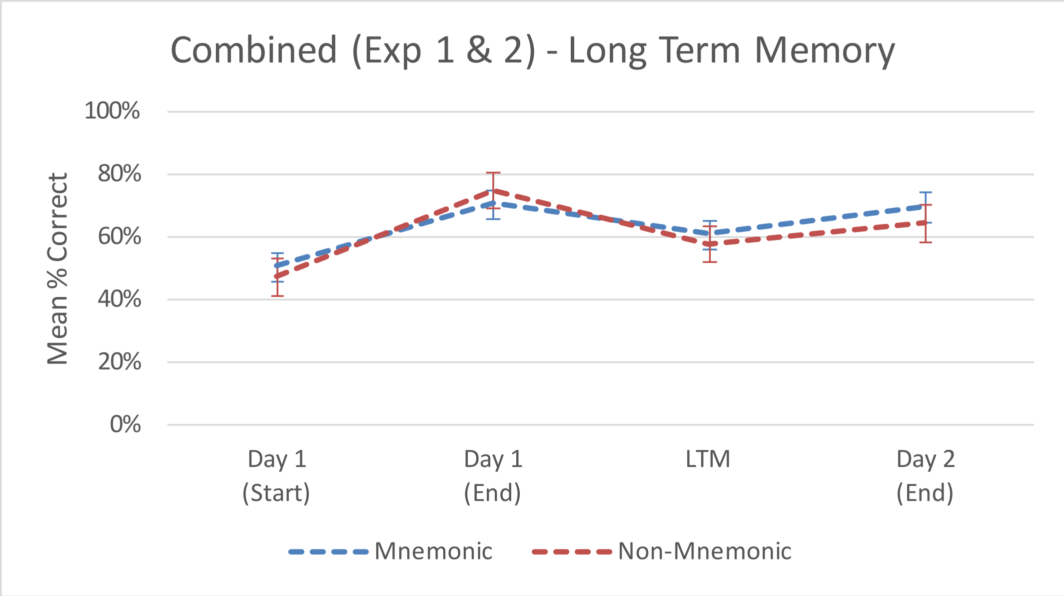


Figure 6-11: Experiment 1 and 2 Results for learning trend at the beginning and end of two days of testing by learning condition

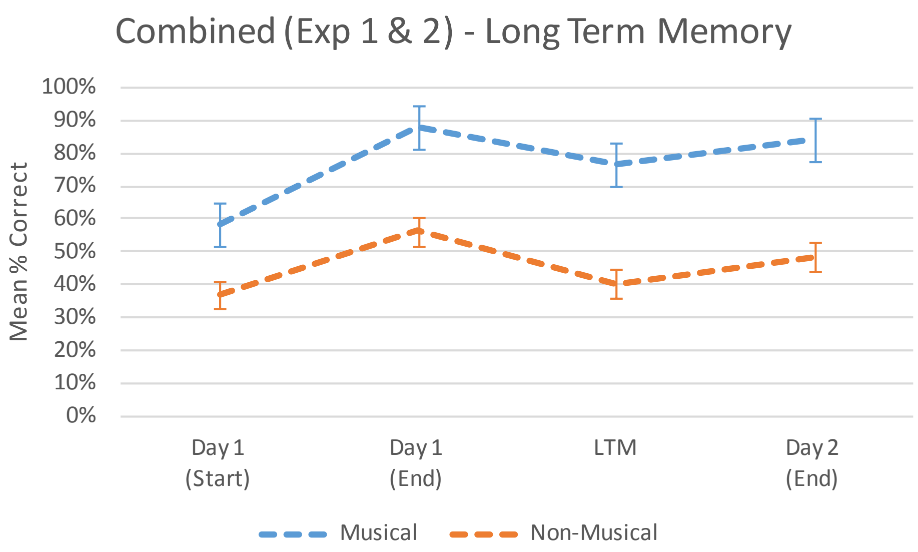


Figure 6-12: Experiment 1 and 2 Results for learning trend at the beginning and end of two days of testing by musical training.

6.4.3 Ability to identify the alarms in a timely manner

There was no observable effect of learning condition on response speed in Day 1 (Table 6-3). Participants in the Mnemonic group achieved a mean response time of 6.36s and participants in the Non-Mnemonic group achieved a mean response time of 6.45s.

Experiment 2 –Day 1 (Response Time) - ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	4.052	1	4.052	0.878	0.361	no
Mnemonics	0.006	1	0.006	0.001	0.971	no
Inter	0.216	1	0.216	0.046	0.831	no
Within	83.117	18	4.618			
Total	87.431	21	4.163			

Mean

	Mnemonic	Non-mnemonic	Mean
Musical	5.927	5.754	5.841
Non-Musical	6.612	6.853	6.732
Mean	6.270	6.304	6.287

Variance

	Mnemonic	Non-mnemonic	Mean
Musical	5.433	2.563	3.436
Non-Musical	6.487	3.367	4.563
Mean	5.642	3.096	4.163

Table 6-3: Experiment 2-Day 1 (Response Times): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

Experiment 2 –Day 2 (Response Time) - ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	11.062	1	11.062	4.471	0.048	yes
Mnemonics	0.186	1	0.186	0.076	0.787	no
Inter	1.643	1	1.643	0.664	0.426	no
Within	44.533	18	2.474			
Total	57.247	21	2.726			

Mean

	Mnemonic	Non-mnemonic	Mean
Musical	4.008	4.767	4.387
Non-Musical	6.050	5.673	5.861
Mean	5.029	5.220	5.124

Variance

	Mnemonic	Non-mnemonic	Mean
Musical	0.462	1.012	0.797
Non-Musical	5.215	1.469	3.123
Mean	4.329	1.394	2.726

Table 6-4: Experiment 2-Day 2 (Response Times): Two factor ANOVAs measuring the effect of learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

Analysis of the breakdown of musically trained and non-musically trained participants on the first day indicates that musical training had no observable effect on a nurse's ability to respond to the alarms Musically Trained = 5.84s; Non-musically trained 6.73s; p =0.319). A marginal effect of musical training was observed in Day 2 but no significant effect of the interaction of musicality and mnemonics was observed.

Unlike Experiment 1, average response time in Experiment 2 was longer for the medium priority alarms (average 6.97s) than for high priority alarms (average 5.84s). However this result was not statistically significant.

Data Set	Participants	Mean Time Taken by Trial	p Value	t Stat	StDev	dF
High Priority	22	5.84	8.54E-02	2.328	1.982	20
Medium Priority	22	6.97		2.328	2.165	20

Table 6-5: Experiment 2 Response time analysis of responses of high priority and Medium priority alarms during the first day of learning.

6.4.4 Ability to correctly identify the alarms while being otherwise occupied

Nurses regularly do a number of tasks while simultaneously monitoring the patient; therefore it is necessary to measure their performance with the melodic alarms in a similar context. As before, we found that there was a significant effect of the phase as accuracy dropped significantly between Day 2 testing and the Timeshared task (Day 2 testing = 66%, Timeshared 61%, p < 0.001).

There was no observable effect of learning condition. Participants in the Mnemonic group achieved a mean accuracy of 9.6 (60%) in each block and participants in the Non-mnemonic group achieved a mean accuracy of 8.96 (56%) in each block (p=0.7).

Experiment 2 Distractor - ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.857	1	0.857	26.353	6.96E-05	yes
Mnemonics	0.003	1	0.003	0.105	0.748	no
Inter	0.021	1	0.021	0.661	0.426	no
Within	0.585	18	0.032			
Total	1.475	21	0.07			

Mean

	Mnemonic	Non-mnemonic	Mean
Musical	0.820313	0.859375	0.839844
Non-Musical	0.474942	0.383929	0.429435
Mean	0.647627	0.621652	0.63464

Variance

	Mnemonic	Non-mnemonic	Mean
Musical	0.01001	0.005534	0.007098
Non-Musical	0.051522	0.038318	0.043695
Mean	0.064279	0.082191	0.070248

Table 6-6: Experiment 2 Accuracy analysis during distracter task of responses of mnemonic and non-mnemonic participants identifying alarms while performing a distracting math task.

Musically trained participants were significantly more accurate (84% accuracy) than non-musically trained participants (43% accuracy) ($p<0.001$)

Response times for alarm identification while performing the math task showed no significant differences between Learning Conditions ($p=0.7$) but was significant for Musical training ($p=0.03$) with musically trained participants responding faster (5.2s) than non musical participants (7.18s).

Experiment 2 – Distractor (Response Time) - ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	11.264	1	11.264	5.417	0.031	yes
Mnemonics	0.203	1	0.203	0.097	0.758	no
Inter	0.7575	1	0.757	0.364	0.553	no
Within	37.431	18	2.079			
Total	49.502	21	2.357			

Mean

	Mnemonic	Non-mnemonic	Mean
Musical	4.383	4.968	4.676
Non-Musical	6.256	6.07	6.163
Mean	5.320	5.519	5.419

Variance

	Mnemonic	Non-mnemonic	Mean
Musical	0.512	0.901	0.704
Non-Musical	3.835	1.695	2.562
Mean	3.348	1.596	2.357

Table 6-7: Experiment 2 Response time analysis during distracter task of responses of mnemonic and non-mnemonic participants identifying alarms while performing a distracting math task.

6.4.5 Self-reported satisfaction and confidence levels

Participants were asked to rate the relative urgency of the alarm sounds on a 7 point Likert scale. Two analyses were conducted combining ratings from the end of Day 1, the end of relearning on Day 2 and after the timeshared task on Day 2. Individual t-tests were used to compare samples across survey responses submitted by participants at the end of Day 1 and Day 2.

Ratings of the ease of distinguishing the alarms from each other raised marginally from Day1 to remain steady for both parts of Day 2 ($p=0.08$). Participants with musical training rated it easier to distinguish the alarms than non-musically trained participants ($p = 0.03$). There was no effect of learning condition ($p = 0.8$). Musical training also had an effect on how easy it was for participants to associate the alarms with their meanings ($p = 0.009$).

Analysis combining ratings from Day 1 with ratings from the timeshared task on Day 2 showed that participants rated the high priority alarms as sounding significantly more urgent than the medium priority alarms ($p < 0.001$). There was no effect of day, learning condition or musical training. However, musically trained participants rated the high priority alarms as relatively more urgent than the medium priority alarms than the non musically trained participants ($p = 0.002$).

At the end of the second day, participants were asked to rank the alarms according to how easy they thought they were to learn (Figure 6-13). There was a consensus that the General alarm was the easiest to learn with nearly all (1 exception) the participants assigning it a value of 1. Oxygen and Powerfail alarms were ranked next easiest to learn followed by Cardiovascular, Infusion, Perfusion and Temperature. The Ventilation Alarm was ranked the hardest alarm to learn.

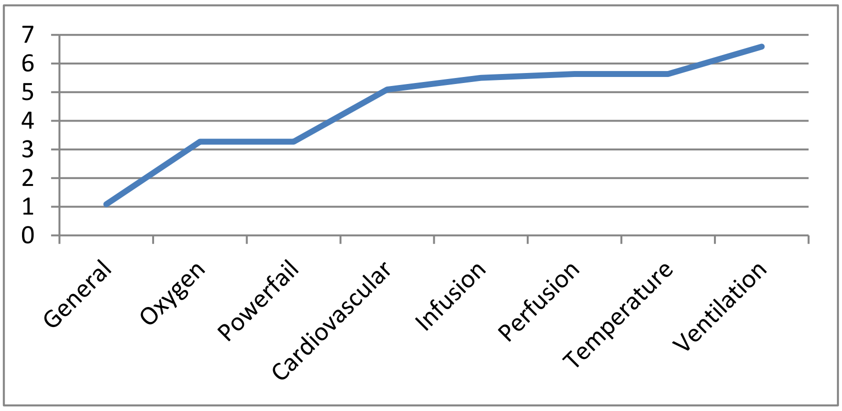


Figure 6-13: Experiment 2 Participants ranking of each alarm on the ease with which they found the alarm to learn (1-Easiest 8 hardest)

Participants were also asked to include additional comments and feedback with regards to the alarms focusing on their thoughts of use of the alarms to replace the ones which they already are familiar with at work. Their responses are presented as a word cloud with the most frequently repeated words shown as larger text. The result is displayed in Figure 6-14.

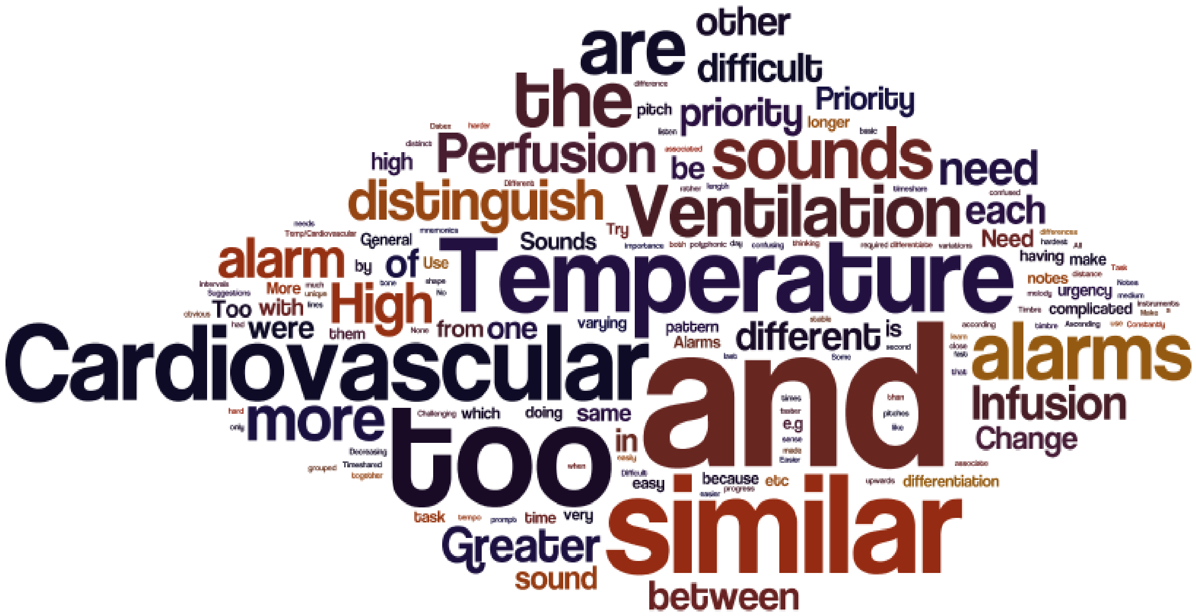


Figure 6-14: Word cloud representing participants' qualitative feedback on the alarms and their feelings about them in the work environment.

Participants in this experiment also reported significant levels of fatigue in feedback forms. This is mainly due to the experiments being conducted during or after a single shift at the hospital.

6.4.6 Analysis of Errors and Confusions

Although there is no difference in overall accuracy between Mnemonic and Non-Mnemonic conditions (see 6.4.1), participants in the Mnemonic condition collectively have a relatively narrower range of confusions (make fewer random guesses) whereas participants in the Non-Mnemonic condition collectively have a broader range of confusions. Figure 6-15 shows the pattern of confusions between alarms in the Mnemonic and Non-Mnemonic conditions for test trials in phase 2 of Day 2. Percentages on links indicated the percentage of participants who on more than 25% of test trials made the confusion indicated. Dotted arrows indicate confusions that were unique to just one participant.

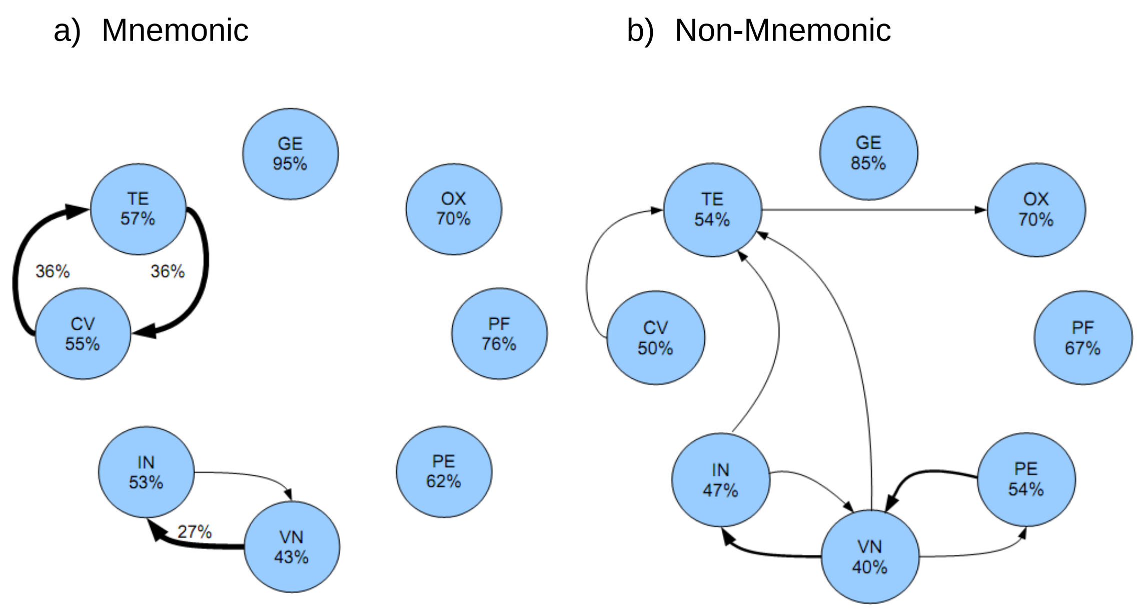


Figure 6-15 - Pattern of confusions between alarms. GE, General; OX, Oxygen; PF, power failure; PE, perfusion; VN, ventilation; IN, infusion; CV, cardiovascular; TE, temperature.

The pattern of confusions was also plotted for musical and non-musical participants. Musical participants were observed to have a narrower range of repeated confusions as compared to non-musically trained participants. Accurate identification of each alarm was also better in musical participants than non-musical participants as demonstrated by the percentage correct values in each alarm circle in Figure 6-16.

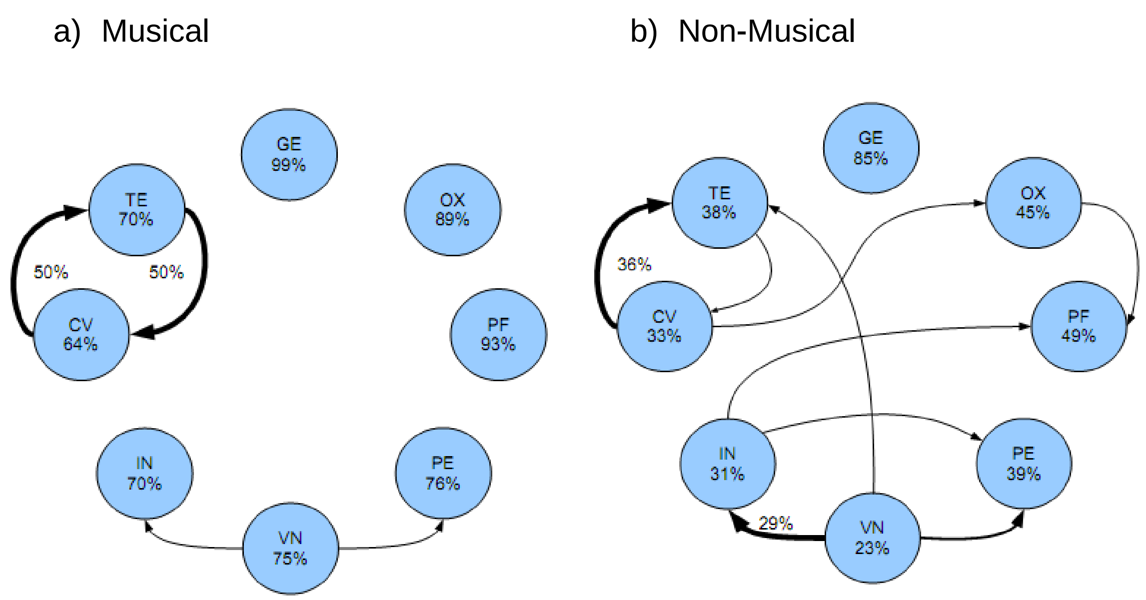


Figure 6-16: Patterns of confusions between alarms from Musical and Non Musical Participants in Experiment 2. Participants misidentified the sound at the origin of the arrow as the alarm at the end of the arrow on more than 25% of the trials.

6.4.7 Analysis of effect of related industry knowledge

In initial comparisons of the data from Experiment 1 and Experiment 2, we found that the nurses performed worse than the non-medically trained students. An analysis comparing accuracy across the Experiment 1 students and Experiment 2 nurses showed that the nurses' accuracy at identifying the melodic alarms (49%) is lower than that of the students (68%) ($p < 0.001$).

ANOVA with Regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	1.0568160	1	1.0568160	43.9639402	1.92121E-08	yes
Mnemonics	0.0236098	1	0.0236098	0.9821754	0.3262516	no
Experiments	0.1779112	1	0.1779112	7.4011732	0.0088434	yes
A x B	0.0045644	1	0.0045644	0.1898804	0.6648198	no
A x C	0.0032249	1	0.0032249	0.1341586	0.7156447	no
B x C	0.0010941	1	0.0010941	0.0455153	0.8318935	no
A x B x C	0.0066401	1	0.0066401	0.2762323	0.6014156	no
Within	1.2499888	52	0.0240382			
Total	2.9429349	59	0.0498803			

Table 6-8: Experiment 1 and 2 Accuracy analysis of responses of nurses and students identifying alarms.

However, the effect of musicality has shown so strongly in both Experiment 1 and Experiment 2 that it is necessary to take this variable into account. The number of musical participants to non-musical participants in Experiment 1 was 21 (64%):12 (36%). In Experiment 2 the number of musical participants to non-musical participants was 8 (36%):14 (64%). Due to the fact that, musically trained participants performed, on average, significantly better than non-musically trained participants, the higher proportion of musical participants in Experiment 1 should have a gross aggregate effect on the overall performance of the group, irrespective of any effect of medical knowledge.

In order to see if prior medical knowledge had any effect on the ability to learn the alarms a comparison was made between the musically trained nurses in Experiment 2 with the musically trained students in Experiment 1, followed by a comparison between the non-musically trained nurses and the non-musically trained students.

There was no observable effect of population on accuracy in the musically trained participants (Students = 78%; Nurses = 68%; $p = 0.07$). A JZS Bayes factor analysis as described in Rouder, Speckman, Sun, Morey and Iverson (2009) of the null hypothesis (that there is no accuracy difference between musically trained students and nurses) indicates that the results favour the null hypothesis by a factor greater than 2.64 to 1.

A slight effect of population was observed in the non-musically trained participants (Students = 51%; Nurses = 38%; $p = 0.044$). However, this is largely irrelevant as non-musically trained participants may have varying degrees of musical talent, pitch perception and informal musical background, thus the high degree of variability within the non-musical participants.

The only deduction that can be made from this finding is that the only way to consistently achieve high performance in alarm identification is to ensure that medical personnel have formal musical training.

6.4.8 Analysis of the effect of musical training on learning patterns

As mentioned in Experiment 1 and Section □the role in musicality cannot be ignored. Confusions in Day 2 alarms were compared between participants who were musically trained and those who were not (Musical = 81%, Non Musical = 49%, $p < 0.001$). The Confusion diagrams in Figure 6-17 present the average percentage accuracy across Day 2 Part 2 for each alarm for musical and non musical participants. Confusions are only displayed if confusions were made by the participant at least 25% of the trials (i.e. repeated confusions). Only confusions made by more than one participant are presented in the diagrams.

In an analysis of the confused alarms, we still see clear indications that the Cardiovascular alarm and the Temperature alarms are mutually confused by musical and non-musically trained participants alike. The Ventilation alarm was also frequently confused for Infusion and Perfusion alarms by participants in both conditions. Musical participants, as expected, showed a narrower range of confusions.

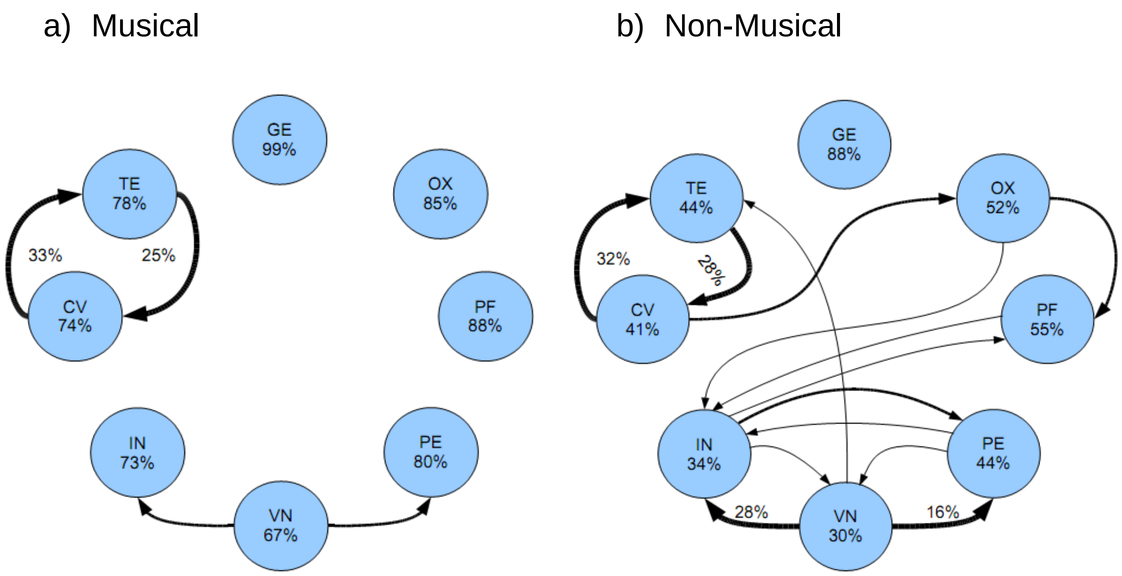
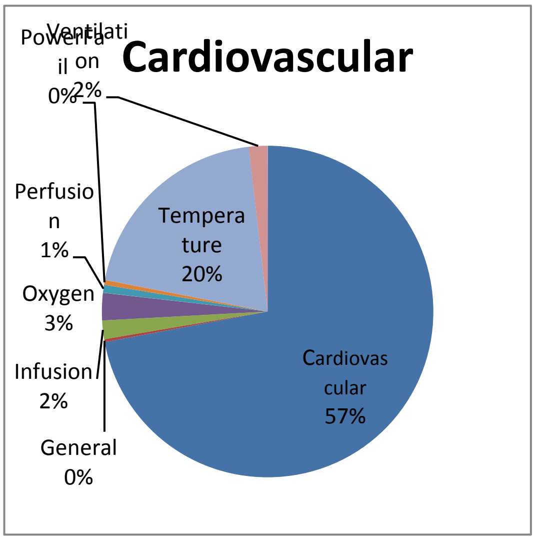
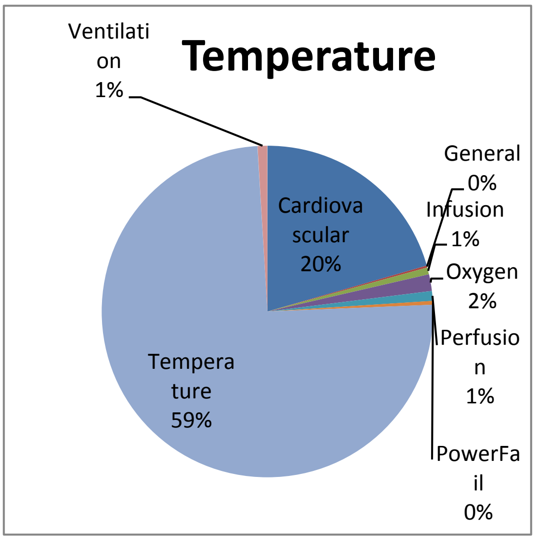
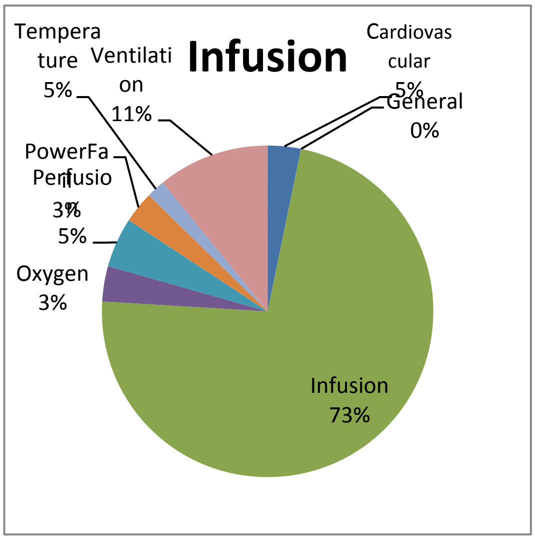


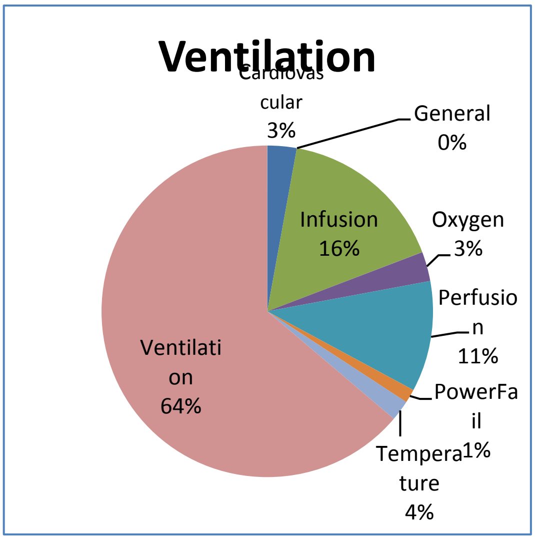
Figure 6-17: Patterns of confusions between alarms from Musical and Non Musical Participants in Experiment 1 and 2. Participants misidentified the sound at the origin of the arrow as the alarm at the end of the arrow on more than 25% of the trials.

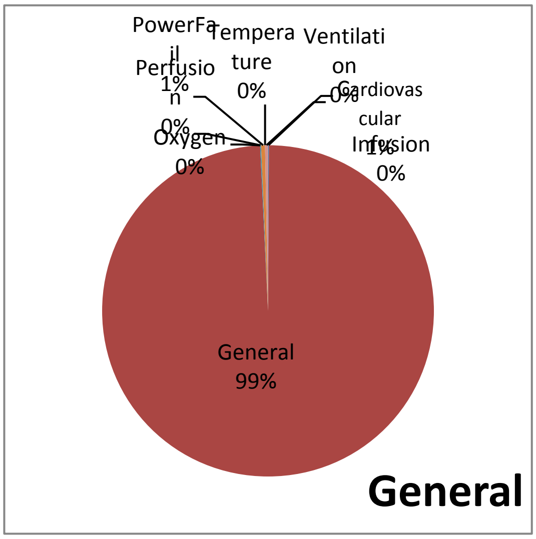
The large variation in performance in non-musical participants makes it difficult to observe performance patterns and to make recommendations for improvements. However, musically trained participants are more uniformed in their responses regardless of whether they are from the general population or medically trained. Their responses to presented stimuli provide a clearer indication of areas of greatest confusion which should be further addressed. Figure 6-18 presents the abovementioned confusions for musically trained participants from Experiment 1 and 2 in greater detail, displaying the actual breakdown of responses to each of the presented stimuli. This illustrates the causes of the errors participants make when responding to each presented stimuli.

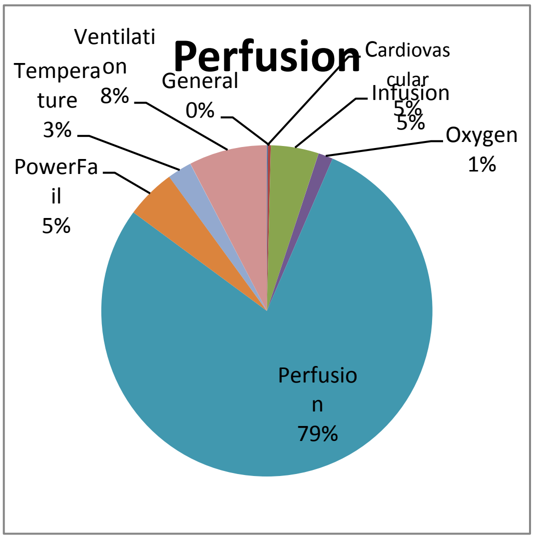

Cardiovascular
Ventilati on 2%
PowerFail 0%
Perfusion 1%
Oxygen 3%
Infusion 2%
General 0%
Temperature 20%
Cardiovascular 57%


Temperature
Ventilation 1%
General 0%
Infusion 1%
Oxygen 2%
Perfusion 1%
PowerFail 0%
Cardiovascular 20%
Temperature 59%


Infusion
Temperature 5%
Ventilation 11%
Cardiovascular 5%
General 0%
PowerFail 3%
Perfusion 5%
Oxygen 3%
Infusion 73%


Ventilation
Cardiovascular 3%
General 0%
Infusion 16%
Oxygen 3%
Perfusion 11%
PowerFail 1%
Temperature 4%
Ventilation 64%


PowerFail 1%
Temperature 0%
Ventilation 0%
Perfusion 0%
Cardiovascular 1%
Oxygen 0%
Infusion 0%
General 99%
General


Perfusion
Ventilation 8%
Temperature 3%
General 0%
Cardiovascular 5%
Infusion 5%
Oxygen 1%
PowerFail 5%
Perfusion 79%

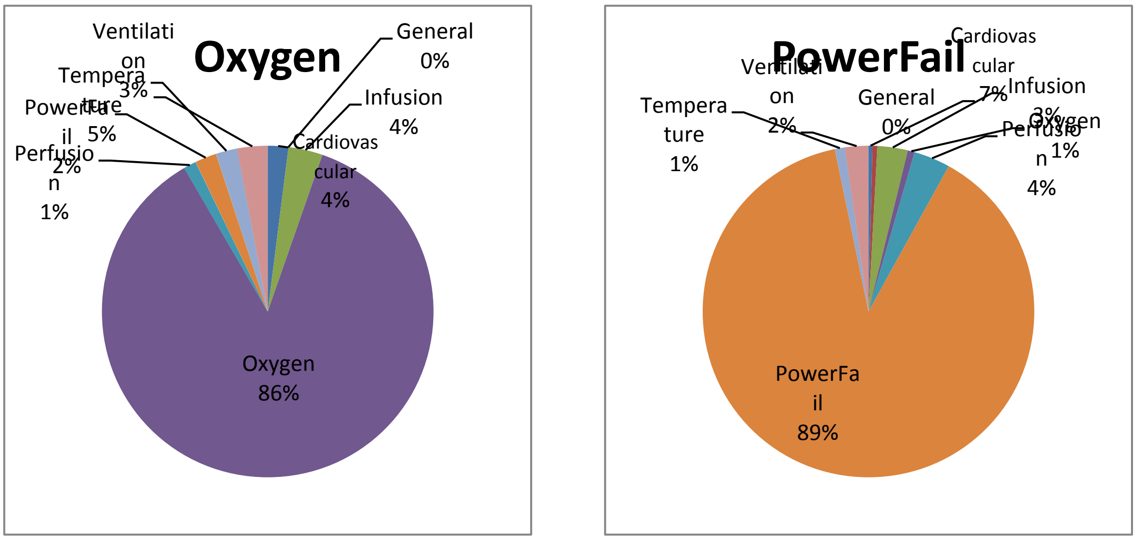


Figure 6-18: Pie graphs of presented stimuli and proportion of each of the avaliable responses for all musically trained particiants in Experiment 1 and 2.

6.4.9 Learning phase variability between Experiment 1 and Experiment 2

A concern was raised around a change in how the experiments were performed from Experiment 1 and Experiment 2. Specifically, the change had to do with the learning phase of each experiment. In the first experiment participants were only permitted to hear each alarm once between tests. This limited (and created uniformity between) the exposure each person had to the alarms. In the second experiment, this limit was removed, allowing each person to determine for him/herself how long they would spend learning the alarms. More importantly, it was now permissible for an individual to listen to a specific alarm (or confusion pair of alarms) as many times and he/she felt in necessary before proceeding with the testing phase. As a result a question was raised*: Is it still possible to compare the two experiments in a meaningful way*.

To begin to answer this question, a few hypotheses must be raised and tested:

H1. When participants have more time to learn the alarms, they will use it;

H2. When participants use more time to learn the alarms, they will perform better at the tests; and

H3. If participants who use more time do better on the tests, then participants in experiment 2 will do better than those in experiment 1.

The following ANOVAs were set up to test these assertions.

6.4.9.1 Testing H1: Did participants spend more time in the learning phase?

To answer this, mean time spent learning was broken up by participant across three factors (Experiment, Learning Condition & Musical Training)

Experiment	Learning Condition	Musicality	Participant 1	Participant 2	Etc..
1	Mnemonic	Musical			
1	Non-Mnemonic	Musical			
1	Mnemonic	Non-Musical			
1	Non-Mnemonic	Non-Musical			
2	Mnemonic	Musical			
2	Non-Mnemonic	Musical			
2	Mnemonic	Non-Musical			
2	Non-Mnemonic	Non-Musical			

This would then determine how participants both within and between factors had used the learning phase provided.

ANOVA With Regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Experiments	37794.83	1	37794.83	8.136797	0.00621	yes
Mnemonics	366.61	1	366.61	0.078927	0.779872	no
Musicality	42168.6	1	42168.6	9.07842	0.00399	yes
A x B	5723.765	1	5723.765	1.232261	0.272075	no
A x C	4565.172	1	4565.172	0.982829	0.326092	no
B x C	386.1659	1	386.1659	0.083137	0.774236	no
A x B x C	256.6256	1	256.6256	0.055249	0.815093	no
Within	241536.2	52	4644.928			
Total	355503.1	59	6025.476			

Table 6-9: Experiment 1 & 2: Three factor ANOVAs measuring the effect of Experiment (E1 vs. E2) learning condition (M vs. MN) and musical training (musically trained vs. non-musically trained)

People in experiment 2 spent 74% longer (124 seconds) in the learning phase than those in experiment 1 (73 seconds). In addition, participants without any formal musical training spent 80% longer (128 seconds) in the learning phases than those with musical training (71 seconds). No significance was found in the between factor analysis.

6.4.9.2 Testing H2: Did time spent in the learning phase effect scores?

In order to address this question, learning times were split into quartiles, with the first quartile being the shortest 25% of learning phases, and the fourth quartile being the longest. Each learning phase was then matched with its subsequent testing phase and the results were weighed against the previous test. The assumption being, the longer someone spends in learning, the greater the effect on the subsequent test's scores. Two tests were then performed: one with musical training as a factor, and another with mnemonics as a factor.

Time in learning was observed to affect accuracy scores, with those in the fourth quartile making (on average) 7% more correct responses on the test immediately following, and those in the first quartile getting 0.1% more. No significant difference was found in the rates of learning (see Section 5.5.1 and 6.4.1) between experiments, musicality or mnemonics (as previously demonstrated by the linear regressions of each).

ANOVA				Alpha	0.05	
	SS	*df*	*MS*	*F*	*p-value*	*Sig*
Experiments	0.029846	1	0.029846	1.35095185	0.245801	No
Musicality	0.073971	1	0.073971	3.34826804	0.068016	No
Time in Learning	0.344802	3	0.114934	5.20240556	0.001555	Yes
A x B	0.00444	1	0.00444	0.20096927	0.654181	No
A x C	0.032668	3	0.010889	0.49289185	0.483046	No
B x C	0.052678	3	0.017559	0.79481467	0.497302	No
A x B x C	0.007138	3	0.002379	0.10769189	0.955579	No
Within	8.88117	402	0.022092			
Total	9.605011	417	0.023034			

Table 6-10: Experiment 1 & 2: Three factor ANOVAs measuring the effect of Experiment (E1 vs. E2) musical training (musically trained vs. non-musically trained) and time spent in learning (Quartiles 1-4)

ANOVA				Alpha	0.05	
	SS	*df*	*MS*	*F*	*p-value*	*sig*
Experiments	0.047376	1	0.047376	2.129829	0.145238	no
Mnemonics	0.003351	1	0.003351	0.150642	0.698128	no
Time in Learning	0.410434	3	0.136811	6.150527	0.000427	yes
A x B	0.017803	1	0.017803	0.800347	0.371525	no
A x C	0.036023	3	0.012008	0.539822	0.462935	no
B x C	0.062466	3	0.020822	0.936078	0.42318	no
A x B x C	0.067665	3	0.022555	1.013991	0.386369	no
Within	8.942034	402	0.022244			
Total	9.605011	417	0.023034			

Table 6-11: Experiment 1 & 2: Three factor ANOVAs measuring the effect of Experiment (E1 vs. E2), learning condition (M vs. MN) and time spent in learning (Quartiles 1-4)

The above tests confirm the following:

- Given the opportunity, people will spend more time learning the alarms (H1 confirmed);
- Spending more time learning the alarms will improve a participant's score (H2 confirmed);
- Participants with no formal musical training spent more time learning the alarms than those with no formal musical training in both experiments;

Given the confirmation of H1 and H2 it is expected that participants who are allowed to take more time will perform better, however, it was not observed that participants in Experiment 2 did better than those in Experiment 1 (H3 not confirmed). On the whole, participants in Experiment 2 did *worse* than those in Experiment 1 (see section 6.5.7). Much of this can be accounted for by the percentage of musically trained participants (Experiment 1 = 63.6%; Experiment 2 = 36.4%). In fact, when controlling for all factors (Table 6-12), there is no statistically significant difference between the student group and the nurses group. The descriptive statistics for this analysis provides further detail in support of the above statements and has been included in Appendix F.5.

Experiment 1 v 2 - ANOVA with Regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	1.0568160	1	1.0568160	43.9639402	1.92121E-08	yes
Mnemonics	0.0236098	1	0.0236098	0.9821754	0.3262516	no
Experiments	0.1779112	1	0.1779112	7.4011732	0.0088434	yes
A x B	0.0045644	1	0.0045644	0.1898804	0.6648198	no
A x C	0.0032249	1	0.0032249	0.1341586	0.7156447	no
B x C	0.0010941	1	0.0010941	0.0455153	0.8318935	no
A x B x C	0.0066401	1	0.0066401	0.2762323	0.6014156	no
Within	1.2499888	52	0.0240382			
Total	2.9429349	59	0.0498803			

Table 6-12: Experiment 1 & 2: Three factor ANOVAs measuring musical training (musically trained vs. non-musically trained), learning condition (M vs. NM), and Experiment (Experiment 1 and Experiment 2)

6.5 Conclusion

This is the first study to test the IEC 60601-1-8 alarms in a way that compares the effects of the mnemonics supplied in the standard and the effects of having a timeshared task that uses a sample of nurses who work in critical care areas and are therefore familiar with patient monitoring.

At the end of the two learning sessions only one nurse (4%) could meet the criteria of 100% accuracy in two successive tests. The mnemonics once again did not seem to have any effect on the accuracy and speed of participant's responses. This was surprising because the mnemonics were designed to be useful to healthcare workers and clinicians, and was meant to make use of the knowledge that these experienced medical professionals would find second nature in order to aid in alarm identification. A possible explanation for this would be that memory aids have been shown to be most effective when people are allowed to generate their own associations (Edworthy & Hellier, 2005) rather than having associations imposed upon them as it is in the IEC 60601-1-8 standard. However, over 80% of participants

The confusions data seems to indicate that certain pairs and groups of alarms are mutually confusable, no matter which condition the participant is in. This is especially so in the early stages of the experiment with some confusions persisting all the way to the end.

This finding is supported by an analysis of the participants learning patterns (i.e. how many times a participant chose to play an alarm during the learning phase). The mean number of observations was very similar across all groups with Non-musical participants, on average, only listening to alarms 10% more during the learning phase than musical participants. (Musical participants = 483.5, Non musical = 531.43).

Participants with at least 1 year of musical training were generally observed to choose to listen to only one of a confusions set, leading to higher variance when there were more alarms in the confusion group (Figure 6-19). This pattern was not observed in the non-musical group which showed little in the way of patterned learning (with the exception of fewer observations of participants in both groups clicking on the General Alarm).

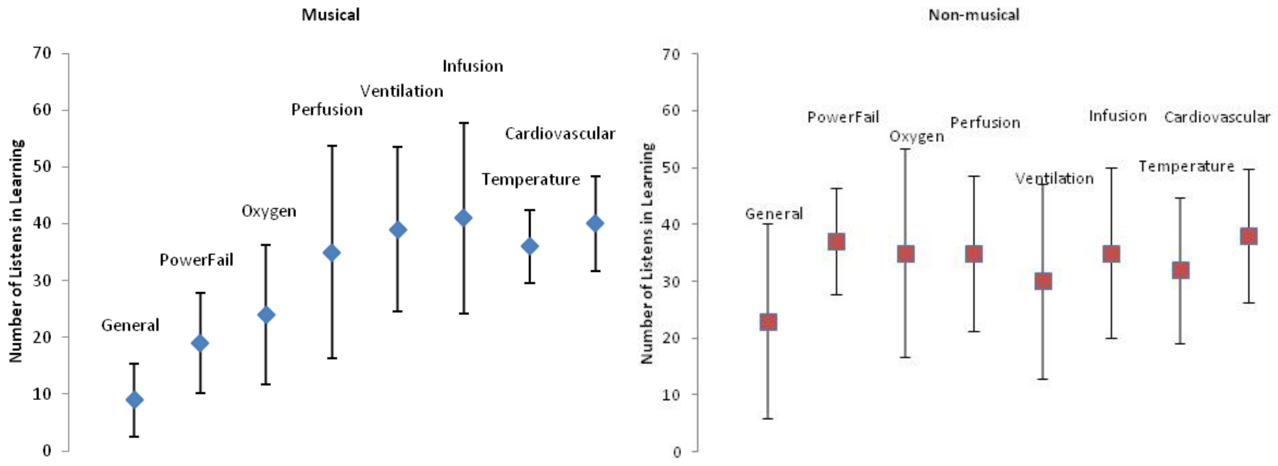


Figure 6-19: Experiment 2-Averaged count of the number of times participants chose to play the alarm during the learning phase with standard deviations.

As with the findings in Experiment 1, during the testing phase, musical training was associated with a greater ability to differentiate between melodies with pitch variations. Musically trained participants were 35% more accurate at identifying alarms correctly. Although we would expect from previous music studies that musical training would be beneficial, the concern arises as to just how dependent on musical ability is the ability to identify the alarms.

In order to address the concern expressed in responses to the publication of the findings of Experiment 1, the data from Experiment 1 was analysed as a comparison to Experiment 2. The effect of population was not observed in musically trained participants. Any observed differences in performance and time spent in learning between experiments is negligible when factors of musicality and learning condition are taken into account. This analysis indicates that it is not unreasonable to assume that there is little difference in the performance of the nurses and the students.

6.6 Discussion

The confusion diagrams in Experiment 1 and 2 suggest that the presence of “persistent confusions” appear to prevent participants from getting other alarms in that cluster correct. This effect was concentrated around alarms only in that cluster. Other alarms were not affected by the presence or absence of that alarm. Therefore, when a participant is presented with the Cardiovascular stimuli there is a higher probability that they will respond with the incorrect answer; as opposed to when presented with the General stimuli where the information transfer is clear and the correct response is highly probable.

An obvious reason for this is that there are essentially only four ways to create variations within the IEC 60601-1-8 melodic alarms, due to the restrictions in rhythm, timbre, and some limitations in pitch. These variations are as follows:

- Level pitch (e.g. General Alarm);
- Ascending pitch (e.g. Cardiovascular alarm and Temperature alarm);
- Descending pitch (e.g. PowerFail alarm and Oxygen alarm); and
- Varying pitch (Ventilation alarm, Perfusion alarm and Infusion alarm).

These groups are identical to the persistent confusion clusters that are observed. The General alarm, being the only alarm of a level pitch, is very rarely confused for any other alarms.

6.6.1 Descending Pitch Cluster

Outside of the General alarm, the next best performing alarms are those of the descending pitch (Powerfail and Oxygen). Although some mutual confusion was observed in Experiment 2, the accuracy in identifying these alarms was significantly better than any others. An explanation for this may be found in the melodic shape of the alarms themselves. The Powerfail (high priority) alarm is a sudden drop of an entire octave (C5-C4-C4—C5-C4) whereas the Oxygen (high priority) alarm is a steady drop of a single note pitch (C5-B4-A4—G4-F4). This further distinguishes the Powerfail alarm from the Oxygen alarm even though they do share the same property of descending pitch.

6.6.2 Ascending Pitch Cluster

The Temperature and Cardiovascular melodic alarms in IEC 60601-1-8 do not have this distinguishing property. Both alarms ascend at a uniform pitch, the Temperature (high priority) alarm in single note steps (C4-D4-Experiment 4—F4-G4) and the Cardiovascular (high priority) alarm in thirds (C4-Experiment 4-G4—G4-C5). Although the interval is different, we found that participants (especially those who were not musically trained) had great difficulty in identifying this distinguishing characteristic and correctly identifying the Cardiovascular alarm from the Temperature alarm.

6.6.3 Varying Pitch Cluster

The varying pitch cluster is the group of alarms which have their musical shape described as variable (the pitch goes up and down). This group is characterised by two distinct features which makes the alarms difficult to distinguish from each other: similarities in pitch variability in the first three notes, and the last two notes.

The Ventilation (high priority) alarm (C4-A4-F4—A4-F4) and the Perfusion (high priority) alarm (C4-F#4-C4—C4-F#4) both share the property of pitch changing in a “Down-Up-Down” pattern in the first 3 notes, while the Infusion (high priority) alarm (C5-D4-G4—C5-D4) is a “Up-Down-Up” pattern in the first 3 notes. Likewise, the last 2 notes of the Ventilation and Infusion alarms are similar in that they both descend whilst the Perfusion alarm ascends.

As the alarm is played twice in under 3 seconds, participants have commented that they are incapable of distinguishing anything more than the fact that the “sound moves up and down”. Of the three alarms in this cluster, accuracy in identifying the Perfusion alarm is consistently better than the other two due to the distinguishing artificial tri-tone created by the F#. However, it is clear from the confusion patterns that these three alarms are very frequently confused with each other.

7 EXPERIMENT 3 – IMPROVING ACCURACY BY REMOVING ONE STIMULI FROM A “PERSISTENT CONFUSION” PAIR

Perhaps the most prominent observation arising from the previous two experiments was the existence of persistent confusion groups. Specifically, these groups were the Cardiovascular and Temperature alarms (Group 1) along with the Infusion, Perfusion and Ventilation alarms (Group 2). These confusions were made by almost all participants in the previous two experiments.

The confusion between Cardiovascular and Temperature alarms were the most prevalent with both occurring in 18.5% of all trials (Cardiovascular for Temperature in 179 from 974 trials; Temperature for Cardiovascular in 180 from 974 trials) (see Figure 5-11 in Chapter 5 and Figure 6-15 in Chapter 6).

Based on this, it was theorised that the removal of one of these alarms would result in the other(s) in its melodic contour group being correctly identified more often, and at a rate greater than chance.

7.1 Feasibility Study

Preliminary analysis was carried out to test if the reduction in one alarm would provide significant improvement in its confusion partner. To do this, Day 1 accuracy from Experiment 2 was manipulated to remove all responses by participants to the Cardiovascular alarm and all responses of the Cardiovascular alarm by participants to any other alarm. (Table 7-1 provides an example of how the data manipulation was carried out.)

Removing the Cardiovascular as a stimulius and as a response resulted in the number of times the alarm was played to be unequal across all the stimuli. This was overcome by normalising the data, presenting the number of correct responses to the alarm as a percentage of the total number of times that Alarm was selected as the correct answer. A similar approach was adopted by Williams and Beatty (2005) when a programming glitch caused the stimuli to present an unequal number of times (email correspondence from Beatty, 2006).

	Alarm Played	Participant Response									% Correct
		CV	GE	IN	OX	PE	PF	TE	VE		
M12	CV	8				1		5		14	57.14
	GE		14							14	100.00
	IN			9	1	2			2	14	64.29
	OX			4	8				2	14	57.14
	PE	1			2	10		1		14	71.43
	PF		1			1	11		1	14	78.57
	TE	6	1					7		14	50.00
	VE	2		1	3	3			5	14	35.71
M12 Total		17	16	14	14	17	11	13	10	112	

	Alarm Played	Participant Response								% Correct
		GE	IN	OX	PE	PF	TE	VE		
M12	GE	14							14	100.00
	IN		9	1	2			2	14	64.29
	OX		4	8				2	14	57.14
	PE			2	10		1		13	76.92
	PF	1			1	11		1	14	78.57
	TE	1					7		8	87.50
	VE		1	3	3			5	12	41.67
M12 Total		16	14	14	17	11	13	10	89	

Table 7-1: Participant data manipulated to simulate responses if CV alarm was absent. The Table on the left is the original data and the table on the right is after CV has been removed.

Accuracy for each alarm with-Cardiovascular and without-Cardiovascular were compared across each alarm for each participant. The purpose of this was to see the effect of the missing Cardiovascular alarm on all Stimuli and not just its confusion pair (Temperature).

Musical Participants			
Stimuli (Alarm)	Experiment 1+Experiment 2 Average (with Cardiovascular)	Experiment 1+Experiment 2 Average (without Cardiovascular)	P
General	1.00	0.99	0.215
Infusion	0.68	0.76	0.056
Oxygen	0.81	0.88	0.112
Perfusion	0.75	0.80	0.264
PowerFail	0.84	0.89	0.079
Temperature	0.73	0.95	0.000
Ventilation	0.62	0.68	0.433
Non-Musical Participants			
Stimuli (Alarm)	Experiment 1+Experiment 2 Average (with Cardiovascular)	Experiment 1+Experiment 2 Average (without Cardiovascular)	P
General	0.87	0.88	0.849
Infusion	0.30	0.36	0.214
Oxygen	0.46	0.54	0.341
Perfusion	0.40	0.45	0.477
PowerFail	0.48	0.57	0.235

Temperature	0.40	0.54	0.049
Ventilation	0.26	0.31	0.255

Table 7-2: Average accuracy (Percentage Correct) under M and MN conditions for Day 1 with-Cardiovascular alarm and without-Cardiovascular alarm.

As shown in Table 7-2, the removal of the Cardiovascular alarm has very little affect on the averages on the General, Infusion, Oxygen, Perfusion, Powerfail, and Ventilation alarms, this is supported by the non-significant P-value. However, a marginally significant value was observed for the effect on performance on the Temperature alarm (Musical $p < 0.0001$; Non-musical $p < 0.05$).

This result was suggestive, but previous research into information transfer Rouder et al. (2004) demonstrated that such findings when done in simulation are not necessarily observable when tested on human participants. It is therefore necessary to test if the Cardiovascular alarm is having a singular effect on the ability to correctly identify the Temperature alarm. A group of participants who have been exposed to a set of IEC 60601-1-8 melodic alarms without the Cardiovascular alarm should perform with similar accuracy to all of the alarms other than Temperature. Also, improvement in the participant's ability to correctly identify the Temperature alarm should be significantly greater to any chance improvement due to the alarm set being reduced by one. It was the aim of Experiment 3 to test this hypothesis on human participants.

7.2 Method

The method used was similar to that which was used in Day 1 of Experiment 1 and 2. Students were used as the participant body as there was no observable effect of occupation from the previous two experiments and they are easier to obtain. Also, given that there was no observable effect of learning condition, only the Non-Mnemonic condition was used. Participants were, however, encouraged to use whatever mnemonics they could come up with to help them in correctly identifying the alarms.

The Java program used to run Experiment 2 was again used with conditions only set for Day 1 testing in the NM condition. The Cardiovascular alarm was removed as a Stimulus and was not visible as an answer in the response panel. In total 7 alarms were presented in two priority resulting in a total of 14 alarms instead of 16. A “Background” and “End of Experiment” Questionnaire was added to the program which removed the need for any paper based questionnaires and would reduce the amount of time and reduce the risk of errors from manual entry of data by the experimenter or data collector.

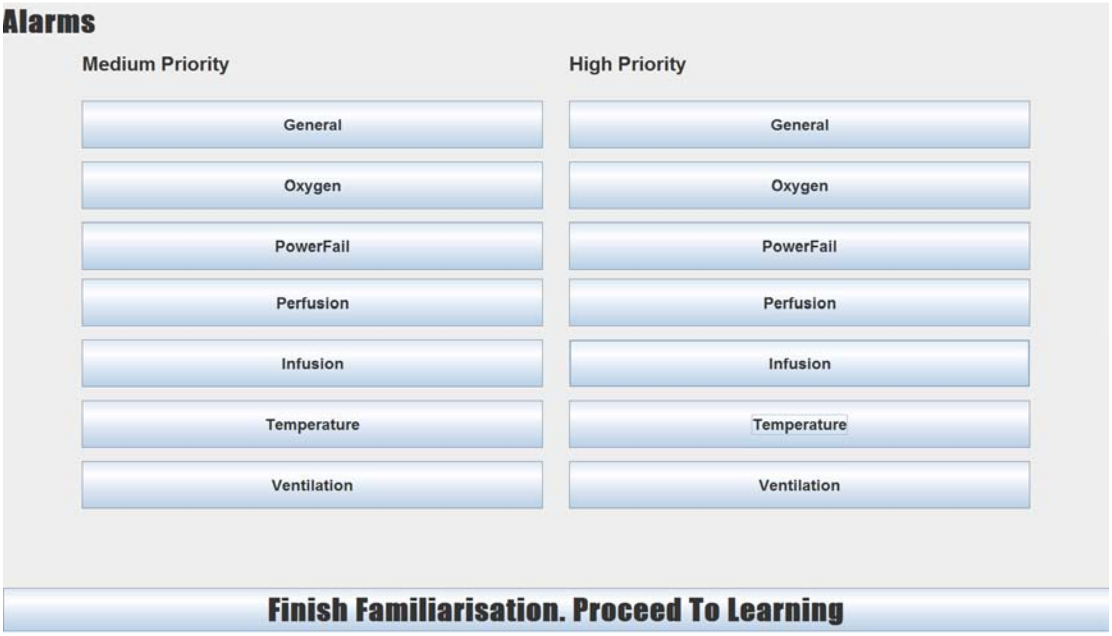


Figure 7-1: Learning Screen from Experiment 3 with the Cardiovascular Alarm option removed.

7.3 Participants

Participants were 23 students from the University of Newcastle in NSW, Australia, recruited through the School of Psychology research sign-up system. The students were between the ages of 18 and 36 years of age (Average 23; Median 20). Participants were asked if they had any formal musical training and for how many years (Musically trained = 8 [34.8%], Non-musically trained = 15 [65.2%]). No participant had previous experience with the IEC 60601-1-8 melodic alarm sounds.

The study was conducted in accordance with the ethical review processes of the University of Newcastle.

7.4 Apparatus and Stimuli

The experimental set up was identical to that described in Experiment 2 with a slightly modified experimenter's protocol which omitted the description of the Cardiovascular alarms.

7.5 Procedure

The procedure remained unchanged from that used on the first day of testing in Experiment 2. With a learning-test phase and a target criterion of achieving 2 consecutive sets of perfect test scores or approximately 45 mins of learning-testing time had passed whichever occurred first. No learn-test cycle was left incomplete.

7.6 Results

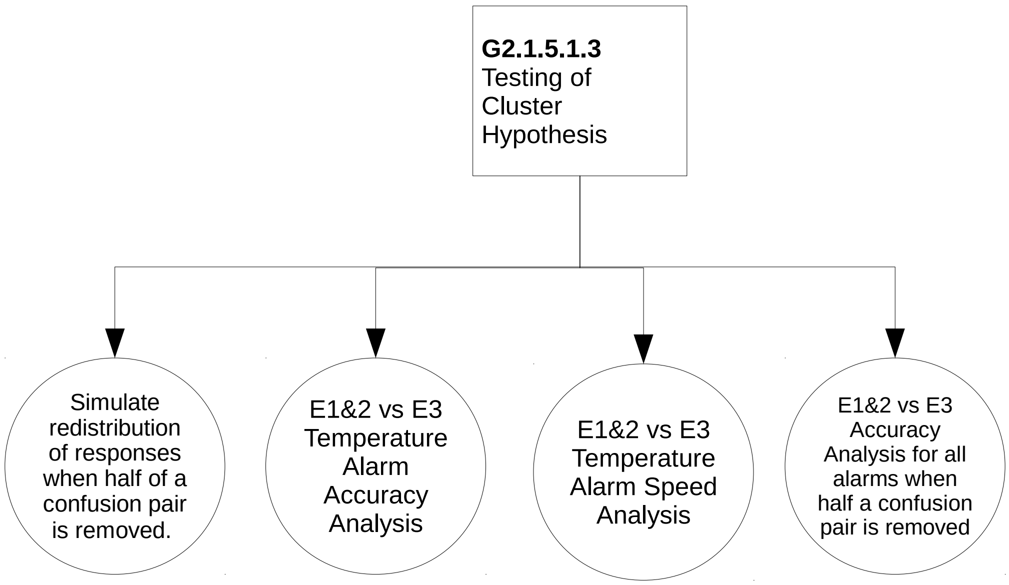


Figure 7-2: Measurement of performance for removal of one half a persistently confused pair

The primary goal of this study was to discover if the removal of one half of a persistently confused pair (Cardiovascular) would have a significant effect on a person's ability to:

- Learn the alarm meaning of the other half of the confusion pair (Temperature); and
- Identify the remaining half of the confusion pair in a timely manner.

Secondary goals of this study included an analysis of:

- If general accuracy performance would be less dependent on musical training; and
- If participants found it easier to learn the alarms with less confusion.

7.6.1 Ability to learn the alarms

At the end of the single session experiment, eight participants managed to achieve the learning criterion of 100% correct identification of the 14 alarms in two consecutive test cycles (Musical = 6; Non Musical = 2). An additional four participants managed to achieve at least one 100% correctly identified test trial.

As with Experiment 1 and Experiment 2, the participants with at least one year of musical training were more accurate in identifying the alarms than those without (Average accuracy musically-trained = 86%, non-musically trained = 62%, $p < 0.001$).

Data Set	Participants	Mean Correct by Block	p Value	t Stat	StDev	dF
Musical	8	86%	8.31E-04	2.560032959	0.095691	6
Non-Musical	15	62%		2.378786266	0.195956	13

Table 7-3: Experiment 3 Accuracy analysis of responses of musical and non-musical participants.

There was a difference in accuracy between when the participants started the experiment and when they finished the experiment. (Average First = 46%, Average Last = 78%, $p < 0.001$).

Figure 7-3 shows the learning trend between the first trial and the last trial of the experiment separated by musical training. Musical participants both started and ended the experiment with better accuracy performances but Musically trained participants were within a narrower band of final results than non-musically trained participants.

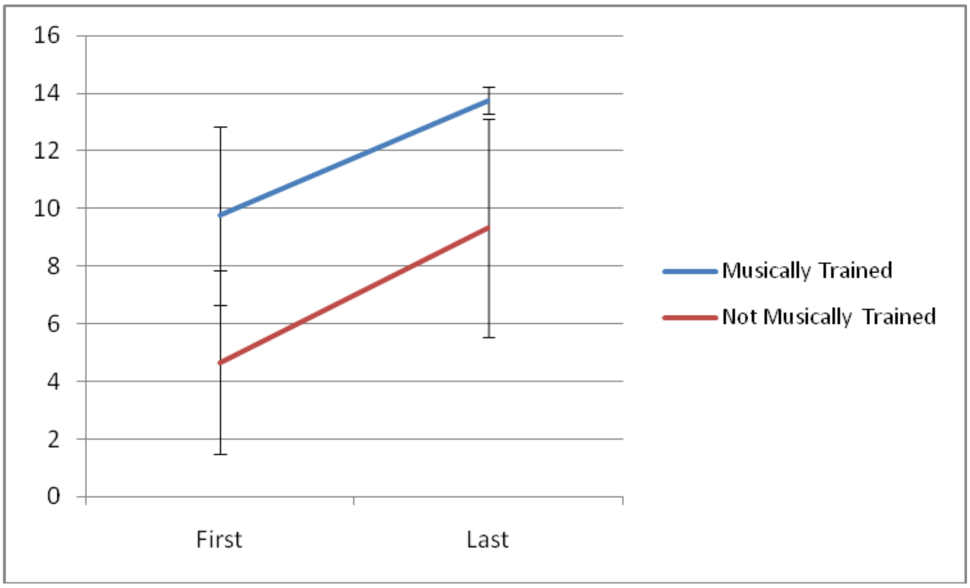


Figure 7-3: Learning trend between the first trial and the last trial of the experiment

An analysis of accurate response by participants to each alarm in Experiment 3 was compared to accuracy response by each participant to each alarm in Experiment 1 and 2. The results across all participants showed that the only significant improvement was observed in accuracy measurements of the temperature alarm.

All Participants			
Stimuli (Alarm)	Experiment 1+Experiment 2 Average (with Cardiovascular)	Experiment 3 Average (No Cardiovascular tested)	P
General	0.92	0.93	0.929
Infusion	0.50	0.55	0.443
Oxygen	0.64	0.76	0.895
Perfusion	0.58	0.62	0.573
PowerFail	0.67	0.74	0.216
Temperature	0.58	0.77	0.007
Ventilation	0.45	0.55	0.217

Table 7-4: Accuracy responses to each alarm by participants in Experiemnt 1 and Experiemtn 2 compared with those in Experiment 3 when the Cardiovascular alarm was removed

The breakdown of the accuracy results across musical and non-musical participants showed that non-musically trained participants were more distributed in their improvements across multiple alarms indicating that the better accuracy was more affected by the fact that there were fewer alarms to select from than any persistent confusion effect.

Musical Participants			
Stimuli (Alarm)	Experiment 1+Experiment 2 Average (with Cardiovascular)	Experiment 3 Average (No Cardiovascular tested)	P
General	1.00	0.99	0.286
Infusion	0.68	0.71	0.781
Oxygen	0.81	0.88	0.363
Perfusion	0.75	0.74	0.868
PowerFail	0.84	0.84	0.991
Temperature	0.73	0.96	<0.001
Ventilation	0.62	0.70	0.516
Non-Musical Participants			
Stimuli (Alarm)	Experiment 1+Experiment 2 Average (with Cardiovascular)	Experiment 3 Average (No Cardiovascular tested)	P
General	0.87	0.91	0.500
Infusion	0.30	0.43	0.026
Oxygen	0.46	0.67	0.042
Perfusion	0.40	0.53	0.167
PowerFail	0.48	0.67	0.034
Temperature	0.41	0.68	0.004
Ventilation	0.26	0.41	0.060

Table 7-5: Accuracy responses to each alarm by participants in Experiment 1 and Experiment 2 (distinguished by musical training) compared with those in Experiment 3 when the Cardiovascular alarm was removed

Significant differences were observed between experiments (p=0.001) and musical and non musical training (p<0.0001).

Experiment 1&2 vs 3 | ANOVA with Regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	1.209909	1	1.209909	44.57038	3.09E-09	yes
Experiments	0.306603	1	0.306603	11.29456	0.0012	yes
Inter	0.015498	1	0.015498	0.570902	0.452148	no
Within	2.144537	79	0.027146			
Total	4.055965	82	0.049463			

Mean

	Exp 1&2	Exp 3	Mean
Musical	0.754547	0.86369	0.809119
Non-Musical	0.443196	0.615649	0.529422
Mean	0.598871	0.73967	0.669271

Variance

	Exp 1&2	Exp 3	Mean
Musical	0.021278	0.010465	0.020747
Non-Musical	0.030951	0.041141	0.040527
Mean	0.04988	0.044101	0.049463

Table 7-6: Experiment 1&2 vs 3 - (Accuracy): Two factor ANOVAs measuring the effect of musical training (musically trained vs. non-musically trained) and Experiment (Experiment 1 & 2 vs. Experiment 3)

A similar effect was observed in the analysis of the Temperature Alarm. Differences were observed between experiments ($p<0.0001$) and musicality ($p<0.0001$). No interaction between the 2 factors was observed.

Experiment 1&2 vs 3 (Accuracy of Temperature Alarm)| ANOVA with Regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	1.374	1	1.374	25.35853	2.94E-06	yes
Experiments	0.930259	1	0.930259	17.16884	8.52E-05	yes
Inter	0.005556	1	0.005556	0.102546	0.749641	no
Within	4.280454	79	0.054183			
Total	6.801931	82	0.08295			

Mean

	Exp. 1&2	Exp. 3	Mean
Musical	0.729952	0.95625	0.843101
Non-Musical	0.412938	0.677144	0.545041
Mean	0.571445	0.816697	0.694071

Variance

	Exp. 1&2	Exp. 3	Mean
Musical	0.053216	0.00817	0.05217
Non-Musical	0.05005	0.087301	0.077509
Mean	0.076302	0.076629	0.08295

Table 7-7: Experiment 1&2 vs 3 - (Accuracy of the Temperature Alarm): Two factor ANOVAs measuring the effect of musical training (musically trained vs. non-musically trained) and Experiment (Experiment 1 & 2 vs. Experiment 3)

7.6.2 Ability to identify the Temperature alarm in a timely manner

How quickly a participant was able to respond to the temperature alarm was measured to determine if the removal of one half of the confusion pair would have any significant effect on a person's ability to quickly respond to the alarm.

Participants in Experiment 3 were significantly quicker at responding to the alarm than those in Experiment 1 and 2 (Average Experiment 1and Experiment 2 = 6.09s; Average Experiment 3 = 4.28s; $p<0.0001$). Musicality was not observed to be a statistically significant factor in participants' response times to the temperature alarm. This suggests that that the only observed impact to affect the response times to the temperature alarm was the removal of the cardiovascular alarm.

Experiment 1&2 vs 3 (Response Time)| ANOVA with Regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	10.0182	1	10.0182	3.079771	0.083149	no
Experiments	55.76832	1	55.76832	17.14417	8.61E-05	yes
Inter	4.730287	1	4.730287	1.454174	0.231459	no
Within	256.9794	79	3.252903			
Total	339.6357	82	4.141898			

<u>Mean</u>

	Exp 1&2	Exp 3	Mean
Musical	5.459745	4.113873	4.786809
Non-Musical	6.817614	4.365668	5.591641
Mean	6.138679	4.239771	5.189225

<u>Variance</u>

	Exp 1&2	Exp 3	Mean
Musical	4.010009	1.770962	3.802558
Non-Musical	4.031057	0.816701	4.261775
Mean	4.418356	1.098243	4.141898

Table 7-8: Experiment 1&2 vs 3 - (Response Time for temperature alarm): Two factor ANOVAs measuring the effect of musical training (musically trained vs. non-musically trained) and Experiment (Experiment 1 & 2 vs. Experiment 3)

7.6.3 Confusions

Participants showed no confusion between high and medium priority alarms. One difference between this and previous studies is that the percentage of correct responses to the Temperature alarm were significantly different in Experiment 3 (Day 1) testing versus

Experiment 1 and 2 (Musical: Experiment 1 and 2 = 73%; Experiment 3 = 97%; p = 0.002; Non Musical: Experiment 1 and 2 = 40%; Experiment 3 = 69%; p = 0.004).

Figure 7-4 shows the pattern of confusions between alarms for all test trials conducted in the experiment. Percentages on links represent the percentage of participants who misidentified the sound at the origin of the arrow as the alarm at the point of the arrow on more than 25% of trials. The numbers in each circle indicate the percentage of accurate identification of that alarm made by participants in all trials. Three arrow weights are shown. The thinnest indicates confusions made by at least 2 participants followed by confusions made by at least 3 participants. Confusions made by 4 or more participants are presented as the thickest arrows with the percentage of participants who misidentified the alarms presented on the links.

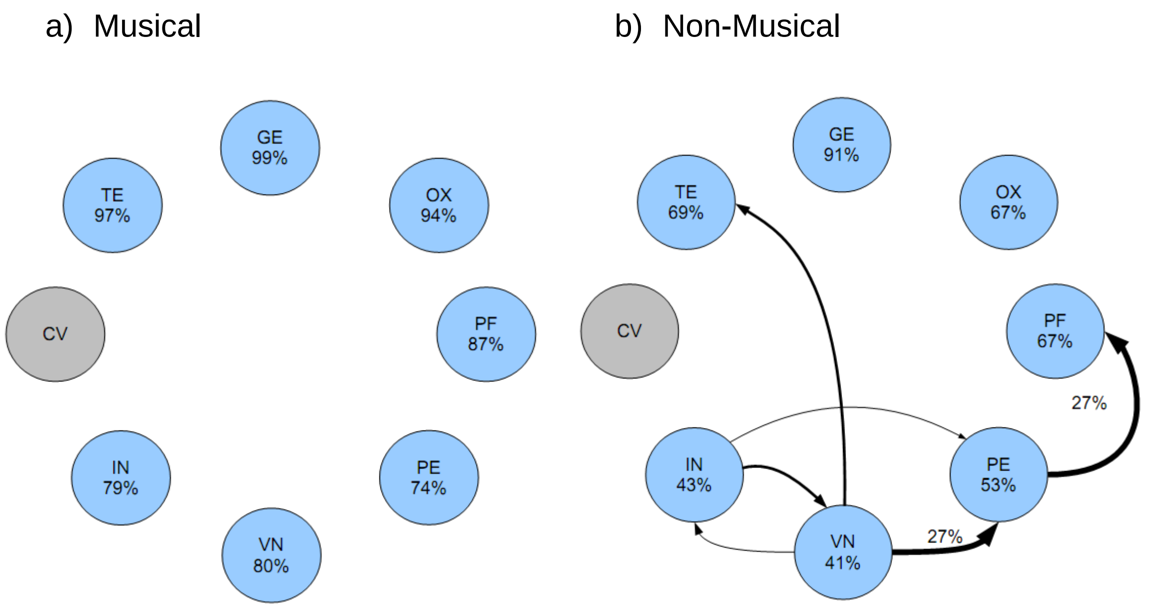


Figure 7-4: Patterns of confusions between alarms from Musical and Non Musical Participants in Experiment 3.

7.7 Conclusion

At the end of the experiment, eight participants managed to achieve the learning criterion of 100% correct identification of the 14 alarms in two consecutive test cycles (Musical = 6 from 8; Non Musical = 2 from 15). Confirming once again that musically trained participants have a significant advantage in being able to discriminate and remember the alarms.

The second “persistent confusion” cluster of Infusion, Ventilation and Perfusion was still evident and were not significantly affected by the removal of the Cardiovascular alarm. The similar musical contours of this cluster (alternate raising and lowering of pitch) still cause confusion between these three alarms by participants.

The analysis of the participants ability to learn the alarms showed that there was no observable difference between participants’ performance in identifying the alarms. It is a significant finding that the effect of Experiment was observed in an analysis of the performance in the Temperature alarm only in both accuracy and response times, with participants in Experiment 1 and 2 achieving 58% accuracy and the participants in Experiment 3 achieving 77% accuracy. The effect of the Temperature alarm was also observed to be independent of musicality as a participant factor in response times, demonstrating that the removal of the Cardiovascular alarm directly affected the performance of the Temperature alarm only.

The results of Experiment 3 demonstrated what was suggested by the confusion diagrams in Experiment 1 and 2: that the presence of “persistent confusions” prevents participants from getting other alarms in that cluster correct.

8 EXPERIMENT 4 – REDUCING CONFUSION BY CHANGING ONE STIMULI FROM A "PERSISTENT CONFUSION" PAIR

Experiment 4 focuses on the differences in accurate identification between the Ascending and Descending clusters of alarms identified in the Experiment 3. Both groups contain two alarms, yet the ability to tell one apart from the other in the descending pattern appears to be easier than the ability to distinguish the alarms in the ascending pattern. The obvious difference between the two clusters is that the interval between the Powerfail alarm makes it more easily discernible from the Oxygen alarm where the uniformed stepped pattern of the Cardiovascular alarm is easily confused with the stepped pattern of the Temperature alarm. Studies into the ability of individuals to make accurate judgements on pitch intervals indicates that greater interval jumps would be more obvious to even untrained ears (Zwicker & Fastl, 1999). It is for this reason that the Powerfail and Oxygen alarms although of a similar musical contour are consistently more accurately identified than the Cardiovascular and Temperature alarms which also share a similar musical contour.

In this experiment, the Temperature alarm is a modification of the one supplied in the IEC 60601-1-8 standard for melodic alarms in order to further distinguish it from the Cardiovascular alarms. Using the Powerfail alarm as a template, the Temperature alarm was modified from the original (C4-D4-E 4—F4-G4) so that the interval between the 2nd and 3rd note would not be the same as the 1st and 2nd note (C4-D4-C5—C5-C5). This would still allow the modified Temperature alarm to conform to the melodic alarm standards, and fulfil the mnemonic assigned to it, whilst being a more distinguishable musical shape. The purpose of this experiment, therefore, is to determine if the test participants' ability to correctly identify the Temperature and Cardiovascular alarms can be improved with alarm redesign.

8.1 Method

The method employed was the same as the previous 3 studies. However, unlike Experiment 3, the Cardiovascular alarm was reintroduced in order to observe if there was still persistent confusion with the Temperature alarm. Additionally, the previous

Temperature alarm was replaced with the modified Temperature alarm as described above (see also Section 8.3).

Like Experiment 3, the Java program used was only set for Day 1 testing in the NM condition. Responses to “background” and “end of experiment” questionnaires were collected electronically via the Java program to reduce the amount of time and potential errors from manual entry of the data by the experimenter or data collector.

8.2 Participants

Participants were 19 students and members of the public from Newcastle, New South Wales and Melbourne, Victoria in Australia, recruited through the School of Psychology research sign-up system (Newcastle University) and announcements and flyers (Melbourne City). Participants were between 16 and 58 years old (Average 26, Median 22). 11 were male and 9 were female. Participants were asked if they had any formal musical training over 1 year (Musically trained =14, Not musically trained = 5). No participants had previous experience with the IEC 60601-1-8 melodic alarm sounds.

The study was conducted in accordance with the ethical review processes of the University of Newcastle.

8.3 Apparatus and Stimuli

The same alarms which were created for Experiment 2 were used for this study. The one exception was the Temperature Alarm which was recreated in Csound to sound like a modified version of the original Temperature alarm used in Experiment 2.

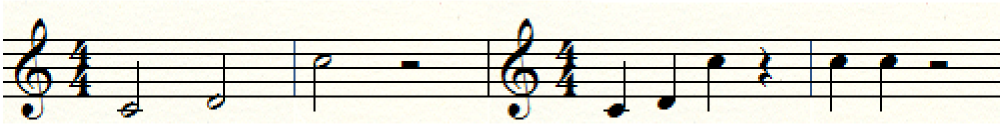


Figure 8-1: Musical Melody of the modified Temperature alarm for Medium Priority and High Priority

The alarm labels were displayed on a laptop screen by the same custom Java program used in Experiment 2 and 3. The screens would be identical to those used in Experiment 2 (Learning: Figure 6-1; Testing **Error! Reference source not found.**). No mnemonics were presented and all participants did the experiment in the Non-Mnemonic condition.

Responses and response times for each participant were recorded as individual comma separate values (.csv) files.

8.4 Procedure

The procedure for the experiment was identical to that described in Experiment 3. Prior to the start of the experiment, participants were asked to complete a short questionnaire which asked about their age, any known hearing impediments, whether they had any formal musical training and if so, what instrument and for how many years. Questionnaire data was collected electronically via the same Java program that presented the alarm stimuli and collected responses.

8.5 Results

The primary goal of this study was to discover if a modified Temperature alarm would have any significant effect on a person's ability to learn the alarm meanings.

Secondary goals of this study included an analysis of:

- Whether the modified alarm would result in an overall better performance than previously observed by musically trained participants; and
- Whether the modified alarm would result in an overall better performance than previously observed by non-musically trained participants.

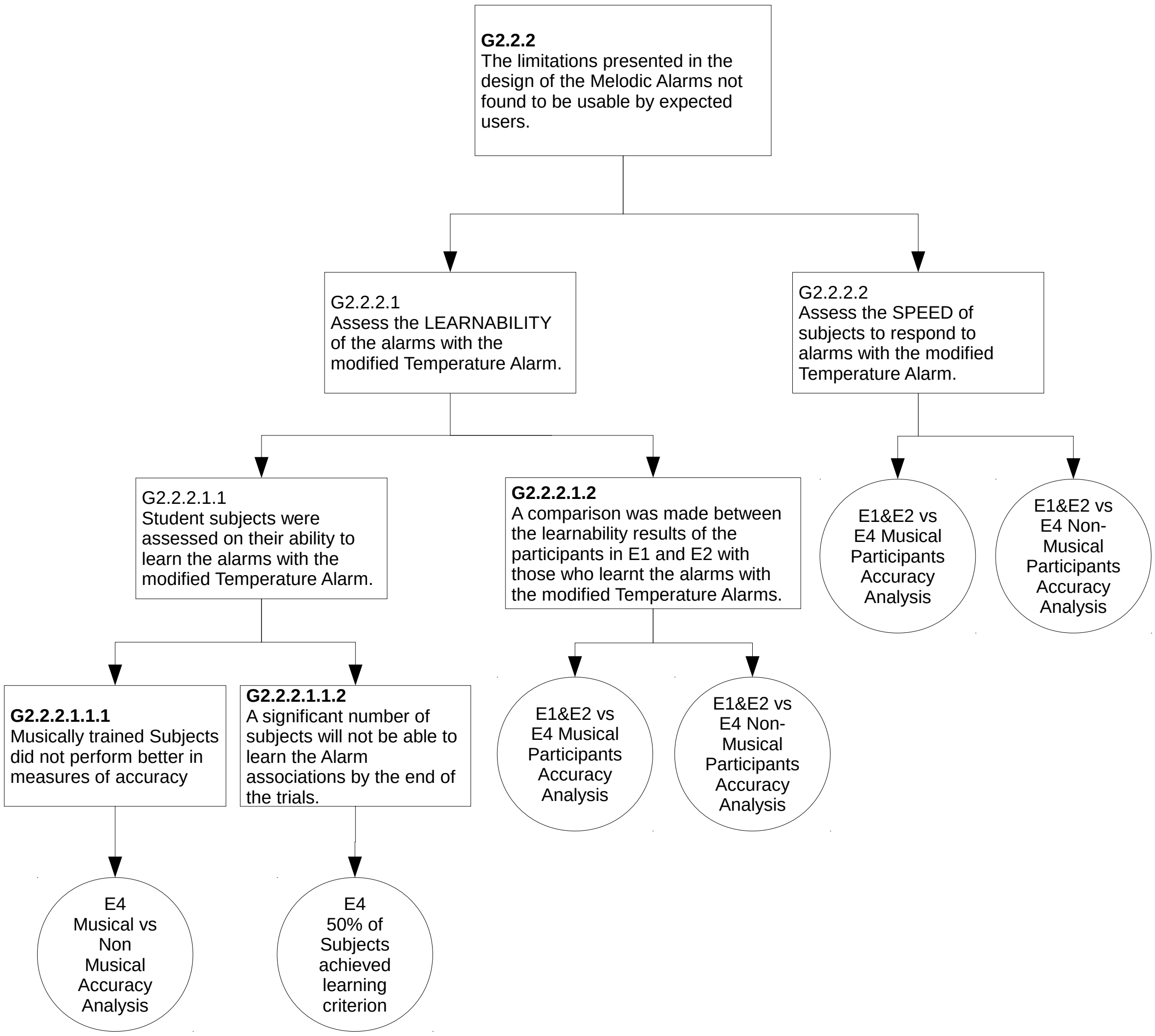


Figure 8-2: Tested hypothesis and Analyses for Experiment 4

8.5.1 Ability to learn the alarms

Of the 20 participants who took part in the study, 10 achieved the learning criterion at the end of the experiment. All participants who achieved 100% correct identification of the 16 alarms in two consecutive test cycles were from the group that identified themselves as having had at least one year of formal musical training. However there was no observable significant difference between musically trained participants and non-musically trained participants (Table 8-1). This may be due to the small number of non-musically trained participants in this group (5 out of 19) and that members of this group have shown (in previous experiments) to have a high variability of innate musical ability. There was a significant difference between the Non musically trained participants in Experiment 1 & 2 and the non musically trained participants in Experiment 4. This is likely due to the sample of non-musically trained participants in Experiment 4 not being high enough to account for this variability (relative pitch performance) across the group.

Data Set	Participants	Mean Correct by Block	p Value	t Stat	StDev	dF
Musical	14	70%	3.64E-01	2.390949315	0.202466	12
Non-Musical	5	60%		2.968686684	0.153194	3

Table 8-1: Experiment 4 Accuracy analysis of responses of musical and non-musical participants

8.5.2 Ability to identify the alarms in a timely manner

There was no observable effect of musicality on response speed (Table 8-2).

Data Set	Participants	Mean Time Taken by Trial	p Value	t Stat	StDev	dF
Musical	14	5.62	2.01E-01	2.390949315	1.214842	12
Non-Musical	5	4.95		2.968686684	0.721159	3

Table 8-2: Experiment 4 Response time analysis of responses of musical and non-musical participants

8.5.3 Does the modified alarm result in an overall better performance than previously observed?

Previous experiments have indicated that comparisons must be made by musical ability due to the strong effect that this variable has on the data. Therefore comparisons were made between musically trained participants in Experiment 1 and 2 with musically trained participants in Experiment 4 to see if the redesigned Temperature alarm had any effect on overall performance.

The modified alarm (Experiment 4) was not observed to result in better overall performance in participants than in Experiment 1 and 2. Musical training was observed to have a significant effect on accuracy of alarm identification (p=0.0001). A between factor effect of Musicality and Experiments is observed due to the high rate of accurate identification by non musical participants in Experiment 4 (44% to 60%, Experiment 1 and 2 vs Experiment 4).

Experiment 1&2 v. 4 (Accuracy) | ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.478377	1	0.478377	16.40868	0.000123	yes
Experiments	0.030656	1	0.030656	1.051512	0.308455	no
Inter	0.143498	1	0.143498	4.922093	0.029541	yes
Within	2.186542	75	0.029154			
Total	3.720323	78	0.047696			

Mean

	Exp. 1&2	Exp. 4	
Musical	0.754547	0.695292	0.72492
Non-Musical	0.443196	0.604301	0.523749
Mean	0.598871	0.649797	0.624334

Variance

	Exp. 1&2	Exp. 4	Mean
Musical	0.021278	0.044146	0.028172
Non-Musical	0.030951	0.029336	0.033223
Mean	0.04988	0.040097	0.047696

Table 8-3: Experiment 1&2 vs. 4 - (Accuracy): Two factor ANOVAs measuring the effect of musical training (musically trained vs. non-musically trained) and Experiment (Experiment 1 & 2 vs. Experiment 4)

It is interesting to note that even though the non musical population did better in Experiment 4 than in Experiment 1 and 2. The overall performance in alarm identification remained the same.

There was no appreciable effect on the overall accuracy (Table 8-4) or response time (Table 8-5) of participants using the new Temperature alarm. This suggests that the modified temperature alarm did not improve performance in identifying the Temperature alarm. All means were observed to be within one standard deviation of their experimental counterparts.

Experiment 1&2 v. 4 (Accuracy Temperature) | ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	0.739484	1	0.739484	13.96301	0.000362	yes
Experiments	0.0019	1	0.0019	0.035872	0.850293	no
Inter	0.052898	1	0.052898	0.998826	0.320808	no
Within	3.972012	75	0.05296			
Total	5.602319	78	0.071825			

Mean

	Exp. 1&2	Exp. 4	
Musical	0.729952	0.650379	0.690165
Non-Musical	0.412938	0.467157	0.440048
Mean	0.571445	0.558768	0.565107

Variance

	Exp. 1&2	Exp. 4	Mean
Musical	0.053216	0.039642	0.049483
Non-Musical	0.05005	0.113901	0.056857
Mean	0.076302	0.060813	0.071825

Table 8-4: Experiment 1&2 vs 4 - (Accuracy of Temperature alarm): Two factor ANOVAs measuring the effect of musical training (musically trained vs. non-musically trained) and Experiment (Experiment 1 & 2 vs. Experiment 4)

Experiment 1&2 v. 4 (Response Time Temperature) | ANOVA with regression

	SS	*df*	*MS*	*F*	*p-value*	*sig*
Musicality	1.397317	1	1.397317	0.408714	0.524571	no
Experiments	8.587366	1	8.587366	2.511798	0.117205	no
Inter	12.15509	1	12.15509	3.555356	0.063226	no
Within	256.4109	75	3.418812			
Total	291.6657	78	3.739304			

Mean

	Exp. 1&2	Exp. 4	
Musical	5.459745	5.62146	5.540603
Non-Musical	6.817614	4.951227	5.88442
Mean	6.138679	5.286344	5.712512

Variance

	Exp. 1&2	Exp. 4	Mean
Musical	4.010009	1.589366	3.227261
Non-Musical	4.031057	0.650087	3.944279
Mean	4.418356	1.384284	3.739304

Table 8-5: Experiment 1&2 vs 4 - (Response time of Temperature alarm): Two factor ANOVAs measuring the effect of musical training (musically trained vs. non-musically trained) and Experiment (Experiment 1 & 2 vs. Experiment 4)

8.5.4 Confusions

In analysis of the confused alarms made on more than 25% of a participant's trials, we see clear indications that the Cardiovascular alarm and the Temperature alarm are still being mutually confused by many participants as are the Infusion-Ventilation-Perfusion alarms. This indicates that the modified Temperature alarm is still too melodically similar to the Cardiovascular alarm so as to cause high levels of confusion between the two.

As expected, the confusions made by non-musically trained participants are far more varied than by those who have self identified as musically trained.

Figure 8-3 shows the pattern of confusions between alarms for all test trials conducted in the experiment. Percentages on links represent the percentage of participants who misidentified the sound at the origin of the arrow as the alarm at the point of the arrow on more than 25% of trials. The numbers in each circle indicate the percentage of accurate identification of that alarm made by participants in all trials. Three arrow weights are shown. The thinnest indicates confusions made by at least 2 participants followed by confusions made by at least 3 participants. Confusions made by 4 or more participants are presented as the thickest arrows with the percentage of participants who misidentified the alarms presented on the links.

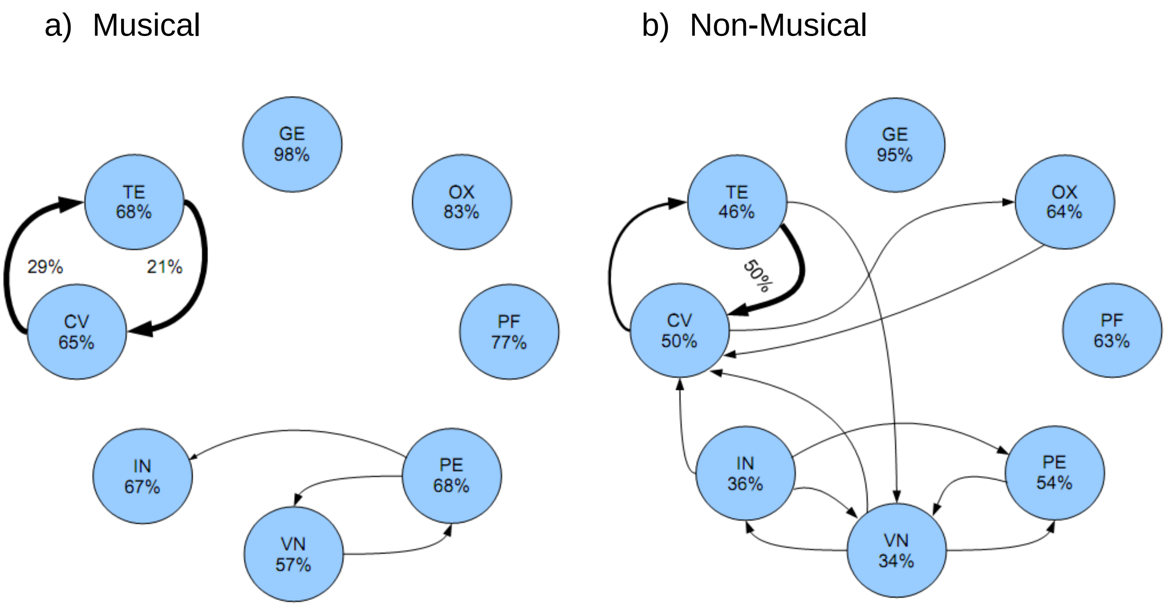


Figure 8-3: Patterns of confusions between alarms from Musical and Non Musical Participants in Experiment 4.

8.6 Conclusion

Although half of the participants achieved the learning criterion at the end of the experiment, there was no significant improvement in accuracy between participants in the Experiment 4 and those from Experiment 1 and Experiment 2.

Musical Participants		Non Musical Participants	
Experiment 1	78%	**Experiment 1**	51%
Experiment 2	68%	**Experiment 2**	38%
Experiment 4	70%	**Experiment 4**	60%

Table 8-6: Average Accuracy by block of responses in Experiment 1, Experiment 2 and Experiment 4 by musical training.

This finding, combined with the observed confusions that still exist between the Cardiovascular and Temperature alarms, indicate that the modified Temperature alarm is still too musically similar in shape to the Cardiovascular alarm to adequately disassociate the two alarms from being confused with each other.

Another recurrent observation is that the ability of participants to achieve better identification accuracy is too dependent on their musical ability. Musically trained participants have an advantage in that respect as they have techniques and practice to support them in identifying the alarms, not to mention a certain amount of interest in music to begin with. Non-musically trained participants are more unpredictable. Although no non-musically trained participant achieved the learning criterion in this experiment, the results of a small minority non-musically trained participants in this experiment and those previous are comparable to that of their musically trained counterparts, indicating that some innate talent or practice in pitch discrimination is possessed by these individuals.

9 DISCUSSION AND CONCLUSIONS.

This chapter summarises the overall findings of the research program.

Section 4.2 has provided the research goals in a top down hierarchy in order to explain how each research question was dissected for examination. The following overview of findings is presented from the bottom up in order to demonstrate how each experimental program (terminal point) in the diagram contributes towards an understanding of the higher order research goal.

9.1 Conclusions from Observations

9.1.1 The Effect of Mnemonics

As predicted in Section 4.6.3, Mnemonics had no observable effect in Experiment 1 when the alarms were taught to 33 students from the University of Queensland for course credit. The credible argument against the absence of an observable effect is that the mnemonics were designed for medical professionals who were familiar with the terminology and that the popular culture references were not well targeted towards an audience outside of the United States of America. Thus not meeting the requirements of useful environmental or ecological cues to aid memory (Leung et al., 1997; Norman, 1988). This hypothesis was tested in Experiment 2 with experienced practicing clinicians (22 Nurses). No significant effect of mnemonics was observed on speed and accuracy of identifying the alarms or in memory retention. In fact, a marginal effect of population was observed with the students performing better than the nurses. However, as described later in Section 9.1.3, this effect could be contributed to the fact that there were more musically trained participants in Experiment 1-Students than in Experiment 2-Nurses. Musicality has been shown by relevant studies (see discussion in section 2.1.3) and the experiments described here to have a strong effect on an individual's ability to learn the alarms.

Memory research also indicates that mnemonics are most useful when people generate their own associations (Edworthy & Hellier, 2005) as opposed to when they are provided with mnemonics. This was not explicitly tested for in Experiment 1 and Experiment 2. However participants in both experiments reported self-designing their own mnemonics. Mnemonic supplied participants reported finding only some of the supplied useful if they used the mnemonics at all. Participants in Experiment 3 and Experiment 4 were encouraged to come up with their own mnemonics or memory aids to help them in remembering the alarms and their associations.

Experiment 2 supports the conclusion of Experiment 1 in that the clinical associations of IEC 60601-1-8 melodic alarms assisted with mnemonics cannot be shown to be any more effective than if the mnemonics and labels had been assigned at random to the melodies. The effect of prior medical knowledge was not observed to aid the nurses in correctly identifying the alarms.

9.1.2 The Effect of Time shared Task

Experiment 1 and 2 were the first studies to evaluate the effects of mnemonics and timeshared tasks on participants' performance in identifying the IEC 60601-1-8 alarms. The timeshared task slowed alarm identification from an average of 5.2s to 6.1s and also led to a decrease in accuracy from 66% to 62% ($p < 0.001$), indicating that identifying the alarms was resource-demanding. The accuracy in identifying high priority alarms was more affected by the timeshared task than the medium priority alarms. Mnemonics were also not observed to aid a participant's ability to correctly identify the alarms while completing the Timeshared task.

9.1.3 The Effect of Musicality

As expected from related research (section 4.6.1), in nearly all the studies in this experimental program, musically trained participants performed significantly better than non-musically trained participants. Musically trained participants identified the alarms much more accurately and in most cases were able to achieve the learning criteria or very close to within the experiment time.

Musical training has been associated with ability to differentiate between melodies with pitch variations (Hyde et al., 2009; Lacherez, Seah, & Sanderson, 2006; Sanderson, Wee, & Lacherez, 2006; Schon, Magne, & Besson, 2004). This association was also found to be highly significant in every study conducted within this experimental program as well as related studies conducted by other researchers (Lacherez et al., 2006; Thompson et al., 2010), with musically trained participants being more accurate at identifying melodically differentiated alarms correctly. Although we would expect from previous music and pitch discrimination studies that musical training would be beneficial, the concern arises as to just how dependent the ability to identify alarms is on musical ability. This should not be interpreted to say that non musically trained individuals would not be able achieve passable performance in the alarm identifying task, as there are many factors which affect an individual's ability to discriminate between melodic tones, such as musical talent, pitch perception and informal musical background. Due to the melodic nature of the alarms and the supporting literature behind pitch sensitivities in musically trained individuals coupled with the evidence of the reported experiments, musical training is highly correlated with high accuracy performance in identifying and retaining the alarms and their associated labels.

This is most certainly not a reasonable assertion to make, however. As suggested by Brewster et al (1992), a successful design should be one in which all participants do well regardless of their musical background.

9.1.4 The Effect of Persistent Confusions

Some "persistent confusions" remained across all experiments even though in Experiment 4, a modified Temperature alarm was created to further distinguish it from having a similar musical contours to the Cardiovascular alarm. Modifying the alarm was also challenging as it was necessary to confirm to consistent rhythmic patterns established in ISO 9703-2 and adopted by IEC 60601-1-8. Experiment 3 also demonstrated that some "persistent confusions" remained even though the set of melodic alarms was reduced in size. Ventilation, Infusion and Perfusion were still mutually confused by a significant proportion of participants who were non-musically trained.

Based on these findings, alarms that are differentiated solely on a limited frequency (single octave) range of melodic (pitch varying) qualities may cause confusion in clinical practice if misidentified alarms cause delays and errors due to inappropriate or missed visual verification.

9.1.5 The issue of perceived urgency of the melodic alarms

A number of researchers have found that perceived urgency of alarms in the patient monitoring environment is inconsistent with actual urgency of the situation which triggers the alarms (Momtahan et al., 1993; Mondor & Finley, 2003). Block et al. (2000) also indicated a concern that conveyance of urgency by length of alarm and repetition at a faster speed might not convey a proper sense of urgency. However, studies into urgency mapping and perceived urgency have shown that increased speed, length and repetition (techniques used in the IEC medical melodic alarms) are useful in conveying a sense of increased urgency but at varying increments (Edworthy, Loxley, & Dennis, 1991; Hellier & Edworthy, 1999).

It is important to note that in the patient monitoring environment, the sense of urgency is largely dependent on context (Seagull & Sanderson, 2001) which is one reason where having each physiological system alarm being distinguished by one dimension of urgency differentiation may be a limiting the scope of potential useful alarm design. The design of the melodic alarms means that all high priority alarms will sound to be the same level of urgency when, in context, they are not as important (i.e. Heart Failure at the same time as a Power Fail warning).

The findings from the experiments outlined in this thesis show that participants, in general, agree with the different levels of urgency by rating the high priority alarms as more urgent sounding that the medium priority alarms (Section 5.5.5 and 6.4.5). Interestingly, musical participants rated the high priority alarms as sounding more urgent than non-musically trained participants did. The findings suggest that a sense of greater urgency was conveyed to participants by the high priority alarms, though it is not possible (in this context), to make any reliable deductions on what the scale might be for measuring how much more urgent the high priority alarms sound as compared to the medium priority alarms. Participants were informed of the priority of the alarms prior to the experiment and did not make any noticeable confusion between high and medium priority alarms, which indicates that the urgency mapping for the IEC alarms is at the very least easily learnt even by naive users.

9.2 Overview of findings from the overall research program

The first line of questioning involved evaluating the usability of the alarms. This was done by testing how easily the general population would be able to learn them, with or without the use of provided mnemonic help, and was accomplished by conducting the study (Experiment 1: The Effects of Mnemonics (Students)) described in Chapter 5.

9.2.1 Can IEC 60601-1-8 Melodic Alarms Be Taught To The General Population?

The results showed that overall learnability was poor but participants managed to easily distinguish between high and medium priority alarms. The majority of participants in both conditions struggled to achieve the passing criteria of two consecutive perfect scores even after 2 days of training sessions of approximately 45 minutes and 1h 30 minutes each. There was no observable effect of having the mnemonics as opposed to not having the mnemonics. However, a strong effect of participants' previous musical training was observed in both speed and accuracy in alarm identification. Musically trained participants also self-rated their confidence in alarm identification as higher than non-musically trained participants.

This initial examination of the IEC 60601-1-8 alarms has identified that serious usability issues in the suggested alarms which need to be rectified if they are to be a useful aid rather than a hindrance to safe medical provision.

9.2.2 Can IEC 60601-1-8 Melodic Alarms Be Taught To Medical Professionals (Nurses)?

In order to completely rule out the effects of mnemonics in the learnability of the alarms, the melodic alarms were required to be tested with the specialist population for which they were designed. However, no participants achieved the learning criterion on Day 1 and only one managed to do so by the end of 2 days of training. In a general comparison of population, the specialist population performed no better in identifying the alarms than then general population. The conclusion is that even prior medical knowledge and years of experience in the domain does not render the provided mnemonics of any use in increasing the learnability of the alarms. The effect of musicality was still strongly observed, indicating that significant musical training, though impractical, would be necessary for greater usability of the alarms.

9.2.3 Can Melodic Alarms Be Designed To Be More Easily Learnt?

Having established that the alarms in their current state did not satisfy the criteria for usability, it was necessary to determine if this was the case for mnemonic alarms in general or just the alarms provided in the standard. A persistent confusion pair was decoupled in order to see if accuracy in identifying its partner was observed to be increased in comparison to the other alarms presented at a rate significantly better than chance.

9.2.4 Can we identify the problems people are experiencing?

Once again musically trained participants have a significant advantage in being able to discriminate and remember the alarms. Experiment 3, conducted to explore this question and demonstrated that the presence of “persistent confusions” was indeed the primary cause of the low rate of accuracy amongst the majority of participants.

9.2.5 Can we identify the specific qualities that drive the error rate?

The results from Experiment 1, Experiment 2 and Experiment 3 clearly identified two clusters of alarms by observing the groupings and using a basic knowledge of musical theory it is very easy to identify how these alarms are cognitively being groups in the participants working memory. Moreover, the nature of the alarms allows us to easily eliminate variability due to rhythm and timbre which simply leaves pitch as a factor. These clusters are:

- Level pitch (e.g. General Alarm);
- Ascending pitch (e.g. Cardiovascular alarm and Temperature alarm);
- Descending pitch (e.g. PowerFail alarm and Oxygen alarm); and
- Varying pitch (Ventilation alarm, Perfusion alarm and Infusion alarm).

The General Alarm is the only alarm in the Level pitch category and as such was the most accurately identified. The alarms in the Ascending pitch category were most frequently confused with one another. In non-musically trained participants the probability of identifying this alarm correctly or confusing it with its confusion partner is practically chance. Alarms in the Descending Pitch category were more accurately identified than those in the Ascending or Varying pitch cluster. This is likely due to the melodic shape of the alarms, with the Powerfail alarm displaying an easily identifiable drop in pitch as compared to the Oxygen alarm’s steady and gradual pitch drop. The Varying Pitch Cluster is made up of 3 alarms which are frequently confused due to the sharp ascending and descending of pitch. Although all participants had significant difficulties in correctly identifying the alarms in this group, the correct identification of these alarms were the worst for non-musically trained participants.

9.2.6 Can the Identified Problems Be Overcome?

Working within the standard's guidelines a different Temperature alarm was designed to make it more auditory distinct from the Cardiovascular alarm. The changes were not made arbitrarily. Temperature and Cardiovascular are similar because they both form an ascending melodic shape. Within the standard there is another pair Oxygen and Powerfail which both form a descending melodic shape. Oxygen and Powerfail don't get confused for each other as often as Temperature and Cardiovascular. The biggest possible reason being that the intervals between the notes are far greater in Oxygen and Powerfail than in Temperature and Cardiovascular. Temperature alarm was changed to mimic this and include a greater interval. Unfortunately observed confusions still exist between the Cardiovascular and modified Temperature alarms indicating that the modified Temperature alarm is still too musically similar in shape to the Cardiovascular alarm to adequately dissociate the alarms from each other.

Based on the constraints of the proposed standard and supporting literature evidence into relative pitch sensitivity in musically trained and untrained individuals (Hyde et al., 2009; Tervaniemi et al., 2005; Van Zuijen et al., 2004) and alarm learning (Patterson & Milroy, 1980), the identified problem of persistent confusions will be difficult to overcome. The constraints of only being able to differentiate between alarms by restrictive pitch melodies without allowing modulation in rhythm, timbre or octave, prevents participants from being able to learn and discriminate between the alarms.

9.2.7 Are the IEC60601-1-8 Melodic Alarms Useful?

The overriding goal of this research was to explore if the proposed melodic alarms achieved the underlying goal of being able to provide encoded information within the alarm sound (Block, 1994). This would be determined by examining the alarms against a number of standard measures of usability and utility to answer the question of:

> *Can the alarms could be taught to medical professionals to a level where they would be able to utilise the auditory displays to aid them in performing their patient monitoring tasks.*

The literature shows significant advantages and uses for informative auditory displays in domains such as aerospace, nuclear power, defence. This thesis does not dispute that informative alarms would provide significant improvements to the field of patient monitoring rather that the current limitations of the proposed standard does not enable practitioners to easily identify the encoded information within the alarm. At best the alarms would in no way perform differently to arbitrary alarms that are currently in prevalence in the patient monitoring workspace, which simply direct attention to the general vicinity of the piece of equipment producing the sound. At worst, the incorrect identification of an alarm may cause a misdiagnosis of the problem and lead to incorrect action or delays in performing the correct action, which could lead to severe to disastrous outcomes.

9.3 Limitations to the experimental program

There are several limitations to the experimental program. These are:

1. The alarms were only tested for absolute identification;
2. Experiment 1 and 2 lacked the necessary power to adequately determine if mnemonics were or were not useful;
3. The standard does not require that all alarms be differentiated;
4. The alarms were not tested under clinical conditions.

First, the IEC 60601-1-8 alarms were only tested for absolute auditory identification and in the absence of redundant visual alarms.

Secondly, retrospective power analysis indicates that Experiment 1 and 2 were not adequately powered to address the question of whether the mnemonics were or were not useful.

Thirdly, the standard does not require that all alarms are different (The standard general alarm is used in a majority of machines currently in clinical operation); therefore the result might apply to only extreme conditions.

Lastly, the alarms were not tested under clinical conditions in the context of an episode of care. Many additional factors would need to be taken into account in order to achieve a holistic view of how the alarms would perform in a genuine medical situation. Several factors would arguably make the alarms easier to identify:

- Availability of visual confirmation;
- Context of the patient care (e.g. Perfusion alarms and Cardiovascular alarms will very unlikely be heard in the same context of a particular patient's care as perfusion machines mainly be in operation when the heart is undergoing bypass and cardiac monitors are not used. It may also be interesting to point out in this context that the Cardiovascular and Temperature alarms, which have been shown to be the most persistently confusable pair would be regularly heard in the same patient care context due to their applications in multiple patient monitoring equipment);
- Area of speciality. (e.g. Operating room nurses may quickly learn to recognise alarms that would frequently occur in the operating room such as Infusion, Cardiovascular and Ventilation and rely on visual confirmation for any unrecognised sounds).

Whilst other factors may actually make the alarms more difficult to identity:

- Overlapping alarms. Lacherez et al (2007) found that overlapping melodic alarms were almost indistinguishable due to participants' inability to segregate the auditory stream.
- Activation of the alarms in areas with ambient noise. The alarms were tested in quiet rooms where participants were undisturbed whilst undergoing testing. No tests were conducted where there was any background noise or with partial masking of any of the sounds.
- Additional workload from high pressure situations. Alarm proliferation generally occurs during periods of high stress and workload for medical personnel.

9.4 Why even have a standard?

The findings from this experimental program contribute along with other related studies into the susceptibility of the proposed melodic alarms being affected by overlapping (Lacherez et al., 2006), masking by ambient music (Sanderson, Tosh, et al., 2005) and difficult to learn (Williams & Beatty, 2005). The findings demonstrate that, at the very least, supporting research needs to be put into determining the requirements for a standardised set of alarms before one can be proposed and implemented. However, another question has arisen of whether a standardised set of alarms are required at all. This argument has a number of points including:

- Standardising limits innovation: Manufacturers and developers will be restricted in what they can use in their products thus make products only in a certain way when other options are as good, or even better with greater scientific knowledge.(Acemoglu, Gancia, & Zilibotti, 2012; Cargill, 2011; Sack, 2011)
- Standards are often heavily influenced by the concerns of the businesses that will be affected by them.
- Standard designs that cannot be shown to be useful or safe may result in the law, industry, developers and customers ignoring aspects of the standard thereby reducing confidence in the regulatory body which developed the standard. (Cargill, 2011)
- Changes in standards once established can be expensive for manufacturers and result incompatibility between systems and equipment.

Weinger (1991) has also pointed out how it can be hard to get the academics and the industry to achieve a consensus on the standard which is likely reason why the proposed Patterson et al. (1986) alarms lost the original acoustic richness and were so heavily simplified.

Standardisation of systems and technical equipment can be useful in increasing compatibility, interoperability (Sack, 2011), assurance of adequate safety levels, familiarisation with users and minimum levels of quality. However, any standard should be able to demonstrate that the restrictions that they are imposing represent the current accepted best practice as supported by industry as well as current scientific understanding.

Unless standards provide clear benefits and serve an obvious need, they will not inspire the necessary confidence and will be difficult to establish in the domain.

10 PRACTICAL IMPROVEMENTS AND RESEARCH MOVING FORWARD

The studies investigated in this thesis represent an initial investigation into a set of proposed auditory alarms, which have been introduced into a Standard without sufficient consideration of the Physical (biological limitations) and Cognitive (psychological limitations) abilities of the general population. Having established the problem in Experiment 1 and Experiment 2, the actual limitations of human musical perception become more obvious in Experiment 3 and Experiment 4. These studies establish that the particular use of limited pitch variation as a form of alarm discrimination in the IEC alarms is not conducive to achieving the goal of informative alarms in medical domains.

However, additional studies are needed to explore if pitch variation in melodic alarms is completely ineffective in conveying information to the general population. The IEC alarms were limited in frequency to one octave. There is suggestive evidence in the better than average performance of participants in the Oxygen and Powerfail identification tasks which suggests that varying intervals between groups of sounds is an effective method of alarm discrimination. A wider frequency range may allow greater flexibility of intervals which in turn may aid in the alarm identification task.

Other dimensions of music should also be explored. The original Patterson et al. (1986) alarms have been very promising in achieving similar high performance in the identification task by musically trained and non-musically trained participants alike (Thompson et al., 2010). These alarms are, however, an acoustically varying and rich set of alarms which incorporate variations in pitch, rhythm, volume and timbre, which may panic and dismay manufacturers and equipment designers who will have to invest heavily in high quality sound technology in order to meet the requirements of this design. It is currently not possible to deduce from the Patterson et al. (1986) alarms what acoustic aspects of the alarm are the most effective at making them easily discriminated. A series of studies examining the effect of each, or combinations of each, acoustic property to determine what achieves a high differentiation and identification performance would definitely be an interesting follow up to this research. A particular goal of this study would be to assess the ability of musically trained participants as opposed to non-musically trained participants with an end to suggesting a simple and practical design aimed at reducing the dependency of accuracy on a trained musical ear or innate musical ability.

As yet, the design of informative auditory alarms has not matured to a point where this mode of information communication can be compared with other forms of monitoring (visual monitoring or sonification). Medical monitoring displays can only be compared to the generic notification alarms which currently exist in use in the medical domain. This is an unfair comparison as the intention and capabilities of these displays are not directly comparable.

Wilcox's (2011) chapter in Weinger, Wiklund, and Gardner-Bonneau (2011) recommends that an important design guideline for Alarm design in Medical devices is to determine which signalling modalities should be allocated to a specific alarm condition. Care must be taken to consider whether the user of the device may have a sensory limitation such as impaired hearing or vision and therefore redundant signalling modalities may be necessary (e.g. Combination of alarm and visual; sonification + visual; etc).

Only after having established a research based ideal of elements to be incorporated into informative alarm design, can we revisit the studies described in:

- Wee (2003) - Alarms vs. Sonification in Anaesthesia;
- Sanderson, Tosh, Philp, Rudie, Watson and Russell (2005) - The effects of ambient music on simulated anaesthesia monitoring; and
- Lacherez et al. (2006) – Overlapping melodic alarms are almost indescriminable

A revisit of these studies should be applied to any new alarm design to establish any limitations or combinations of displays would best support the work of medical practitioners. The aims of these research studies would be to eventually reduce the Alarm problem as described by Woods (1995) and effectively reduce the cognitive workload currently necessary for active patient monitoring.

10.1 Practical Significance of the work

This research program in conjunction with the studies of other researchers (independent and related to this research program) (Lacherez et al., 2006; Sanderson, Wee, Seah, et al., 2006; Thompson et al., 2010; Thompson et al., 2005; Weinger, 1991; Williams &

Beatty, 2005) has demonstrated that the IEC 60601-1-8 melodic alarms have a serious design flaw when it comes to an individual's ability to utilise these melodic tones to garner information about an abnormal parameter or to decrease workload in medical staff. The heavy dependence on the musical ability of an individual to be able to perform well in the identification task is an unreasonable expectation for your typical medical practitioner.

Moreover, medical equipment manufacturers, who aim to keep their products competitive and up to date and have implemented these suggestions from the IEC 60601-1-8, may have inadvertently added to the alarm problem described by Woods (1995). This may have been done at great expense to the equipment designer as many pieces of medical equipment, up to recently, still used low cost piezoelectric audible alarms for their audible alarm signalling. The same kinds of alarms used in smoke detectors and checkout counters at grocery stores. These alarms would no longer be able to be used as they do not meet the complex frequency requirements of the IEC 60601-1-8 melodic alarms.

10.2 Conclusion

The results of this study and other supporting literature (Lacherez et al., 2006; Thompson et al., 2010; Williams & Beatty, 2005) strongly indicate that the learnability and discriminability of the IEC melodic alarms are issues that need to be addressed before they are introduced into healthcare settings. Moreover, there is an inherent danger that due to some persistent confusion between certain alarm groups, there is a possibility of alarm mix-ups in clinical practice. If an alarm does not indicate its meaning rather only that something is wrong, then the clinician must investigate the cause of the alarm by using other sources of information such as visually checking the monitors of patient, checking the patient's temperature with touch or listening to their breathing, in order to determine the reason for the alarm. If however, alarms are meant to indicate the meaning of the alarm by its melody, a clinician may confuse one alarm for another, fail to do a visual check and potentially embark on an inappropriate treatment pathway resulting in the possibility of wasting time or in extreme cases potential harm to the patient.

It is reasonable to deduce that the intent of the IEC melodic alarms was to introduce the benefits of conveying basic information of which physiological system the alarm was referring to, thus elevating the alarm function from a basic alerting to convey meaning similar to the more informative alarm or earcon. This noble endeavour, to achieve greater ubiquitous information transfer which does not get missed due to inattention or distraction, has, in the case of the IEC 60601-1-8 melodic alarms, failed.

This is not to say that information conveying alarms are too difficult to implement and should therefore be abandoned. Patterson et al. (1986) designed a set of alarms distinguished by pitch, rhythm, dynamics, and the sonic qualities of timbre and texture for each of the physiological systems (General, Cardiovascular, Temperature, etc) described in the four experiments of this study. Compared to this acoustically rich design, the IEC 60601-1-8 melodic alarms were only distinguished by pitch (having consistent rhythm and being restricted in the sonic qualities by a frequency limitation). It is therefore no surprise that in a 2010 study by Thompson et al, for the Australian and New Zealand Collage of Anaesthetists Annual Scientific meeting, found that while the IEC 60601-1-8 alarms (as well as a modified set of alarms which were designed as a compromised base on the IEC set) were greatly affected by the musical ability of the listener, the Patterson et al. (1986) alarms were identified as accurately with those who had no prior musical training as those who had. Moreover, it was observed in the same study that users of the Patterson alarms performed better than those participants who were tested on the IEC or IEC related alarms.

10.2.1 Addressing the Aims of the research

The aims of the research as listed in Section 1.1 were designed to answer the overarching question of whether the IEC 60601-1-8 melodic alarms were currently suitable to be presented as an industry standard by which manufacturers and international bodies should conform to. This was achieved by examining the psychoacoustic qualities of the alarms and examining the relevant literature which addresses the benefits and attributes of these qualities and how they are interpreted by a typical human brain. Following this, tests were developed to test the theoretical expectations of an individual’s alarm identification performance based on standard measures of Usability and Utility commonly used in Human Factors and Interface design studies.

Weinger (1991) expressed a valid concern that it was premature to include a standard for melodic alarms by physiological system without a holistic examination into the people (capabilities), environment and events associated with using the alarms. The research presented in this thesis supports this view and has presented evidence that the cognitive, and aural capabilities of potential users of the system would be hard pressed to achieve the desired effect of automatic identification of the alarms to the overall benefit of the patient and field of medical science.

11 REFERENCES (ENDNOTE)

Acemoglu, D., Gancia, G., & Zilibotti, F. (2012). Competing engines of growth: Innovation and standardization. *Journal of Economic Theory, 147*(2), 570-601. e573.

Bacon, S. P. (2006). Auditory Compression and Hearing Loss. *Acoustics Today, 2*(2), 30-34.

Bellas, J. A., & Howard, J., J H. (1987). Interpreting the language of enviromnetal sounds. *Environment and Behaviour, 19*, 91-114.

Blattner, M. M., Sumikawa, D., & Greenberg, R. (1989). Earcons and Icons: Their Structure and Common Design Principles. *Human Computer Interaction, 4*, 11-44.

Block, F. E. (1992). Evaluation of users' abilities to reconize musical alarm tones. *Journal of Clinical Monitoring, 8*, 285-290.

Block, F. E. (Ed.). (1994). *Human Factors and Alarms* (2 ed.). Philadelphia: W.B. Saunders.

Block, F. E., Rouse, J. D., Hakala, M., & Thompson, C. L. (2000). A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. *Journal of Clinical Monitoring, 16*, 541-546.

Blom, J. A. (1988). Real-time expert systems in patient monitoring. *Journal of Clinical Monitoring*(4), 130-131.

Boquet, G., Bushman, J., & Davenport, H. (1980). The anesthetic machine-a study of function and design. *British Journal of Anaesthesia, 52*(1), 61-67.

Brewster, S. A. (1997). Using non-speech sound to overcome information overload. *Displays, 17*(3), 179-189.

Cargill, C. F. (2011). Why Standardization Efforts Fail. *Journal of Electronic Publishing, 14*(1).

Chou, R., Dana, T., Bougatsos, C., Fleming, C., & Beil, T. (2011). Screening adults aged 50 years or older for hearing loss: a review of the evidence for the U.S. preventive services task force. *Ann Intern Med, 154*(5), 347-355. doi: 10.7326/0003-4819-154-5-201103010-00009

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality and Safety in Health Care, 14*, 212-215.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 33*(2), 205-231.

Edworthy, J., & Meredith, C. (1997). Influence of Verbal Labeling and Acoustic Quality on the Learning and Retention of Medical Alarms. *International Journal of Cognitive Ergonomics, 1*(3), 229-243.

Ehmer, R. H. (1959). Masking by tones vs. noise bands. *Journal of Acoustical Society of America, 31*, 1253-1256.

Fitch, T., & Kramer, G. (Eds.). (1994). *Sonifying the body electric: Superiority of an auditory over a visual display in a complex, multivariate system*.

Gaba, D., & Howard, S. (1995). Situation awareness in anesthesiology. *Human Factors, 37*, 20-31.

Gaver, W. (1997). Auditory Interfaces. In M. Helander, T. Landauer & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction* (2nd ed.). Amsterdam.

Gelfand, S. A. (2004). *Hearing: And Introduction to Psychologycal and Physiological Acoustics* (4th Edition ed.): Taylor & Francis.

Guillaume, A., Jacob, E., Bourgeon, L., Valot, C., Blancard, C., Pellieux, L., . . . Cazalaà, J. B. (2005, 6-9 July 2005). *Analysis of auditory warning signals in the operating room.* Paper presented at the International Conference on Auditory Displays, Limerick, Ireland.

Hakkinen, M. T., & Willignes, B. H. (1984). Synthesized warming messages: Effects of an alerting cue in single- and multiple-function voice synthesis systems. *Human Factors, 26*, 185-195.

Hellier, E., & Edworthy, J. (1999). On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics, 30*(2), 167-171.

Huang, C., & Shachter, R. (1997). Alarms for monitoring: A decision-theoretic framework. *Section on Medical Informatics* (Vol. 97): Stanford University, Standord, CA.

Hyde, K., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A., & Schlaug, G. (2009). The Effects of Musical Training on Structural Brain Development - A Longitydinal Study. *The Neurosciences and Music III: Disorders and Plasticity: Ann. N.Y. Acad. Sci, 1169*, 182-186.

IEC60601-1-8. (2003-08-14). Medical Electrical equipment Part 1-8. Geneva, Switzerland: International Electrotechnical Commission.

ISO-9703. (1992). Anaesthesia and respiratory care alarm Signals - Part 2.

Jesteadt, W., Wier, C. C., & Green, D. M. (1977). Intensity discrimination as a function of frequency and sensation level. *The Journal of the Acoustical Society of America, 61*(1), 169-177. doi: doi:http://dx.doi.org/10.1121/1.381278

Kelly, T., & Weaver, R. (2004). *The goal structuring notation–a safety argument notation.* Paper presented at the Proceedings of the dependable systems and networks 2004 workshop on assurance cases.

Kemeny, J. G. (1979). The Need for Change: The Legacy of Three Mile Island. Report of the President's Commission on Three Mile Island: New York: Pergamon Press.

Kramer, G. (1994). *Some organising principles for representing data with sound.*: Sante Fe Institute studies in the sciences of complexity, Proceedings Volume XVIII, Reading, MA: Addison-Wesley.

Lacherez, P., Seah, E., & Sanderson, P. (2006). Overlapping melodic alarms are almost indiscriminable. *Human Factors*.

Leung, Y. K., Smith, S., Parker, S., & Martin, R. (1997). *Learning and retention of auditory warnings.* Paper presented at the Proceedings of the Third International Conference on Auditory Display (ICAD).

Loeb, R. G., Brunner, J. X., Westenskow, D., Feldman, B., & Pace, N. L. (1989). The Utah anaesthesia workstation. *Anaesthesiology, 70*, 999-1007.

Loeb, R. G., & Fitch, T. (2000, 13-15 January 2000). *Laboratory evaluation of an auditory display designed to enhance intra-operative monitoring.* Paper presented at the The society for Technology in Anaesthesia Annual Meeting, Orlando.

Loeb, R. G., Jones, B. R., Leonard, R. A., & Behrman, K. (1992). Reconition Accuracy of Current Operating Room Alarms. *Anesthesia Analgesia, 75*, 499-505.

Macleod, K., & Carr, C. (2005). Synaptic Physiology in the cochlear nucleus angularis of the chick. *Journal of Neurophysiology, 93*(5), 2520-2529.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science, 19*(12), 1263-1271.

Mcgookin, D., & Brewster, S. A. (2004). Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Trans. Appl. Percept., 1*, 130-155.

Mendelson, J. R., & Ricketts, C. (2001). Age-related temporal processing speed deterioration in auditory cortex. *Hear Res, 158*(1-2), 84-94.

Meredith, C., & Edworthy, J. (1995). Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of Advanced Nursing, 21*, 15-20.

MIL-HDBK-759C. (1995). Handbook for Human Engineering Design Guidelines. USA: US Department of Defence.

Miller, G. (1956). The magical number seven plus or minus two: some limited on our cappacity for processing information. *Psychological Review, 63*(2), 81-97.

Momtahan, K., Hetu, R., & Tansley, B. (1993). Audibility and identification of auditory alarms in the operating room. *Ergonomics, 36*(10), 1159-1176.

Mondor, T. A., & Finley, G. A. (2003). The perceived urgency of auditory warning alarms used in the hospital operating room is inappropriate. *Canadian Journal of Anesthesia, 50*(3), 211-228.

Mylrea, K., Orr, J., & Westenskow, D. (1993). Integration of monitoring for intelligent alarms in anaesthesia: Neural networks- can theyt help. *Journal of Clinical Monitoring*(9), 31-37.

Navabi, M. J., Mylrea, K., & Watt, R. C. (1989). *Detection of false alarms using an integrated anaesthesia monitor.* Paper presented at the IEEE Engineering Medical Biological Society 11th Annual International Conference.

Nielsen, J. (2001, 21 January 2001). Usability metrics. from http://www.nngroup.com/articles/usability-metrics/

Norman, D. A. (1988). *The psychology of everyday things*. New York: Harper & Row.

Orr, J. A., & Westenskow, D. R. (1994). A breathing circuit alarm system based on neural networks. *Journal of Clinical Monitoring, 10*, 101-109.

Patterson, R. D. (1982). *Guidlines for auditory warning systems on civil aircraft*.

Patterson, R. D. (1990). Auditory warning sounds in the work environment. *Philosophical Transaction of the Royal Society of London, B327*, 482-492.

Patterson, R. D., Edworthy, J., Shailer, M. J., Lower, M. C., & Wheeler, P. D. (1986). Alarm Sounds for medical equipment in intensive care areas and operating theatres *Institute of Sound and Vibration Research Report No. AC598*: Institute of Sound and Vibration.

Patterson, R. D., & Milroy, R. (1980). [Auditory warnings on civil aircraft: the learning and retention of warnings].

Philp, S. (2004). *Attentional resource allocation in sonification research.* (Honours), The University of Queensland, St Lucia, Queensland, Australia.

Potter, P., Wolf, L., Boxerman, S., Grayson, D., Sledge, J., Dunagan, C., & Evanoff, B. (2005). Understanding the Cognitive Work of Nursing in the Acute Care Environment. *Journal of Nursing Administration, 35*(7/8), 327-335.

Rasmussen, J., & Vicente, K. (1989). Coping with human errors through system design: implications for ecological interface design. *International Journal of Man-Machine Studies, 31*, 517-534.

Rheineck-Leyssius, A. T., & Kalkman, C. J. (1999a). Advanced pulse oximeter signal processing technology compared to simple averaging 1. Effect on frequency of alarms in the Operating Room. *Journal of Clinical Anesthesia, 11*, 192-195.

Rheineck-Leyssius, A. T., & Kalkman, C. J. (1999b). Advanced pulse oximeter signal processing technology compated to simple averageing 2: Effect on Frequency of alarms in the Postanesthesia care unit. *Journal of Clinical Anesthesia, 11*, 196-200.

Rich, K., Blanchard, H., & McCloskey, J. (2007). The use of goal structuring notation as a method for ensuring that human factors is represented in a safety case.

Ross, B., Herdman, A. T., & Pantev, C. (2005). Stimulus induced desynchronization of human auditory 40-Hz steady-state responses. *Journal of Neurophysiology, 94*(6), 4082-4093.

Rouder, J. N., Morey, R. D., Cowen, N., & Pfaltz, M. (2004). Learning in a unidimensional absolute identification task. *Psychonomic Bulletin and Review, 11*, 932-938.

Sack, J. (2011). Innovation and Standardization: Friends not Foes. *Information Standards Quarterly, 23*(4), 22-23.

Sanderson, P., Anderson, J., & Watson, M. (2000). *Extending exological interface design to auditory displays.* Paper presented at the Proceedings of the 10th Australasian Conference on Computer-Human Interaction OzChi2000, Sydney, Australia.

Sanderson, P., Crawford, J., Savill, A., Watson, M., & Russell, W. J. (2004). Visual and auditory attention in patient monitoring: A formative analysis. *Cognition Technolgy & Work, 6*(3), 172-185.

Sanderson, P., Tosh, N., Philp, S., Ridue, J., Watson, M., & Russell, W. J. (2005). Effects of ambient music on simulated anaesthesia monitoring with visual and auditory displays. *Anaesthesia, 60*, 1073-1080.

Sanderson, P., Watson, M., & Russell, W. J. (2005). Advanced patient monitoring displays: Tools for continuous informing. *Anesthesia and Analgesia, 101*(1), 161-168.

Sanderson, P., Wee, A., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia, 61*, 142-147.

Sanderson, P., Wee, A., Seah, E., & Lacherez, P. (2006, 20-23 June). *Auditory alarms, medical standards and urgency.* Paper presented at the International Conference on Auditory Displays (ICAD2006), Queen Mary University of London.

Sarter, N. B. (2000). The need for multisensory interfaces in support of effective attentional allocation in highly dynamic event-driven domains: the case of cockpit automation. *The International Journal of Acatiation Psychology, 10*(3), 231-241.

Sauro, J. (2010). *A practical guide to measuring usability*. CreateSpace Independent Publishing Platform.

Schmidt, S. I., & Baysinger, C. L. (1986). Alarms: Help or Hindrance? (letter). *Anesthesiology, 64*, 654-655.

Schon, D., Magne, C., & Besson, M. (2004). The Music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*, 341-349.

Seagull, F. J., & Sanderson, P. (2001). Anaesthesia alarms in surgical context: An Observational study. *Human Factors, 43*(1), 66-77.

Seagull, F. J., Wickens, C. D., & Loeb, R. G. (2001). *When is less more? Attention and workload in Auditory, Visual and redundant Patient-Minitoring Conditions.* Paper presented at the Proceedings of the Human Factors and Ergonics Society 45th Annual Meeting.

Seagull, F. J., Xiao, Y., Mackenzie, C., & Wickens, C. D. (2000). *Auditory alarms: From alerting to informing.* Paper presented at the Proceedings of the joint meeting of the Human Factors and Ergonics Society and the International Ergonics Association (IEA2000/HFES2000), Santa Monica, CA.

Seidman, M. D., Ahmad, N., & Bai, U. (2002). Molecular mechanisms of age-related hearing loss. *Ageing research reviews, 1*(3), 331-343.

Shackel, B. (Ed.). (1991). *Usability-context, framework, defination,design and evaluation*. Cambridge UK: Cambridge University Press.

Sharpe, L. T., Stockman, A., & Jägle, H. (Eds.). (1999). *Opsin genes, cone photopigments, color vision and color blindness.* Cambridge: Cambridge University Press.

Stanton, N. (Ed.). (1994). *A human factors approach* (Vol. Human Factors in Alarm Design): Taylor & Francis.

Stanton, N., & Barber, C. (1997). Comparing verbal alarm displays: Speech verses textual systmes. *Ergonomics*(40), 1240-1254.

Sumikawa, D. (1985). *Guidelines for the integration of audio cues into computer user interfaces.* (Masters of Philisophy Masters), University of California, Davis.

Sumikawa, D., Blattner, M. M., Joy, K., & Greenberg, R. (1986). Guidelines for the syntactic design of audio cues in computer interfaces: Lawrence Livermore National Laboratory.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental brain research, 161*(1), 1-10.

Thompson, C. L., Sanderson, P., Watson, M., Thompson, M., Muthukrishna, M., & Murphy, S. (2010, 1-5 May 2010). *Testing auditory alarm effectiveness with three different alarm sets.* Paper presented at the Australian and New Zealand College of Anaesthetists Annual Scientific Meeting 2010, Christchurch. New Zealand.

Thompson, C. L., Watson, M., & Sanderson, P. (2005). Auditory alarms in anaesthesia - can they be improved? ANZCA application: ANZCA.

Tsien, C. (1997, 25-29 Oct). *Reducing false alarms in the intensive care unit: a systematic comparison of four algorithms.* Paper presented at the AMIA Fall Symposium, Nashville, TN.

Van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. *Journal of cognitive neuroscience, 16*(2), 331-338.

Viemeister, N. F., & Bacon, S. P. (1988). Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones. *The Journal of the Acoustical Society of America, 84*(1), 172-178. doi: doi:http://dx.doi.org/10.1121/1.396961

Walsh, T., & Beatty, P. (2002). Human Factors error and patient monitoring. *Physiological Measurement, 23*(3), R111-R132.

Watson, M., Russell, W. J., & Sanderson, P. (1999). *Ecological interface design for anaesthesia monitoring.* Paper presented at the Proceedings of the 9th Australasian Conference on Computer-human Interaction OzChi99, Wagga Wagga, Australia.

Watson, M., & Sanderson, P. (2001, 8-12 October). *Intelligibility of Sonification for respiration monitoring in anaesthesia.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 45th meeting, Minneapolis, Minnesota.

Watson, M., & Sanderson, P. (2004). Respiratory Sonification helps eyes-free physiological monitoring and task timesharing. *Human Factors, 6*(3), 497-517.

Watson, M., Sanderson, P., & Russell, W. J. (2004). Tailoring reveals information requirements: the case of anaesthesia alarms. *Interacting with Computers, 16*(2004), 271-293.

Wee, A. (2003). *Alarms vs. Sonification in Anaesthesia.* (Honours Thesis), University of Queensland, Brisbane.

Wee, A., & Sanderson, P. (2005, 12-16 September). *Testing New Alarms for Medical Electrical Equipment.* Paper presented at the Tenth IFIP TC13 International Conference on Human-Computer Interaction (Interact2005), Rome.

Wee, A., & Sanderson, P. (2006, 16-20 October). *Effects of mnemonics in learnability of melodic alarms with registered nurses.* Paper presented at the 50th Annual Meeting of the Human Factors and Ergonomics Society, San Francisco, CA.

Weinger, M. B. (1991). Proposed new alarm standards may make a bad situation worse (Letter). *Anesthesiology, 74*, 791-792.

Weinger, M. B., & Englund, C. E. (1990). Ergonomic and Human Factors Affecting Anesthetic Vigilance and Monitoring Performance in the Operating Room Environment. *Anesthesiology, 73*, 995-1021.

Weinger, M. B., Wiklund, M. E., & Gardner-Bonneau, D. J. (2011). *Handbook of human factors in medical device design*: CRC Press.

Westenskow, D., & Orr, J. (1992). Intelligent Alarms Reduce Anesthesiologist's Response time to Critical Faults. *Journal of Anesthesiology, 77*, 1074-1079.

Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An Introduction to Human factors Engineering*.

Wiklund, M. E., & Hoffman, L. R. (1998). *When electronic devices outnumber flower bouquets in the hospital room.* Paper presented at the Human Factors Society 32nd Annual Meeting, Human Factors Society, Anaheim.

Wilcox, S. (2011). Alarm Design. In M. B. Weinger, M. E. Wiklund & D. J. Gardner-Bonneau (Eds.), *Handbook of Human Factors in medical device design*. Hoboken: Taylor and Francis.

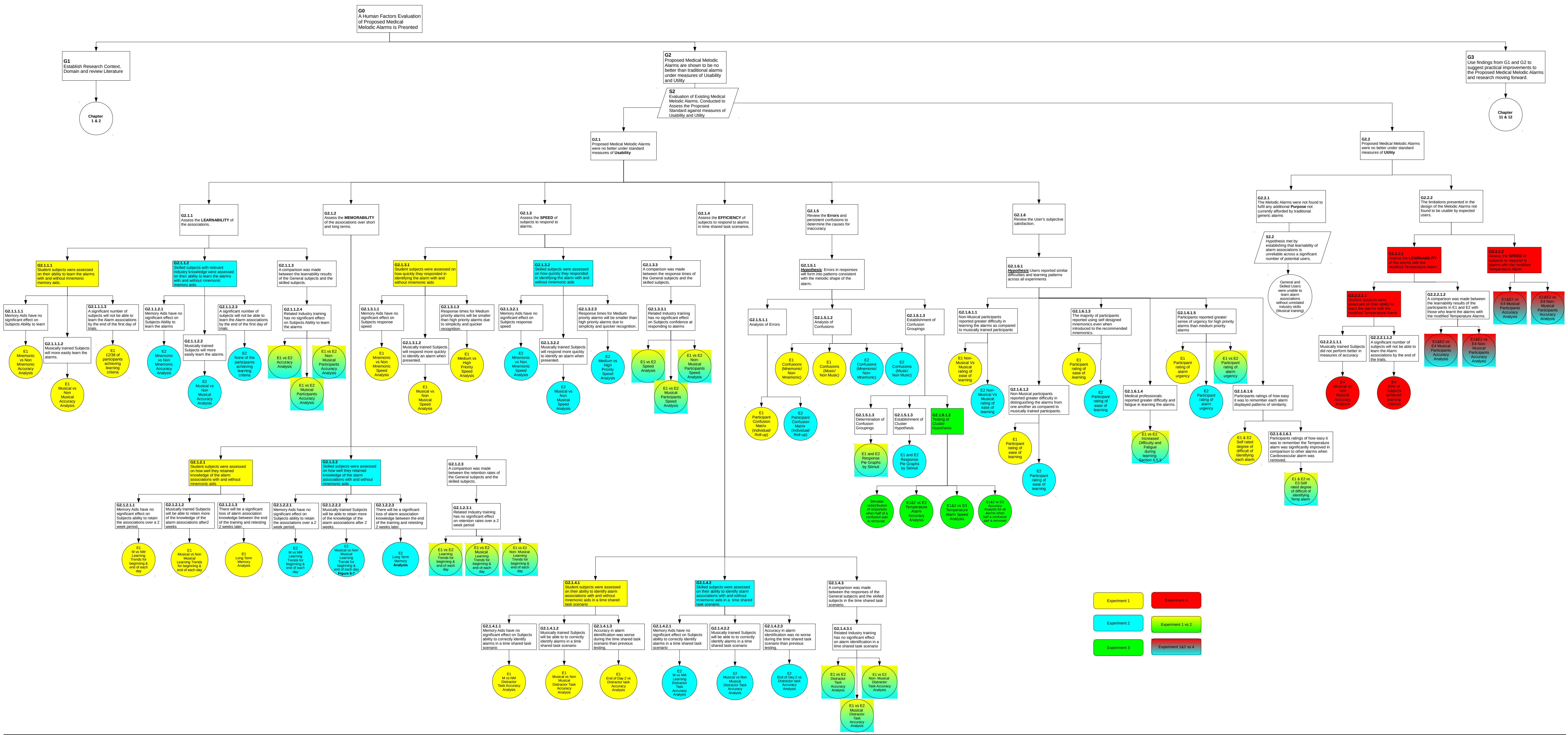
Williams, S., & Beatty, P. (2005). Measureing the performance of audiable alarms for anaesthesia. *Physiological Measurement, 26*, 571-581.

Woods, D. D. (1995). The alarm problem and direct attention in dynamic fault management. *Ergonomics, 38*, 2371-2393.

Xiao, Y. (1994). *Interacting with complext work environments: A field study and a planning model. .* (PhD PhD), University of Toronto, Toronto Canada.

Zwicker, E., & Fastl, H. (1999). *Psychoacoustics - Facts and Models* (2nd Edition ed.): Springer-Verlag BmbH.

APPENDIX A Goal Structured Notation of A Human Factors Evaluation of Medical Melodic Alarms



Full high resolution diagram is available for download at https://dl.dropboxusercontent.com/u/68065118/Experimental%20Program%20GSN.pdf

APPENDIX B Experiment 1 Instructions and Questionnaires

Appendix B.1 Experimenter's protocol (Session 1)

Personnel involved

1. Experimenter (Alexandra Wee) interacting with participant

Resources needed

1. Protocol – "Sonification and Alarms in the Anaesthesia Environment"
2. Paper Work
    a. Information Sheet (on UQ letterhead)
    b. Informed consent form (on UQ letterhead)
    c. Background questionnaire
    d. End of experiment questionnaire
    e. Debriefing Sheet (on UQ letterhead)
3. 1x DELL computer with Alarms program installed
4. Clipboard for experimenter
5. 2 x pens
6. video tapes
7. 2 x chairs
8. Book usability lab and meeting room

Ensure that participant information is filled into headings of all answer sheets.

Conduct of the Sonification and Alarms in the Anaesthesia Environment

Set up room – have two chairs, lights working, tidy etc…

Set up computer and Program

Test that the sound is working. Take sound pressure reading

Label questionnaires with participant code

Seat the participant in front of the computer

"Welcome to the UQ Usability Lab and thanks for taking part in this study.

I'm going to read out loud most of the instruction today. This is so that everyone who takes part gets the same information in the same way. However, please feel free to ask questions as we go along.

In our session today, we will explore the effectiveness of different Alarm sounds to convey information about different danger events that could occur while a patient is under in an operating room.

More details are provided in this information sheet. I suggest that you read the information sheet to understand what will be expected of you during this experiment.

Hand participant the information sheet (let them read it on their own)

Before proceeding further, we need you to read and sign the informed consent form. Please take your time reading it and ask any questions you might have. A copy of the consent form will be made available for you.

Hand participant the informed consent form

Collect the signed consent form

If there are no further questions and you are happy to continue, we now begin the experiment.

Answer any questions

Before we proceed with the experiment, can I also ask you to turn off your mobile phone, your beeping watch or anything else that will make noise or distract you from the experiment."

Wait for the participant to turn things off

"Thank you. We also have a background questionnaire, could you please fill this in now."

Hand participant the background questionnaire

Collect the background questionnaire
Enter participant's group and number in the program and click "Start"

As mentioned in the information sheet this experiment is concerned with alarms in medical equipment. You will be taught to identify 8 different alarm categories in two priorities.

High Priority alarms require immediate operator action whereas *Medium Priority* alarms simply require prompt operator action while

The 8 alarm categories are as follows:

Play each alarm in medium priority as explaining

General Alarms are used in equipment which does not readily fall into any of the other categories, for example electrical or non-oxygen gas supply systems.

This is the medium priority general alarm **<play sound>**

Oxygen Alarms are used in any equipment which measures delivered oxygen to the patient such as pulse oximeters, tissue oxygen monitors, oxygen analysers, oxygen concentrators and oxygen gas supply lines.

This is the medium priority oxygen alarm **<play sound>**

Ventilation Alarms are used in workstations that support the patients breathing such as artificial ventilators, spirometers, CO2 monitors, ventilator disconnect monitors etc.

This is the medium priority ventilation alarm **<play sound>**

Cardiovascular Alarms are used in workstations that monitor the heart. For example cardiac monitors, heart rate monitors, invasive or non-invasive blood pressure monitors.

This is the medium priority cardiovascular alarm **<play sound>**

Temperature Alarms are used in workstations that measure temperature of the patient. For example temperature monitors, neonatal incubators, patient heating or cooling systems, ultrasound, X-ray or MRI systems etc.

This is the medium priority temperature alarm **<play sound>**

Infusion Alarms are used in workstations that deliver fluids intravenously to the patient. For example volumetric infusion pumps, syringe drivers, anaesthetic agent delivery systems or analysers.

This is the medium priority Infusion alarm **<play sound>**

Perfusion Alarms are used in workstations that artificially support the patient’s circulation. For example, Cardio-pulmonary perfusion pump, ventricular assist devices; artificial hearts; renal dialysis systems etc.

This is the medium priority perfusion alarm **<play sound>**

Power Fail Alarms occur when any device experiences loss of power or other major failure of the device.

This is the medium priority power failure alarm **<play sound>**

USE THE DIAGRAM TO ILLUSTRATE THE FOLLOWING

The high priority alarms for each sound are similar to their medium priority counter parts except that they are quicker and have 2 extra notes played after the initial 3 note sequence. This complete sequence is repeated twice.

PLAY THE HIGH PRIORITY SOUNDS

You will now proceed to the learning screen which will be similar to this screen and allowed to learn the alarms at your own pace. After this you will be tested on your ability to name the alarm sounds as they are played to you. Try your best to get as many right as possible in the fastest time. If you cannot remember the answer try to make your best guess. At the end of the test a score card will be displayed on the screen with your correct answers in blue and your incorrect answers in red.

After the test you will be allowed to learn the sounds again. You do not need to listen to all the sounds again but please take this opportunity to review any sounds that you may not be sure about. When you feel that you are ready. Hit the take test button in order to start the next test. You will have to run through this cycle until you have achieved 2 consecutive sets of correct tests or until approximately 35 minutes have passed.

Are there any questions so far?"

Answer questions (if any)

ALLOW PARTICIPANT TO DO THE EXPERIMENT

Finally please fill out this brief questionnaire.

AT END OF EXPERIMENT

Thank you! This is the end of the experiment for today and we thank you for your time. Next week we will be doing more work on these same alarms. If you have any other questions or if you are interested in learning more about the study I will be happy to answer any questions you may have.

If nothing

Then Show participant out… set up for next one

Appendix B.2 Experimenter's protocol (Session 2)

Personnel involved

2. Experimenter (Alexandra Wee) interacting with participant

Resources needed

9. Protocol – "Sonification and Alarms in the Anaesthesia Environment"
10. Paper Work
    a. Information Sheet (on UQ letterhead)
    b. Informed consent form (on UQ letterhead)
    c. Background questionnaire
    d. End of experiment questionnaire
    e. Debriefing Sheet (on UQ letterhead)
11. 1x DELL computer with Alarms program installed
12. Clipboard for experimenter
13. 2 x pens
14. video tapes
15. 2 x chairs
16. Book usability lab and meeting room

Ensure that participant information is filled into headings of all answer sheets.

Conduct of the Sonification and Alarms in the Anaesthesia Environment

Set up room – have two chairs, lights working, tidy etc…

Set up computer and Program

Test that the sound is working. Take sound pressure reading

Label questionnaires with participant code

Seat the participant in front of the computer

Welcome back to the UQ usability lab. Our experiment today is going to once again look at the alarm sounds that you were introduced to last week. There will be three stages to today's experiment.

In the first stage you will be tested on how well you remember the alarm sounds from before.

In the second stage you will once again be allowed to learn the alarm sounds until you reach a certain level of expertise.

In the third stage you will be asked to identify the alarm sounds as a secondary task while at the same time taking part in another primary task.

I will explain each stage in more detail as we get to it.

Stage 1 – 10 mins

First of all we would like to test how much of the alarm sounds you have remembered from last week. You will participate in a series of tests, similar to the test phases that you have experienced before. The only difference is that this time they will not be broken up by learning phases.

Remember we are not testing you but the alarms but please try your best to get as many answers right as possible. If you cannot remember the answer try to make your best guess. At the end of each test a score card will be displayed on the screen with your correct answers in blue and your incorrect answers in red.

You will go through 2 cycles of this phase. At the end of the 2 cycles you will be shown the learning screen again. Stop here and indicate to the experimenter that you are ready for Stage 2

Stage 2 – 30 mins

This stage is similar to experiment 1. This is the learning screen where you will be allowed to review the alarm sounds. Same rules apply, you can listen to them in any order you wish however you can only listen to each alarm once. When you're done listening to the alarms. Click on the **Take Test** Button to take a test. If you complete the test with no mistakes, you will proceed to the next stage. If you made any errors, you will be allowed to learn the sounds again. This cycle continues until you have learnt all the alarms or if you go through it 10 times whichever comes first. When you have finished this stage a screen will appear informing you to stop and wait for further instructions.

Stage 3 – 20 mins

In this stage you will be asked to identify alarms as they are played to you. At the same time you will be asked to complete a mathematics task on the laptop you see in front of you. The math task is a series of simple addition and subtraction equations for which you are required to answer **TRUE** or **FALSE**.

For example

2+2=4 True

4-2=10 False

An indicator on the graph on the same screen plots your accuracy and speed of answering the questions. Please try to keep the indicator as close to the top right hand corner of the graph as possible.

At various intervals throughout the test, an alarm sound will be played. When this occurs please select the button on the touch screen which you think identifies the sound you just heard. Once again please try your best to get as many answers right as possible. If you cannot remember the answer try to make your best guess.

Any questions?

AT END OF EXPERIMENT

Thank you! This is the end of the experiment for today and we thank you for your time. If you have any other questions or if you are interested in learning more about the study I will be happy to answer any questions you may have.

If nothing
Then Show participant out… set up for next one

Appendix B.3 Participant Information Sheet

Study: Alarms in Medical Electrical Equipment

This study has been cleared in accordance with the ethical review processes of the University of Queensland and within the guidelines of the National Health and Medical Research Council. You are, of course, free to discuss your participation with project staff (contactable on 3365 3988 for Professor Penelope Sanderson; or 3365 1636 for Alexandra Wee). If you would like to speak to an officer of the University not involved in the study, you may contact the School of Psychology Ethics Review Officer directly on 3346 9517, or by e-mail: stone@psy.uq.edu.au or contact the University of Queensland Ethics Officer on 3365 3924, e-mail: humanethics@research.uq.edu.au

Overview of the intention of the experiment

In this experiment we will be evaluating a new standard for alarms which will be used in Medical Electrical equipment. We are interested in assessing the alarms for their ability to convey the correct information and whether they are easily distinguishable from each other. Our goal is to test the alarms design, **NOT** to test you or your knowledge. However we do ask that you try to do your best in the evaluations.

You will receive information regarding the alarms, which will involve listening to and knowing some background as to which machines in an operating theatre would emit these alarms.

The experimenter may be taking notes throughout the experiment. However no notes will bear any information by which you may be identified. Your responses to the recall tests at the end of every learning session will be recorded. Both the notes and any recordings will be analysed and archived under lock and key within the UQ Cognitive Engineering Research Laboratory and its confidentiality will be maintained.

In general, you can withdraw from the study at any point without prejudice or penalty of any kind. For instance, if you withdraw early for any reason, you will still be paid for the number of hours you have completed. If you do decide to withdraw early, all of the data already collected from you will be destroyed.

The experiment will run for a maximum of one hour and a half, with opportunities for rest breaks if necessary. There are no foreseeable added risks to you above the risks of everyday living.

Before proceeding further, we need you to read and sign the informed consent form. Please take your time reading it and ask any questions you might have.

If there are no further questions and you are happy to continue, we now begin the experiment.

Appendix B.4 Participant's Consent Form

Study: **Alarms in Medical Electrical Equipment**

Participant's Name: ______________________

Address:

____

Email/Phone: ____

I have been provided with information about the procedure for the **Alarms in Medical Electrical Equipment** study and I am happy to take part. I understand that I will receive the necessary credits for my participation.

I understand that I can withdraw from the study at any point without prejudice or penalty of any kind. A responsible UQ staff member will be in attendance or available nearby during my session.

The handling of my data from this study has been explained to me. No notes or logs will bear any information by which I might be identified. In addition, unless I agree to the Special Release below, only researchers working on this will view any data collected during the study. All material will be analysed and archived under lock and key within the UQ Cognitive Engineering Research Laboratory and its confidentiality will be maintained.

Signature:____ Date: ___________

Experimenter's name: **Alexandra Wee** Date: ___________

Experimenter's signature: _________________________________

Experimenter:

Alexandra Wee

University of Queensland

St Lucia, Queensland, 4065, AUSTRALIA

Tel: +61 7 3365 1636

Email: alexwee@itee.uq.edu.au

Supervisor:

Professor Penelope Sanderson

University of Queensland

St Lucia, Queensland, 4065, AUSTRALIA

Tel: +61 7 3365 3988

Email: psanderson@humanfactors.uq.edu.au

Appendix B.5 Background Questionnaire

Participant Code:_______________	Date: ____/_____/_____	Group:_______________
Experimenter: **Alexandra Wee**		

Age: ________ years

Gender: Male / Female

B1. Have you had your hearing tested? Yes / No

*If **Yes** please specify results of the test:*

____

____

B2. If you have any hearing impediments please describe them here:

____

____

B3. Have you had any medical or physiological training? Yes / No

*If **Yes** please complete the following questions:*

a. Number of years of training: ______ years
b. Describe any professional medical training:

____

____

____

B4. Is English your first language? Yes / No

B5. Do you have any musical training? Yes / No

If Yes, Please Specify (Instrument and Grade if applicable):

____

Appendix B.6 End of Experiment Questionnaire Session 1 (M)

Participant No.: ________________	**Group:** ________________________
Experimenter: AW	**Date:** __________ **Time:** _________

PLEASE CIRCLE YOUR ANSWERS

E1. How easy was it to distinguish the alarm sounds from one another?

1 2 3 4 5 6 7

Difficult Moderate Easy

E2. How easy was it to remember the alarm sounds with their associated meanings?

1 2 3 4 5 6 7

Difficult Moderate Easy

E3. How would you rate the pleasantness of the alarm sounds?

1 2 3 4 5 6 7

Annoying Moderate Pleasant

E4. How would you rate the sense of urgency conveyed by the MEDIUM priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E5. How would you rate the sense of urgency conveyed by the HIGH priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E6. How useful were the Mnemonics in general when helping you to learn the alarm sounds.

1 2 3 4 5 6 7
Not Useful Moderate Very Useful

E7. For which alarms (if any) did you use the supplied Mnemonic to help you. (Please tick)

- ☐ General Alarms
- ☐ Cardiovascular Alarms
- ☐ Infusion Alarms
- ☐ Oxygen Alarms
- ☐ Perfusion Alarms
- ☐ Power Fail Alarms
- ☐ Temperature Alarms
- ☐ Ventilation Alarms

E8. Did you utilise any other learning aids of your own (such as sound or word associations) to remember any or all of the alarm sounds? If so please describe as comprehensively as possible.
Or describe them to the experimenter if you don't really know how to put them in words.

E9. Any suggestions as to how to improve the overall display?

E10. Any other comments?

Appendix B.7 End of Experiment Questionnaire Session 1 (NM)

Participant No.: ________________	**Group:** ______________________
Experimenter: AW	**Date:** _________ **Time:** ________

PLEASE CIRCLE YOUR ANSWERS

E1. How easy was it to distinguish the alarm sounds from one another?

1 2 3 4 5 6 7

Difficult Moderate Easy

E2. How easy was it to remember the alarm sounds with their associated meanings?

1 2 3 4 5 6 7

Difficult Moderate Easy

E3. How would you rate the pleasantness of the alarm sounds?

1 2 3 4 5 6 7

Annoying Moderate Pleasant

E4. How would you rate the sense of urgency conveyed by the MEDIUM priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E5. How would you rate the sense of urgency conveyed by the HIGH priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E6. Did you utilise any learning aids of your own (such as sound or word associations) to remember any or all of the alarm sounds? If so please describe as comprehensively as possible.

Or describe them to the experimenter if you don't really know how to put them in words.

E7. Any suggestions as to how to improve the overall display?

E8. Any other comments?

Appendix B.8 End of Experiment Questionnaire (Day 2)

Participant No.: ________________	**Group:** ______________________
Experimenter: AW	**Date:** _________ **Time:** ________

PLEASE CIRCLE YOUR ANSWERS

E11. How easy was it to distinguish the alarm sounds from one another when performing the math task?

1 2 3 4 5 6 7
Difficult Moderate Easy

E12. How easy was it to remember the alarm sounds with their associated meanings while performing the math task?

1 2 3 4 5 6 7
Difficult Moderate Easy

E13. How would you rate the sense of urgency conveyed by the MEDIUM priority alarm sounds while performing the math task?

1 2 3 4 5 6 7
Not Urgent Moderate Urgent

E14. How would you rate the sense of urgency conveyed by the HIGH priority alarm sounds while performing the math task?

1 2 3 4 5 6 7
Not Urgent Moderate Urgent

E15. Any suggestions as to how to improve the alarm sounds to make them easier to learn/distinguish/remember?

E16. Any other comments?

Appendix B.9 Debriefing Sheet

Study: Alarms in Medical Electrical Equipment

Many thanks for your participation in this study!

The research you have just participated in investigates the effectiveness of Operating Room alarms in conveying the correct information. In particular we are interested in assessing the alarms for their discernability, learnability and memorability as well as whether mnemonics (learning aids) increases a person's ability to learn and remember what each alarm signifies.

This research will enable us to evaluate current alarms and identify their weaknesses and strengths. The results from this experiment will allow us to better recommend whether current alarm standards are sufficient for patient safety or if further improvements are necessary. We will also be able to suggest improved methods of introducing these alarms in training sessions for anesthesia and nursing staff.

If you are interested in the subject you might consult the following paper:

Block, F. E., Rouse, J. D., Hakala, M., and Thompson, C. L. (2000). A proposed new set of alarm sounds which satisfy standards and rationale to encode source information. *Journal of Clinical Monitoring, 16*, 541-546.

APPENDIX C Experiment 2 Instructions and Questionaires

Appendix C.1 Participant Information Sheet

Study: Alarms in Medical Electrical Equipment

This study has been cleared in accordance with the ethical review processes of the University of Queensland and the University of Newcastle and within the guidelines of the National Health and Medical Research Council. You are, of course, free to discuss your participation with project staff (contactable on 0422177840 Alexandra Wee). If you would like to speak to an officer of the University not involved in the study, you may contact the School of Psychology Ethics Review Officer directly on 3346 9517, or by e-mail: stone@psy.uq.edu.au or contact the University of Queensland Ethics Officer on 3365 3924, e-mail: humanethics@research.uq.edu.au

Overview of the intention of the experiment

In this experiment we will be evaluating a new standard for alarms which will be used in Medical Electrical equipment. We are interested in assessing the alarms for their ability to convey the correct information and whether they are easily distinguishable from each other. Our goal is to test the alarms design, **NOT** to test you or your knowledge. However we do ask that you try to do your best in the evaluations.

You will receive information regarding the alarms, which will involve listening to and knowing some background as to which machines in an operating theatre would emit these alarms.

The experimenter may be taking notes throughout the experiment. However no notes will bear any information by which you may be identified. Your responses to the recall tests at the end of every learning session will be recorded. Both the notes and any recordings will be analysed and archived and its confidentiality will be maintained.

In general, you can withdraw from the study at any point without prejudice or penalty of any kind. For instance, if you withdraw early for any reason, you will still be paid for the number of hours you have completed. If you do decide to withdraw early, all of the data already collected from you will be destroyed.

The experiment will run for a maximum of one hour, with opportunities for rest breaks if necessary. There are no foreseeable added risks to you above the risks of everyday living.

Before proceeding further, we need you to read and sign the informed consent form. Please take your time reading it and ask any questions you might have.

If there are no further questions and you are happy to continue, we now begin the experiment.

Appendix C.2 Consent Form

Study: **Alarms in Medical Electrical Equipment**

Participant's Name: ______________________

Address:

____

Email/Phone: ____

I have been provided with information about the procedure for the **Alarms in Medical Electrical Equipment** study and I am happy to take part. I understand that I will receive $40 for my participation.

I understand that I can withdraw from the study at any point without prejudice or penalty of any kind.

The handling of my data from this study has been explained to me. No notes or logs will bear any information by which I might be identified. All material will be analysed and archived under lock and key within the UQ Cognitive Engineering Research Laboratory and its confidentiality will be maintained.

Signature:____ Date: ____/____/_______

Experimenter's name: **Alexandra Wee** Date: ___________

Experimenter's signature: _________________________________

Experimenter:
Alexandra Wee
University of Queensland
St Lucia, Queensland, 4065, AUSTRALIA
Tel: +61 7 3365 1636
Email: alexwee@itee.uq.edu.au

Supervisor:
Professor Penelope Sanderson
University of Queensland
St Lucia, Queensland, 4065, AUSTRALIA
Tel: +61 7 3365 3988
Email: psanderson@humanfactors.uq.edu.au

Appendix C.3 Background Questionnaire

Official Use Only

Participant Code:_______________ Date: ____/_____/_____ Group:_______________
Experimenter: **Alexandra Wee**

Age: ________ years

Gender: Male / Female

B6. Have you had your hearing tested? Yes / No

*If **Yes** please specify results of the test:*

____

____

B7. If you have any hearing impediments please describe them here:

____

____

B8. Are you participating in this experiment after a shift or any other mentally fatiguing activity? Yes / No

B9. How many years have you been in the nursing profession? ______________ years

B10. What is your primary nursing domain? ____________________________________

B11. Is English your first language? Yes / No

B12. Do you have any musical training? Yes / No

If Yes, Please Specify (Instrument, number of years and Grade if applicable):
Instrument: ____
Years playing: ____
Grade: ____
Current frequency of play: ____

Appendix C.4 End of Experiment Questionnaire Day 1 (Mnemonic)

Participant No.: ________________	**Group:** ______________________
Experimenter: AW	**Date:** _________ **Time:** ________

PLEASE INDICATE YOUR ANSWERS WITH AN **X** ON THE BAR

E1. How easy was it to distinguish the alarm sounds from one another?

1 2 3 4 5 6 7

Difficult Moderate Easy

E2. How easy was it to remember the alarm sounds with their associated meanings?

1 2 3 4 5 6 7

Difficult Moderate Easy

E3. How would you rate the pleasantness of the alarm sounds?

1 2 3 4 5 6 7

Annoying Moderate Pleasant

E4. How would you rate the sense of urgency conveyed by the MEDIUM priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E5. How would you rate the sense of urgency conveyed by the HIGH priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

Please Turn over ⟶

E6. How useful were the Mnemonics in general when helping you to learn the alarm sounds.

1	2	3	4	5	6	7
Not Useful			Moderate			Very Useful

E7. For which alarms (if any) did you use the supplied memory aid to help you. (Please tick) (Also circle which one you used)

- ☐ General Alarms (Singing / Musical / General)
- ☐ Cardiovascular Alarms (Singing / Musical / General)
- ☐ Infusion Alarms (Singing / Musical / General)
- ☐ Oxygen Alarms (Singing / Musical / General)
- ☐ Perfusion Alarms (Singing / Musical / General)
- ☐ Power Fail Alarms (Singing / Musical / General)
- ☐ Temperature Alarms (Singing / Musical / General)
- ☐ Ventilation Alarms (Singing / Musical / General)

E8. Did you utilise any other learning aids of your own (such as sound or word associations) to remember any or all of the alarm sounds? If so please describe as comprehensively as possible.
Or describe them to the experimenter if you don't really know how to put them in words.

E9. Any suggestions as to how to improve the overall alarm sounds?

E10. Any other comments?

Appendix C.5 End of Experiment Questionnaire Day 1 (Non-Mnemonic)

Participant No.: ________________	**Group:** ________**NM**______________
Experimenter: AW	**Date:** _________ **Time:** ________

PLEASE INDICATE YOUR ANSWERS WITH AN **X** ON THE BAR

E1. How easy was it to distinguish the alarm sounds from one another?

1 2 3 4 5 6 7

Difficult Moderate Easy

E2. How easy was it to remember the alarm sounds with their associated meanings?

1 2 3 4 5 6 7

Difficult Moderate Easy

E3. How would you rate the pleasantness of the alarm sounds?

1 2 3 4 5 6 7

Annoying Moderate Pleasant

E4. How would you rate the sense of urgency conveyed by the MEDIUM priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E5. How would you rate the sense of urgency conveyed by the HIGH priority alarm sounds?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E6. Did you utilise any learning aids of your own (such as sound or word associations) to remember any or all of the alarm sounds? If so please describe as comprehensively as possible.
Or describe them to the experimenter if you don't really know how to put them in words.

E7. Any suggestions as to how to improve the overall alarm sounds?

E8. Any other comments?

Appendix C.6 End of Experiment Questionnaire (Day 2)

Participant No.: ________________	**Group:** _______________________
Experimenter: AW	**Date:** _________ **Time:** ________

PLEASE INDICATE YOUR ANSWERS WITH AN **X** ON THE BAR

E1. How easy was it to distinguish the alarm sounds from one another when performing the math task?

1 2 3 4 5 6 7

Difficult Moderate Easy

E2. How easy was it to remember the alarm sounds with their associated meanings while performing the math task?

1 2 3 4 5 6 7

Difficult Moderate Easy

E3. How would you rate the sense of urgency conveyed by the MEDIUM priority alarm sounds while performing the math task?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E4. How would you rate the sense of urgency conveyed by the HIGH priority alarm sounds while performing the math task?

1 2 3 4 5 6 7

Not Urgent Moderate Urgent

E5. Any suggestions as to how to improve the alarm sounds to make them easier to learn/distinguish/remember?

E6. Any other comments?

APPENDIX D Questionnaires for Experiments 3 and 4

Appendix D.1 Participant's Consent Form

A. Prof. Scott Brown
School of Psychology/ Faculty of Science & IT
Aviation Building
The University of Newcastle
University Drive, Callaghan, NSW, 2308
Ph: (02) 49215760
Email: Scott.Brown@newcastle.edu.au

Consent Form for the Research Project:
Learning in Absolute Identification
Alexandra Morey, Pennie Dodds, Babette Rae, Peter Cassey, Daniel Conway & Guy Hawkins

Document Version 15; dated 24/01/11

I agree to participate in the above research project and give my consent freely.

I understand that the project will be conducted as described in the Information Statement, a copy of which I have retained.

I understand I can withdraw from the project at any time and do not have to give any reason for withdrawing.

I consent to participating in experiments as described in the accompanying information sheet

I understand that my personal information will remain confidential to the researchers.

I have had the opportunity to have questions answered to my satisfaction.

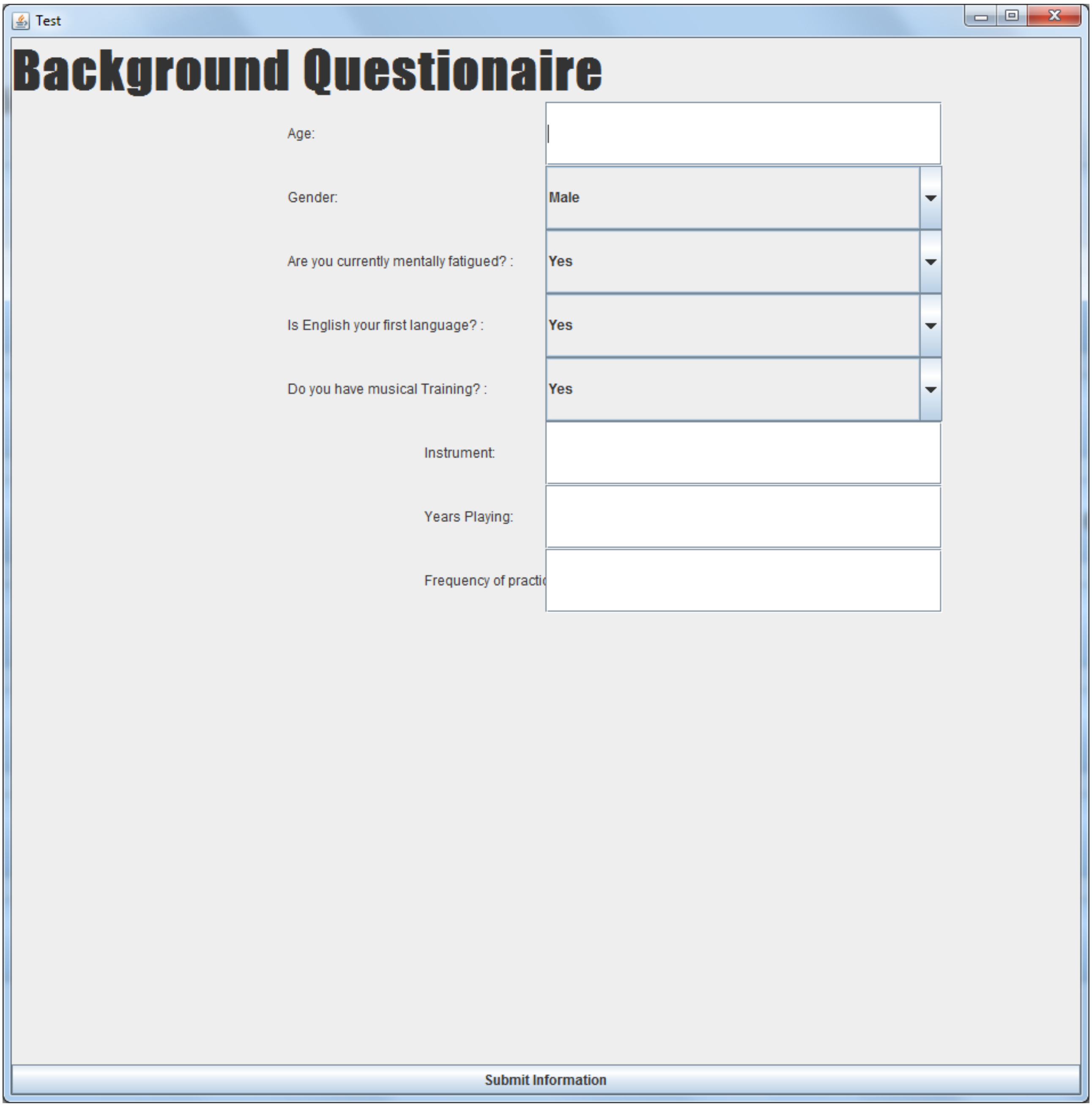
Print Name: ____

Signature: ________________ **Date:** __________________

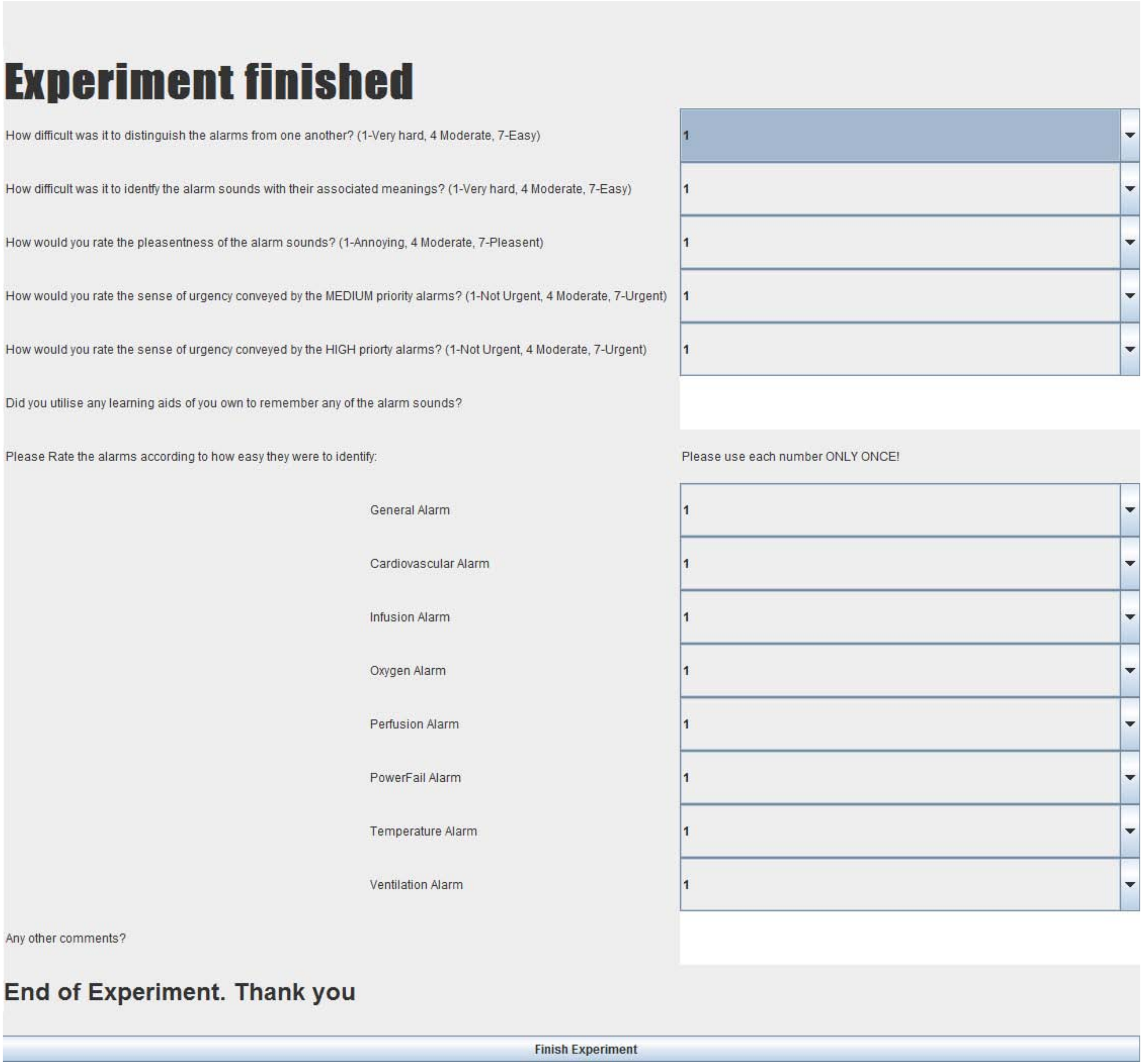
Enquiries about this study may be directed to A. Prof. Scott Brown, School of Psychology, Faculty of Science and Information Technology, The University of Newcastle, telephone: (02) 492 15760

Complaints about this research
This project has been approved by the University's Human Research Ethics Committee, Approval No. H-440-0407. Should you have concerns about your rights as a participant in this research, or you have a complaint about the manner in which the research is conducted, it may be given to the researcher, or, if an independent person is preferred, to the Human Research Ethics Officer, Research Office, The Chancellery, The University of Newcastle, University Drive, Callaghan NSW 2308, Australia, telephone (02 49216333, email Human-Ethics@newcastle.edu.au.

Appendix D.2 Background Questionnaire



Appendix D.3 End of Experiment Questionnaire



Experiment finished

How difficult was it to distinguish the alarms from one another? (1-Very hard, 4 Moderate, 7-Easy)	1
How difficult was it to identfy the alarm sounds with their associated meanings? (1-Very hard, 4 Moderate, 7-Easy)	1
How would you rate the pleasentness of the alarm sounds? (1-Annoying, 4 Moderate, 7-Pleasent)	1
How would you rate the sense of urgency conveyed by the MEDIUM priority alarms? (1-Not Urgent, 4 Moderate, 7-Urgent)	1
How would you rate the sense of urgency conveyed by the HIGH priorty alarms? (1-Not Urgent, 4 Moderate, 7-Urgent)	1
Did you utilise any learning aids of you own to remember any of the alarm sounds?	
Please Rate the alarms according to how easy they were to identify:	Please use each number ONLY ONCE!
General Alarm	1
Cardiovascular Alarm	1
Infusion Alarm	1
Oxygen Alarm	1
Perfusion Alarm	1
PowerFail Alarm	1
Temperature Alarm	1
Ventilation Alarm	1
Any other comments?	

End of Experiment. Thank you

Finish Experiment

APPENDIX E Experiment 1 Data

Appendix E.1 Confusion Tables Day 1

Mnemonic Participants

Subject	Real Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
		Response								
1	CV	10		1				1		12
	GE		12							12
	IN	2		8					2	12
	OX				12					12
	PE					11	1			12
	PF						12			12
	TE							12		12
	VE								12	12
1 Total		12	12	9	12	11	13	13	14	96
2	CV	13		2	1		1	7		24
	GE		24							24
	IN			11	5		2		6	24
	OX			4	14		6			24
	PE			2		12	2		8	24
	PF		1	4			18	1		24
	TE	11				2		11		24
	VE	1		8	4			1	10	24
2 Total		25	25	31	24	14	29	20	24	192
3	CV	6	1		1	2	2	3	1	16
	GE	3	6	2	1	3	1			16
	IN			12	1		1		2	16
	OX			8	4	1		1	2	16
	PE	2	3	3		6	1		1	16
	PF	1		1		1	10		3	16
	TE	8			2			6		16
	VE	1		5	3	1	2	1	3	16
3 Total		21	10	31	12	14	17	11	12	128
4	CV	12						4		16
	GE		16							16
	IN			15					1	16
	OX				16					16
	PE					16				16
	PF						16			16
	TE	1						15		16
	VE					1			15	16
4 Total		13	16	15	16	17	16	19	16	128
5	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4

5 Total		4	4	4	4	4	4	4	4	32
6	CV	2		3	5	3		8	3	24
	GE		24							24
	IN	3		5	3	6	2		5	24
	OX	3		3	6	2	5	1	4	24
	PE	1		5	1	5	8		4	24
	PF	1		3		5	13		2	24
	TE	4		1	3	3	1	12		24
	VE	3		5	9	2	1		4	24
6 Total		17	24	25	27	26	30	21	22	192
7	CV	8			3	2		7	6	26
	GE		24				2			26
	IN	2		5	10			1	8	26
	OX	4		12	5	2		2	1	26
	PE	2				17	3	1	3	26
	PF					1	22	2	1	26
	TE	6		2	2	5		8	3	26
	VE	6		3	6	2		4	5	26
7 Total		28	24	22	26	29	27	25	27	208
8	CV	12	1					3		16
	GE		15				1			16
	IN			15					1	16
	OX				15				1	16
	PE					16				16
	PF			1		1	14			16
	TE							16		16
	VE	1		2					13	16
8 Total		13	16	18	15	17	15	19	15	128
9	CV	8						4		12
	GE		12							12
	IN			10		1			1	12
	OX				9	1		1	1	12
	PE				1	8			3	12
	PF					1	10		1	12
	TE	2		1				9		12
	VE				1	3			8	12
9 Total		10	12	11	11	14	10	14	14	96
10	CV	15			1	1		1	2	20
	GE		20							20
	IN			14		2	1	1	2	20
	OX				18	1		1		20
	PE					17		2	1	20
	PF		1				18	1		20
	TE	2				2		15	1	20
	VE			2		1		1	16	20
10 Total		17	21	16	19	24	19	22	22	160
11	CV	14					1	1		16
	GE		16							16
	IN			14			1		1	16
	OX				16					16
	PE			1		15				16
	PF						16			16
	TE	1						15		16

	VE			1					15	16
11 Total		15	16	16	16	15	18	16	16	128
12	CV	7		1	3			3		14
	GE		14							14
	IN			6		1	1		6	14
	OX			2	9		1	2		14
	PE			7		6			1	14
	PF						14			14
	TE	6			1			7		14
	VE	3		5					6	14
12 Total		16	14	21	13	7	16	12	13	112
13	CV	6						4		10
	GE		10							10
	IN			5			2		3	10
	OX				8				2	10
	PE					7	1		2	10
	PF			1	1	1	6		1	10
	TE	4						6		10
	VE	1			1	4		1	3	10
13 Total		11	10	6	10	12	9	11	11	80
14	CV	11			1			2		14
	GE		14							14
	IN	1		11		2				14
	OX				12			1	1	14
	PE			2		12				14
	PF					1	13			14
	TE	2				1		11		14
	VE	1		1		4			8	14
14 Total		15	14	14	13	20	13	14	9	112
15	CV	16						6		22
	GE		22							22
	IN			20			1		1	22
	OX			2	20					22
	PE			4		13			5	22
	PF			2			20			22
	TE	6						16		22
	VE			4	1	4			13	22
15 Total		22	22	32	21	17	21	22	19	176
16	CV	14		2		1	1	10	2	30
	GE	1	16	1	2	4	6			30
	IN	3	1	7	5		6	1	7	30
	OX	2		5	12	1	7		3	30
	PE	2	2	8	2	7	1		8	30
	PF	1	1	7	7		8	2	4	30
	TE	7		2	3			14	4	30
	VE	4	1	7	1	3	2	4	8	30
16 Total		34	21	39	32	16	31	31	36	240
17	CV	16			1			7		24
	GE		24							24
	IN	5		14		1		1	3	24
	OX				24					24
	PE					24				24
	PF						23		1	24

	TE	2						22		24
	VE					1		2	21	24
17 Total		23	24	14	25	26	23	32	25	192
18	CV	19				1		6		26
	GE		25		1					26
	IN			22	1	3				26
	OX				25			1		26
	PE			1		25				26
	PF			1			25			26
	TE	7						19		26
	VE			2		4			20	26
18 Total		26	25	26	27	33	25	26	20	208
19	CV	16		1	2			7	4	30
	GE		30							30
	IN	3		9	4	1	4	4	5	30
	OX	2		2	18			5	3	30
	PE	1			1	25	3			30
	PF	1	1		1	1	20	3	3	30
	TE	6		2	1	2	1	14	4	30
	VE	5		15	2			1	7	30
19 Total		34	31	29	29	29	28	34	26	240

Non-Mnemonic Participants

Subject	Real Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total
1	CV	4			4	1	1	2	2	14
	GE		14							14
	IN	4		5		2		1	2	14
	OX	1		1	9	1	1		1	14
	PE	1	1	1		8	1		2	14
	PF	3			1	3	7			14
	TE	1		1	1	1		10		14
	VE	2		4	2	1			5	14
1 Total		16	15	12	17	17	10	13	12	112
2	CV	26		1	1			4		32
	GE		32							32
	IN	2		27	1	1			1	32
	OX				27	3	1	1		32
	PE					19	9	3	1	32
	PF					3	19	2	8	32
	TE	13						19		32
	VE	1			2	8	2	2	17	32
2 Total		42	32	28	31	34	31	31	27	256
3	CV	3	1	2	1		4		1	12
	GE		9	2	1					12
	IN			4	1	2	2	2	1	12
	OX		1	1	1	1	1	3	4	12
	PE			3	3	3			3	12
	PF	3	2	1		2	1	1	2	12

	TE	1	1	2	2		1	5		12
	VE	2		1	3	2		3	1	12
3 Total		9	14	16	12	10	9	14	12	96
4	CV	6		2	1	1	2	2	4	18
	GE		18							18
	IN	5		2	1	2	1	3	4	18
	OX	3		4	4				7	18
	PE			2	2	10	2	1	1	18
	PF	1		1		4	7	4	1	18
	TE	4			1		2	9	2	18
	VE	4		3	1		1	1	8	18
4 Total		23	18	14	10	17	15	20	27	144
5	CV	3		4	2	1	3	8	1	22
	GE	3	17	1		1				22
	IN	3	1	3	2	2	6	1	4	22
	OX	2		6	9	2	1		2	22
	PE	2	3	3	2	3	2	5	2	22
	PF	4	2	4	1	3	5	1	2	22
	TE	4	2	2	1	2	2	4	5	22
	VE	2		6	3	3	3	3	2	22
5 Total		23	25	29	20	17	22	22	18	176
6	CV	22							2	24
	GE		24							24
	IN	1		18		1			4	24
	OX	2			22					24
	PE			3	1	20				24
	PF						18	6		24
	TE	1				1		19	3	24
	VE			3	1	4			16	24
6 Total		26	24	24	24	26	18	25	25	192
7	CV	5			1					6
	GE		6							6
	IN	1		5						6
	OX				5	1				6
	PE					5	1			6
	PF			1		1	4			6
	TE							6		6
	VE								6	6
7 Total		6	6	6	6	7	5	6	6	48
8	CV	10			2			2		14
	GE		14							14
	IN			6		5	1		2	14
	OX			3	7		3	1		14
	PE			4		9	1			14
	PF			2	1		8	1	2	14
	TE	3						11		14
	VE			2	3	1		1	7	14
8 Total		13	14	17	13	15	13	16	11	112
9	CV	13			1			10	2	26
	GE		26							26
	IN			17		2	3	1	3	26
	OX				25				1	26
	PE			2		20	1		3	26

	PF			4		2	19		1	26
	TE	2			1			22	1	26
	VE	1		4		1	1		19	26
9 Total		16	26	27	27	25	24	33	30	208
10	CV	11			1			5	3	20
	GE		20							20
	IN			3		4	2	4	7	20
	OX	1			16	1		2		20
	PE	1	1	2		6	3	2	5	20
	PF			4	1	4	6	4	1	20
	TE	12				2		6		20
	VE	2		3		6			9	20
10 Total		27	21	12	18	23	11	23	25	160
11	CV	16			1			1		18
	GE		18							18
	IN			12	2	3			1	18
	OX				8	3	4		3	18
	PE			2	4	10	1		1	18
	PF					4	10		4	18
	TE							18		18
	VE			2	2	2			12	18
11 Total		16	18	16	17	22	15	19	21	144
13	CV	15			1			2		18
	GE		17				1			18
	IN	1		6	2	2	2	1	4	18
	OX	1		1	16					18
	PE		1	1		10	1	1	4	18
	PF			2		2	12		2	18
	TE	12				2		4		18
	VE	1		5		6		1	5	18
13 Total		30	18	15	19	22	16	9	15	144
14	CV	6								6
	GE		6							6
	IN			5			1			6
	OX				6					6
	PE					6				6
	PF						6			6
	TE							6		6
	VE			2					4	6
14 Total		6	6	7	6	6	7	6	4	48
15	CV	10						7	1	18
	GE		18							18
	IN			15		2	1			18
	OX				18					18
	PE					14	1		3	18
	PF						18			18
	TE	4						13	1	18
	VE			2		5			11	18
15 Total		14	18	17	18	21	20	20	16	144
16	CV	10								10
	GE		10							10
	IN	2		7		1				10
	OX				10					10

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
	PE					9			1	10
	PF			1			9			10
	TE							10		10
	VE								10	10
16 Total		12	10	8	10	10	9	10	11	80
17	CV	9		1	2	1		4	1	18
	GE		17			1				18
	IN			6	5	1	1	3	2	18
	OX	2			12			4		18
	PE			4		10	4			18
	PF	1		1		2	13		1	18
	TE	11		1			2	4		18
	VE	3		4	2	2		1	6	18
17 Total		26	17	17	21	17	20	16	10	144
18	CV	5						5		10
	GE		10							10
	IN			5		2	1	1	1	10
	OX			1	8		1			10
	PE			2		3			5	10
	PF						8	1	1	10
	TE	1						9		10
	VE			1		5			4	10
18 Total		6	10	9	8	10	10	16	11	80
19	CV	14						3	1	18
	GE		18							18
	IN			12	1	3	2			18
	OX	2		2	11				3	18
	PE			1		13	4			18
	PF			1	2	2	13			18
	TE							18		18
	VE	1			2				15	18
19 Total		17	18	16	16	18	19	21	19	144
20	CV	31	2					1		34
	GE		34							34
	IN			27	1	2	1	1	2	34
	OX	1			27	2	3	1		34
	PE			1	2	25	1	1	4	34
	PF				1		33			34
	TE	24						10		34
	VE	1		27	2	2			2	34
20 Total		57	36	55	33	31	38	14	8	272

Appendix E.2 Confusion Tables Day 2 (Long Term Memory Test)

Mnemonic Participants		Response								
Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
1	CV	4								4
	GE		4							4

	IN			4						4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE							4		4
	VE			1		2			1	4
1 Total		4	4	5	4	5	4	4	2	32
2	CV	2		1		1				4
	GE		4							4
	IN			1	1		1	1		4
	OX				4					4
	PE					2			2	4
	PF						4			4
	TE							4		4
	VE			2	1				1	4
2 Total		2	4	4	6	3	5	5	3	32
4	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
4 Total		4	4	4	4	4	4	4	4	32
5	CV	4								4
	GE		4							4
	IN	1		3						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE					1			3	4
5 Total		5	4	3	4	5	4	4	3	32
6	CV				2			1	1	4
	GE		4							4
	IN			1	1	1	1			4
	OX	1					1		2	4
	PE			1	2		1			4
	PF			1		2	1			4
	TE				1			3		4
	VE	1				1			2	4
6 Total		2	4	3	6	4	4	4	5	32
7	CV	2						1	1	4
	GE		4							4
	IN			2	1				1	4
	OX			1	2			1		4
	PE	1				1		2		4
	PF						4			4
	TE	3				1				4
	VE				2				2	4
7 Total		6	4	3	5	2	4	4	4	32
8	CV	4								4

	GE		4							4
	IN					1			3	4
	OX				4					4
	PE			3					1	4
	PF						4			4
	TE							4		4
	VE					1	3			4
8 Total		4	4	3	4	2	7	4	4	32
9	CV	4								4
	GE		4							4
	IN			3					1	4
	OX	1			3					4
	PE					4				4
	PF						4			4
	TE				1			3		4
	VE					1			3	4
9 Total		5	4	3	4	5	4	3	4	32
10	CV	1						2	1	4
	GE		4							4
	IN			2		2				4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							3	1	4
	VE			3		1				4
10 Total		1	4	5	4	7	4	5	2	32
11	CV	3						1		4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE			1		1			2	4
11 Total		4	4	5	4	5	4	4	2	32
12	CV	2			1			1		4
	GE		4							4
	IN			2		1		1		4
	OX				3			1		4
	PE			1		1			2	4
	PF						4			4
	TE	2			1			1		4
	VE	1							3	4
12 Total		5	4	3	5	2	4	4	5	32
13	CV	4								4
	GE		4							4
	IN			1			2	1		4
	OX			1	2				1	4
	PE				1				3	4
	PF			2		1	1			4
	TE	2						2		4
	VE					2	1		1	4
13 Total		6	4	4	3	3	4	3	5	32

14	CV	2						2		4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE			3					1	4
14 Total		2	4	6	4	4	4	6	2	32
15	CV	3						1		4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE				1			3		4
	VE							1	3	4
15 Total		3	4	3	5	4	4	5	4	32
16	CV	1		1				2		4
	GE		4							4
	IN			1		2			1	4
	OX	1					3			4
	PE					3			1	4
	PF	1		1			2			4
	TE	1			1			1	1	4
	VE	1		1	2					4
16 Total		5	4	4	3	5	5	3	3	32
17	CV	2		1				1		4
	GE		4							4
	IN	4								4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE							1	3	4
17 Total		6	4	1	4	4	4	6	3	32
18	CV	3						1		4
	GE		4							4
	IN			2		2				4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
18 Total		4	4	2	4	6	4	4	4	32
19	CV	4								4
	GE		2						2	4
	IN			1	1	1	1			4
	OX			2			1		1	4
	PE					3	1			4
	PF			2		1	1			4
	TE							4		4
	VE	1			2				1	4

19 Total		5	2	5	3	5	4	4	4	32

Non-Mnemonic Participants

Subject	R Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total
2	CV	3						1		4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
2 Total		4	4	4	4	4	4	4	4	32
3	CV			2			2			4
	GE		4							4
	IN			1	1				2	4
	OX				1	1		1	1	4
	PE	1			1	1	1			4
	PF	2		1		1				4
	TE		2	1					1	4
	VE	1				1	2			4
3 Total		4	6	5	3	4	5	1	4	32
4	CV			2	1		1			4
	GE		4							4
	IN				2				2	4
	OX			3					1	4
	PE						2	2		4
	PF					4				4
	TE	2						2		4
	VE	2			1				1	4
4 Total		4	4	5	4	4	3	4	4	32
5	CV	1		1		2				4
	GE		3						1	4
	IN				1	1	2			4
	OX	1				1		1	1	4
	PE		1			1	1	1		4
	PF	1		1	1			1		4
	TE	1		2			1			4
	VE	1			1	1	1			4
5 Total		5	4	4	3	6	5	3	2	32
6	CV	4								4
	GE		4							4
	IN			1					3	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							3	1	4
	VE								4	4
6 Total		4	4	1	4	4	4	3	8	32
7	CV	4								4
	GE		4							4

	IN	1		3						4
	OX				2			2		4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
7 Total		6	4	3	2	4	4	5	4	32
9	CV	4								4
	GE	1	3							4
	IN			3					1	4
	OX				4					4
	PE					3			1	4
	PF		1				3			4
	TE	1						3		4
	VE			1					3	4
9 Total		6	4	4	4	3	3	3	5	32
10	CV	2				1		1		4
	GE		4							4
	IN			1		1			2	4
	OX				4					4
	PE		1			1	1		1	4
	PF						2	2		4
	TE	4								4
	VE							2	2	4
10 Total		6	5	1	4	3	3	5	5	32
13	CV	2					1	1		4
	GE		4							4
	IN			3		1				4
	OX				4					4
	PE				1	2			1	4
	PF						4			4
	TE	4								4
	VE				1	1	1		1	4
13 Total		6	4	3	6	4	6	1	2	32
14	CV	4								4
	GE		4							4
	IN			3		1				4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
14 Total		4	4	3	4	5	4	4	4	32
15	CV	3						1		4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					2			2	4
	PF						4			4
	TE							4		4
	VE								4	4
15 Total		3	4	4	4	2	4	5	6	32
16	CV	4								4

	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE			1					3	4
16 Total		4	4	5	4	4	4	4	3	32
17	CV	2						2		4
	GE		4							4
	IN					2			2	4
	OX				4					4
	PE			1		3				4
	PF						4			4
	TE	2		1				1		4
	VE			2					2	4
17 Total		4	4	4	4	5	4	3	4	32
18	CV	3			1					4
	GE		4							4
	IN	1		3						4
	OX				2			1	1	4
	PE					3		1		4
	PF						4			4
	TE				1			2	1	4
	VE			2					2	4
18 Total		4	4	5	4	3	4	4	4	32
20	CV	4								4
	GE		2						2	4
	IN			3		1				4
	OX				4					4
	PE			1		2		1		4
	PF						4			4
	TE	4								4
	VE			4						4
20 Total		8	2	8	4	3	4	1	2	32

Appendix E.3 Confusion Tables Day 2 (Test)

Mnemonic Participants

Subject	R Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total
1	CV	6								6
	GE		6							6
	IN			6						6
	OX				6					6
	PE					6				6
	PF					1	5			6
	TE							6		6
	VE								6	6
1 Total		6	6	6	6	7	5	6	6	48
2	CV	8		1				5		14
	GE		14							14
	IN			9	3	1	1			14
	OX			4	9	1				14
	PE			1		9	1		3	14
	PF			1			13			14
	TE		1			2		11		14
	VE			7	1				6	14
2 Total		8	15	23	13	13	15	16	9	112
4	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
4 Total		4	4	4	4	4	4	4	4	32
5	CV	8						6		14
	GE		14							14
	IN			14						14
	OX				14					14
	PE					14				14
	PF						14			14
	TE	2						12		14
	VE						1		13	14
5 Total		10	14	14	14	14	15	18	13	112
6	CV	7			1	1		4	1	14
	GE		14							14
	IN	3		4	1	2	3		1	14
	OX	3		1	6	1	3			14
	PE			1	1	6	4		2	14
	PF			2	1	4	6		1	14
	TE	2		3			1	7	1	14
	VE	1		7	1	1			4	14
6 Total		16	14	18	11	15	17	11	10	112
7	CV	4		3		1		5	1	14
	GE		14							14

	IN	1		9	2				2	14
	OX			5	9					14
	PE	2				11			1	14
	PF						14			14
	TE	1		2	1	1		8	1	14
	VE	3		2	1	3		1	4	14
7 Total		11	14	21	13	16	14	14	9	112
8	CV	14								14
	GE		14							14
	IN			12					2	14
	OX				14					14
	PE			2		7			5	14
	PF						14			14
	TE							14		14
	VE			2					12	14
8 Total		14	14	16	14	7	14	14	19	112
9	CV	12						2		14
	GE		14							14
	IN			14						14
	OX				11			3		14
	PE					14				14
	PF						14			14
	TE							14		14
	VE								14	14
9 Total		12	14	14	11	14	14	19	14	112
10	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
10 Total		4	4	4	4	4	4	4	4	32
11	CV	12						2		14
	GE		14							14
	IN			13			1			14
	OX				14					14
	PE					14				14
	PF			1			13			14
	TE							14		14
	VE					2			12	14
11 Total		12	14	14	14	16	14	16	12	112
12	CV	12						2		14
	GE		14							14
	IN	1		8		1			4	14
	OX				14					14
	PE			1		12			1	14
	PF						14			14
	TE	1						13		14
	VE	1		5					8	14
12 Total		15	14	14	14	13	14	15	13	112
13	CV	10						4		14

	GE		14							14
	IN	1		8	3		1		1	14
	OX				13				1	14
	PE					13			1	14
	PF			1			13			14
	TE	1						13		14
	VE	1		1		2			10	14
13 Total		13	14	10	16	15	14	17	13	112
14	CV	5						1		6
	GE		6							6
	IN			6						6
	OX				6					6
	PE					6				6
	PF						6			6
	TE							6		6
	VE								6	6
14 Total		5	6	6	6	6	6	7	6	48
15	CV	12						2		14
	GE		14							14
	IN			14						14
	OX				14					14
	PE			2		7			5	14
	PF			2			12			14
	TE	5						9		14
	VE			5		2			7	14
15 Total		17	14	23	14	9	12	11	12	112
16	CV	8				1		5		14
	GE		12				2			14
	IN	3		2	1	1	1	1	5	14
	OX			1	11				2	14
	PE			3	2	2	4	1	2	14
	PF			6	4	1	1		2	14
	TE	5		1				7	1	14
	VE	2		2		3		3	4	14
16 Total		18	12	15	18	8	8	17	16	112
17	CV	12						2		14
	GE		14							14
	IN	2		7				2	3	14
	OX				14					14
	PE					14				14
	PF						14			14
	TE	2						12		14
	VE								14	14
17 Total		16	14	7	14	14	14	16	17	112
18	CV	11						3		14
	GE		14							14
	IN			13		1				14
	OX				14					14
	PE					14				14
	PF						14			14
	TE	2						12		14
	VE								14	14
18 Total		13	14	13	14	15	14	15	14	112

19	CV	14								14
	GE		14							14
	IN			11		3				14
	OX				14					14
	PE			1		10	3			14
	PF			1		4	9			14
	TE	1						13		14
	VE				2				12	14
19 Total		15	14	13	16	17	12	13	12	112

Non-Mnemonic Participants		Response								
2	CV	9						1		10
	GE		10							10
	IN			10						10
	OX				10					10
	PE					10				10
	PF					1	9			10
	TE	1						9		10
	VE								10	10
2 Total		10	10	10	10	11	9	10	10	80
3	CV	5		6	2		1	4	2	20
	GE	1	17		1			1		20
	IN	5	1	4	2	2		3	3	20
	OX	3		2	2	2	4	4	3	20
	PE	4		3	2	1	5	4	1	20
	PF	7	2	2		4	3	1	1	20
	TE	5		2	3	3	2	3	2	20
	VE		1	2	1	4	6	6		20
3 Total		30	21	21	13	16	21	26	12	160
4	CV	7		3	1			2	1	14
	GE		14							14
	IN	1		5	1		2	1	4	14
	OX	3		3	4			1	3	14
	PE				1	12	1			14
	PF					1	13			14
	TE			2	1			10	1	14
	VE	2		3	1	2	1		5	14
4 Total		13	14	16	9	15	17	14	14	112
5	CV	5		2		2		4	1	14
	GE		13	1						14
	IN			1	5	1	2	3	2	14
	OX			2	5	2	1	3	1	14
	PE	5		1	1	4	2	1		14
	PF	1	1	1		3	6		2	14
	TE	5		5	1	1		2		14
	VE	3		3	1	3		4		14
5 Total		19	14	16	13	16	11	17	6	112
6	CV	10								10
	GE		10							10

	IN			7			1		2	10
	OX				10					10
	PE			1		9				10
	PF						10			10
	TE							10		10
	VE			1					9	10
6 Total		10	10	9	10	9	11	10	11	80
7	CV	14								14
	GE		14							14
	IN			14						14
	OX				10			4		14
	PE					14				14
	PF		1			1	12			14
	TE							14		14
	VE			1					13	14
7 Total		14	15	15	10	15	12	18	13	112
9	CV	11						3		14
	GE		14							14
	IN			13					1	14
	OX				14					14
	PE					13	1			14
	PF						14			14
	TE	5						9		14
	VE								14	14
9 Total		16	14	13	14	13	15	12	15	112
10	CV	12						2		14
	GE		14							14
	IN			7		1	1	1	4	14
	OX	1			13					14
	PE					5	2		7	14
	PF			1		1	11	1		14
	TE	3						11		14
	VE			3		1			10	14
10 Total		16	14	11	13	8	14	15	21	112
13	CV	8						6		14
	GE		14							14
	IN	4		4		1		1	4	14
	OX				14					14
	PE		1			10	1		2	14
	PF						14			14
	TE	7						7		14
	VE			1	2	3			8	14
13 Total		19	15	5	16	14	15	14	14	112
14	CV	10								10
	GE		10							10
	IN			10						10
	OX				10					10
	PE					9	1			10
	PF						10			10
	TE							10		10
	VE					1			9	10
14 Total		10	10	10	10	10	11	10	9	80
15	CV	13						1		14

	GE		14							14
	IN			14						14
	OX				14					14
	PE					14				14
	PF						14			14
	TE	2						12		14
	VE					2			12	14
15 Total		15	14	14	14	16	14	13	12	112
16	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
16 Total		4	4	4	4	4	4	4	4	32
17	CV	7		1				6		14
	GE		14							14
	IN	1		10		2			1	14
	OX	2			12					14
	PE					12	2			14
	PF					1	13			14
	TE	2				1		11		14
	VE	1		5		2			6	14
17 Total		13	14	16	12	18	15	17	7	112
18	CV	10			2			2		14
	GE		14							14
	IN			6		5	1		2	14
	OX				14					14
	PE			8		5			1	14
	PF						13		1	14
	TE	2	1					11		14
	VE			1		6			7	14
18 Total		12	15	15	16	16	14	13	11	112
20	CV	12						2		14
	GE		13						1	14
	IN			14						14
	OX				14					14
	PE					14				14
	PF						14			14
	TE	12						2		14
	VE	1		5					8	14
20 Total		25	13	19	14	14	14	4	9	112

Appendix E.4 Confusion Tables Distracter Task

Mnemonic
Participants Response

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
1	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
1 Total		5	4	4	4	4	4	3	4	32
2	CV	2				1		1		4
	GE		4							4
	IN			3	1					4
	OX				3				1	4
	PE					3			1	4
	PF			1	3					4
	TE	1						3		4
	VE			1			1		2	4
2 Total		3	4	5	7	4	1	4	4	32
4	CV	3						1		4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
4 Total		5	4	4	4	4	4	3	4	32
6	CV	2						1	1	4
	GE		4							4
	IN			2			1		1	4
	OX				4					4
	PE				1	1	1		1	4
	PF			1		1	1		1	4
	TE	1						2	1	4
	VE			1			1		2	4
6 Total		3	4	4	5	2	4	3	7	32
7	CV			1	1			2		4
	GE		4							4
	IN			2				1	1	4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE	2						2		4
	VE	2		1			1			4

7 Total		4	4	4	5	3	5	5	2	32
8	CV	4								4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE							4		4
	VE								4	4
8 Total		4	4	3	4	3	4	4	6	32
11	CV	4								4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
11 Total		5	4	3	4	4	4	3	5	32
12	CV			1				2	1	4
	GE	1				1	1		1	4
	IN	1	1			1	1			4
	OX			1	2				1	4
	PE	1		1		1	1			4
	PF	1	3							4
	TE	1		1				1	1	4
	VE	2		1					1	4
12 Total		7	4	5	2	3	3	3	5	32
13	CV	2						2		4
	GE		4							4
	IN			3		1				4
	OX				4					4
	PE					3	1			4
	PF						4			4
	TE	1						3		4
	VE			1				1	2	4
13 Total		3	4	4	4	4	5	6	2	32
14	CV	4								4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE	3						1		4
	VE					1			3	4
14 Total		7	4	3	4	4	4	1	5	32
15	CV	1						3		4
	GE		4							4
	IN			3		1				4
	OX				4					4
	PE			1		3				4
	PF					1	3			4
	TE							4		4

	VE			1		2			1	4
15 Total		1	4	5	4	7	3	7	1	32
16	CV			1			1	2		4
	GE		2				1		1	4
	IN						2		2	4
	OX	1		2			1			4
	PE				2	1	1			4
	PF			1			2	1		4
	TE	1		1				2		4
	VE	1		2					1	4
16 Total		3	2	7	2	1	8	5	4	32
17	CV	3						1		4
	GE		4							4
	IN	1		3						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
17 Total		6	4	3	4	4	4	3	4	32
18	CV	3						1		4
	GE		4							4
	IN			3				1		4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
18 Total		5	4	3	4	4	4	4	4	32
19	CV	3			1					4
	GE		4							4
	IN			3		1				4
	OX			1	3					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE				2				2	4
19 Total		5	4	4	6	5	4	2	2	32
		66	58	61	63	56	61	56	59	480

Non-Mnemonic Participants

Subject	R Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total

2	CV	3						1		4
	GE		4							4
	IN			3	1					4
	OX				4					4
	PE					4				4
	PF					1	3			4
	TE	1						3		4
	VE								4	4
2 Total		4	4	3	5	5	3	4	4	32
4	CV	2			1				1	4
	GE		4							4
	IN	1		2	1					4
	OX				3				1	4
	PE					2	1		1	4
	PF						4			4
	TE	1						3		4
	VE	1							3	4
4 Total		5	4	2	5	2	5	3	6	32
5	CV				1	2			1	4
	GE		4							4
	IN				1		1	1	1	4
	OX			1	2				1	4
	PE	1				2		1		4
	PF					1	2	1		4
	TE			2	1				1	4
	VE			1	1	1		1		4
5 Total		1	4	4	6	6	3	4	4	32
6	CV	4								4
	GE		4							4
	IN			3			1			4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE					1			3	4
6 Total		4	4	3	4	5	5	4	3	32
7	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
7 Total		4	4	4	4	4	4	4	4	32
9	CV	4								4
	GE		4							4
	IN			2					2	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4

9 Total		6	4	2	4	4	4	2	6	32
13	CV	3						1		4
	GE		4							4
	IN	1		3						4
	OX				4					4
	PE					3	1			4
	PF						4			4
	TE	3						1		4
	VE					2		1	1	4
13 Total		7	4	3	4	5	5	3	1	32
14	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
14 Total		4	4	4	4	4	4	4	4	32
15	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE	2						2		4
	VE								4	4
15 Total		6	4	4	4	3	4	2	5	32
16	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
16 Total		4	4	4	4	4	4	4	4	32
17	CV	4								4
	GE		3			1				4
	IN			2					2	4
	OX	1			3					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE			1		1			2	4
17 Total		6	3	3	3	6	4	3	4	32
18	CV	3						1		4
	GE		4							4
	IN			1			1		2	4
	OX	1			2	1				4
	PE		1	2					1	4
	PF					1	3			4
	TE							4		4

	VE			1		2			1	4
18 Total		4	5	4	2	4	4	5	4	32
20	CV	1			1	2				4
	GE	2		1		1				4
	IN		1					1	2	4
	OX	1	1				2			4
	PE	2		1		1				4
	PF			1	1	1			1	4
	TE		1	1			2			4
	VE	2	1				1			4
20 Total		8	4	4	2	5	5	1	3	32
		63	52	44	51	57	54	43	52	416

Appendix E.5 Confusion Tables Distracter Task

Mnemonic Participants

Subject	R Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total
1	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
1 Total		5	4	4	4	4	4	3	4	32
2	CV	2				1		1		4
	GE		4							4
	IN			3	1					4
	OX				3				1	4
	PE					3			1	4
	PF			1	3					4
	TE	1						3		4
	VE			1			1		2	4
2 Total		3	4	5	7	4	1	4	4	32
4	CV	3						1		4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
4 Total		5	4	4	4	4	4	3	4	32
6	CV	2						1	1	4
	GE		4							4
	IN			2			1		1	4
	OX				4					4
	PE				1	1	1		1	4
	PF			1		1	1		1	4
	TE	1						2	1	4
	VE			1			1		2	4
6 Total		3	4	4	5	2	4	3	7	32
7	CV			1	1			2		4
	GE		4							4
	IN			2				1	1	4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE	2						2		4
	VE	2		1			1			4
7 Total		4	4	4	5	3	5	5	2	32
8	CV	4								4
	GE		4							4

	IN			3					1	4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE							4		4
	VE								4	4
8 Total		4	4	3	4	3	4	4	6	32
11	CV	4								4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE								4	4
11 Total		5	4	3	4	4	4	3	5	32
12	CV			1				2	1	4
	GE	1				1	1		1	4
	IN	1	1			1	1			4
	OX			1	2				1	4
	PE	1		1		1	1			4
	PF	1	3							4
	TE	1		1				1	1	4
	VE	2		1					1	4
12 Total		7	4	5	2	3	3	3	5	32
13	CV	2						2		4
	GE		4							4
	IN			3		1				4
	OX				4					4
	PE					3	1			4
	PF						4			4
	TE	1						3		4
	VE			1				1	2	4
13 Total		3	4	4	4	4	5	6	2	32
14	CV	4								4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE	3						1		4
	VE					1			3	4
14 Total		7	4	3	4	4	4	1	5	32
15	CV	1						3		4
	GE		4							4
	IN			3		1				4
	OX				4					4
	PE			1		3				4
	PF					1	3			4
	TE							4		4
	VE			1		2			1	4
15 Total		1	4	5	4	7	3	7	1	32
16	CV			1			1	2		4

	GE		2				1		1	4
	IN						2		2	4
	OX	1		2			1			4
	PE				2	1	1			4
	PF			1			2	1		4
	TE	1		1				2		4
	VE	1		2					1	4
16 Total		3	2	7	2	1	8	5	4	32
17	CV	3						1		4
	GE		4							4
	IN	1		3						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
17 Total		6	4	3	4	4	4	3	4	32
18	CV	3						1		4
	GE		4							4
	IN			3				1		4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
18 Total		5	4	3	4	4	4	4	4	32
19	CV	3			1					4
	GE		4							4
	IN			3		1				4
	OX			1	3					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE				2				2	4
19 Total		5	4	4	6	5	4	2	2	32

Non-Mnemonic
Participants

		Response								
Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
2	CV	3						1		4
	GE		4							4
	IN			3	1					4
	OX				4					4
	PE					4				4
	PF					1	3			4
	TE	1						3		4
	VE								4	4
2 Total		4	4	3	5	5	3	4	4	32
4	CV	2			1				1	4
	GE		4							4
	IN	1		2	1					4
	OX				3				1	4

	PE					2	1		1	4
	PF						4			4
	TE	1						3		4
	VE	1							3	4
4 Total		5	4	2	5	2	5	3	6	32
5	CV				1	2			1	4
	GE		4							4
	IN				1		1	1	1	4
	OX			1	2				1	4
	PE	1				2		1		4
	PF					1	2	1		4
	TE			2	1				1	4
	VE			1	1	1		1		4
5 Total		1	4	4	6	6	3	4	4	32
6	CV	4								4
	GE		4							4
	IN			3			1			4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE					1			3	4
6 Total		4	4	3	4	5	5	4	3	32
7	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
7 Total		4	4	4	4	4	4	4	4	32
9	CV	4								4
	GE		4							4
	IN			2					2	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	2						2		4
	VE								4	4
9 Total		6	4	2	4	4	4	2	6	32
13	CV	3						1		4
	GE		4							4
	IN	1		3						4
	OX				4					4
	PE					3	1			4
	PF						4			4
	TE	3						1		4
	VE					2		1	1	4
13 Total		7	4	3	4	5	5	3	1	32
14	CV	4								4
	GE		4							4
	IN			4						4

	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
14 Total		4	4	4	4	4	4	4	4	32
15	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					3			1	4
	PF						4			4
	TE	2						2		4
	VE								4	4
15 Total		6	4	4	4	3	4	2	5	32
16	CV	4								4
	GE		4							4
	IN			4						4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE								4	4
16 Total		4	4	4	4	4	4	4	4	32
17	CV	4								4
	GE		3			1				4
	IN			2					2	4
	OX	1			3					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE			1		1			2	4
17 Total		6	3	3	3	6	4	3	4	32
18	CV	3						1		4
	GE		4							4
	IN			1			1		2	4
	OX	1			2	1				4
	PE		1	2					1	4
	PF					1	3			4
	TE							4		4
	VE			1		2			1	4
18 Total		4	5	4	2	4	4	5	4	32
20	CV	1			1	2				4
	GE	2		1		1				4
	IN		1					1	2	4
	OX	1	1				2			4
	PE	2		1		1				4
	PF			1	1	1			1	4
	TE		1	1			2			4
	VE	2	1				1			4
20 Total		8	4	4	2	5	5	1	3	32

APPENDIX F Experiment 2 Data

Appendix F.1 Confusion Tables Day 1

Mnemonic Participants

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
		Response								
M01	CV	4		1	5			1	1	12
	GE	1	9		1		1			12
	IN	1		3	1		2	3	2	12
	OX	2		3	2	2	2	1		12
	PE			3		4	3	1	1	12
	PF	1		3		2	3	1	2	12
	TE	1		1	5		1	4		12
	VE	3		3	3	1	1	1		12
M01 Total		13	9	17	17	9	13	12	6	96
M02	CV	5	1	1	2	4	4	4	1	22
	GE		21		1					22
	IN	1		5	5	3	2	4	2	22
	OX	3		2	4	1	8		4	22
	PE	1	4	1	4	2	5	3	2	22
	PF			8	3	2	4	2	3	22
	TE	2		1	3	6	5	5		22
	VE	3		3	7	2	1	2	4	22
M02 Total		15	26	21	29	20	29	20	16	176
M03	CV	10		1			1	6	2	20
	GE		20							20
	IN			8	1	6	2		3	20
	OX	1			15		1	2	1	20
	PE			1	2	12	1	1	3	20
	PF			1	1	1	14	1	2	20
	TE	4						16		20
	VE			9		2		1	8	20
M03 Total		15	20	20	19	21	19	27	19	160
M06	CV	4		1		1		6		12
	GE		12							12
	IN			8		2	1		1	12
	OX	1			8	1	2			12
	PE					5	1		6	12
	PF					1	11			12
	TE	3			1			8		12
	VE	1							11	12
M06 Total		9	12	9	9	10	15	14	18	96
M07	CV	4		2	5	2	1	2	8	24

	GE		24							24
	IN	2		7	3	3	2	4	3	24
	OX	2		7	5		6	4		24
	PE	1		3	4	8	2	2	4	24
	PF	1		2	1	1	17		2	24
	TE	5		4	5	3	1	4	2	24
	VE	5		1	4	7	2	3	2	24
M07 Total		20	24	26	27	24	31	19	21	192
M11	CV			1	1		1	3	2	8
	GE		7				1			8
	IN	1		1	1	4		1		8
	OX	1			1	2		1	3	8
	PE	1	2		2	1	1	1		8
	PF		2			2	4			8
	TE	3			1			4		8
	VE	1			1	1		1	4	8
M11 Total		7	11	2	7	10	7	11	9	64
M12	CV	8				1		5		14
	GE		14							14
	IN			9	1	2			2	14
	OX			4	8				2	14
	PE	1			2	10		1		14
	PF		1			1	11		1	14
	TE	6	1					7		14
	VE	2		1	3	3			5	14
M12 Total		17	16	14	14	17	11	13	10	112
M14	CV	9			1			9	1	20
	GE		20							20
	IN			10		2		1	7	20
	OX	1			18		1			20
	PE			1		16	2	1		20
	PF			1		1	17	1		20
	TE	7			1			12		20
	VE	2		6		2			10	20
M14 Total		19	20	18	20	21	20	24	18	160
M19	CV	15						10	1	26
	GE		26							26
	IN			16	1		1		8	26
	OX	2			24					26
	PE					26				26
	PF	1				1	23		1	26
	TE	13						12	1	26
	VE	1		9			1		15	26
M19 Total		32	26	25	25	27	25	22	26	208

M25	CV	9			1			4		14
	GE		14							14
	IN			8		3			3	14
	OX				9	1	4			14
	PE			2	1	8	1	2		14
	PF					3	7		4	14
	TE	2			2			10		14
	VE			4		5		1	4	14
M25 Total		11	14	14	13	20	12	17	11	112
M24	CV				2	1		3	2	8
	GE		8							8
	IN	1		1		2	2		2	8
	OX	2		2	1		3			8
	PE	2		2		1	2		1	8
	PF	1		2	1	1	2		1	8
	TE	1				3		3	1	8
	VE			1	2	1	1	1	2	8
M24 Total		7	8	8	6	9	10	7	9	64

Non-Mnemonic Participants

Subject	R Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total
N04	CV	15						11		26
	GE		26							26
	IN			23					3	26
	OX				26					26
	PE			4		22				26
	PF			1			25			26
	TE	3						23		26
	VE			4		7			15	26
N04 Total		18	26	32	26	29	25	34	18	208
N05	CV	2		1	6	1		8		18
	GE		17				1			18
	IN	3		6	5			2	2	18
	OX		1	1	4		12			18
	PE		1	2		9	3		3	18
	PF	2		1		5	10			18
	TE	2			7	3		5	1	18
	VE	3		4	3	1		5	2	18
N05 Total		12	19	15	25	19	26	20	8	144
N08	CV	2	1	1	3		2		1	10
	GE	1	5	2	1			1		10
	IN	2			2	4		1	1	10
	OX			2	2		4	1	1	10
	PE	3		1		2	2	1	1	10
	PF			2	1	2		3	2	10

	TE	2			2		3	2	1	10
	VE		2	1			3	1	3	10
N08 Total		10	8	9	11	8	14	10	10	80
N09	CV	8						2		10
	GE		10							10
	IN	1		3	1	2	2		1	10
	OX				9				1	10
	PE	1		1		4	1	2	1	10
	PF	1		1	1		5	1	1	10
	TE	3				1		6		10
	VE	1		1		6		1	1	10
N09 Total		15	10	6	11	13	8	12	5	80
N10	CV	8						4		12
	GE		12							12
	IN	1		5		1	2		3	12
	OX	1		4	4		2		1	12
	PE			2	1	3	4		2	12
	PF	1				1	9		1	12
	TE	3			1			8		12
	VE			3	2	5	1	1		12
N10 Total		14	12	14	8	10	18	13	7	96
N16	CV	8		2	2					12
	GE		12							12
	IN	2		4	2	1	2	1		12
	OX	1		2	7		1		1	12
	PE			2	1	7		1	1	12
	PF			2	2		2	3	3	12
	TE	1			4	1		5	1	12
	VE	2		4		3		2	1	12
N16 Total		14	12	16	18	12	5	12	7	96
N20	CV	2		1	1			3	3	10
	GE		5		3		1	1		10
	IN	1		2			6		1	10
	OX						6	3	1	10
	PE			3		7				10
	PF	1	1		2		5	1		10
	TE	3		1	1		1	2	2	10
	VE			1		1		4	4	10
N20 Total		7	6	8	7	8	19	14	11	80
N21	CV			4	3	3	2	1	3	16
	GE		15				1			16
	IN	1		4	1		6	2	2	16
	OX			2	9	2		1	2	16
	PE	5	2			2	1	2	4	16
	PF	3		1	1	1	10			16
	TE	3	2		1	1	4	2	3	16
	VE	4		2		2	4	4		16
N21 Total		16	19	13	15	11	28	12	14	128

N22	CV	5	1	1	2	3	4	6		22
	GE		15		2		4		1	22
	IN	4		4	4		8	1	1	22
	OX	1	3	1	13	1		2	1	22
	PE	4	4	2		2	8		2	22
	PF	5		1	3	2	5	1	5	22
	TE	6		2	3	1	4	2	4	22
	VE	4	1	2	3	2	3	3	4	22
N22 Total		29	24	13	30	11	36	15	18	176
N23	CV	13		2	6			3		24
	GE		24							24
	IN			17				2	5	24
	OX	1			23					24
	PE			1		19	3		1	24
	PF						23		1	24
	TE	1		1	4	1		17		24
	VE			5		1		1	17	24
N23 Total		15	24	26	33	21	26	23	24	192
N26	CV	11			5			5	3	24
	GE		24							24
	IN	3		11	3	2			5	24
	OX	1			20			1	2	24
	PE			3		12		7	2	24
	PF			1	1		22			24
	TE	12			2			8	2	24
	VE	1		4		8	1	5	5	24
N26 Total		28	24	19	31	22	23	26	19	192

Appendix F.2 Confusion Tables Day 2 (Long Term Memory Test)

Mnemonic Participants — Response

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
M01	CV	2		1			1			4
	GE		3					1		4
	IN			1			3			4
	OX				2			1	1	4
	PE				1	1	1	1		4
	PF	1						1	2	4
	TE	1		1		1		1		4
	VE	2				1			1	4
M01 Total		6	3	3	3	3	5	5	4	32
M02	CV			1	1	1	1			4
	GE		3		1					4
	IN	1		1	1				1	4
	OX			2	1			1		4
	PE					1	2		1	4
	PF			1	1	2				4
	TE	1						2	1	4
	VE			1	3					4
M02 Total		2	3	6	8	4	3	3	3	32
M03	CV	2				1		1		4
	GE		4							4
	IN			1		1	1		1	4
	OX	1		1	1			1		4
	PE					3		1		4
	PF				2		2			4
	TE	1						3		4
	VE			3					1	4
M03 Total		4	4	5	3	5	3	6	2	32
M06	CV	1						3		4
	GE		4							4
	IN			2					2	4
	OX				4					4
	PE			1		3				4
	PF						4			4
	TE	1			1			2		4
	VE					2			2	4
M06 Total		2	4	3	5	5	4	5	4	32
M07	CV			1	2				1	4
	GE		3				1			4
	IN	1		1			1	1		4
	OX				1		1	2		4
	PE				2	1	1			4
	PF		1				2		1	4
	TE			1	1				2	4
	VE	1		1		2				4

M07 Total		2	4	4	6	3	6	3	4	32
M11	CV				2	1			1	4
	GE		2				2			4
	IN				2			1	1	4
	OX		1	1	1		1			4
	PE	1		2		1				4
	PF	1		1			2			4
	TE	1						3		4
	VE				1	2			1	4
M11 Total		3	3	4	6	4	5	4	3	32
M12	CV	2			1			1		4
	GE		3				1			4
	IN	1		2	1					4
	OX			1	2				1	4
	PE		1			3				4
	PF						4			4
	TE	1		1				2		4
	VE			1		2			1	4
M12 Total		4	4	5	4	5	5	3	2	32
M14	CV	2						2		4
	GE		4							4
	IN			2		1		1		4
	OX				4					4
	PE				1	3				4
	PF						4			4
	TE	2			1			1		4
	VE			2					2	4
M14 Total		4	4	4	6	4	4	4	2	32
M19	CV	3						1		4
	GE		4							4
	IN			3			1			4
	OX				4					4
	PE					4				4
	PF						4			4
	TE	1						3		4
	VE			1					3	4
M19 Total		4	4	4	4	4	5	4	3	32
M24	CV				1	1			2	4
	GE		2		1				1	4
	IN	1	1			1	1			4
	OX			2			1	1		4
	PE	1			1	1	1			4
	PF	1					1	1	1	4
	TE		1	1	1				1	4
	VE		1	1	1			1		4
M24 Total		3	5	4	5	3	4	3	5	32
M25	CV				1			3		4
	GE		4							4

	IN	1				2			1	4
	OX				1		3			4
	PE			2					2	4
	PF					2	2			4
	TE				2			2		4
	VE	1		1		1			1	4
M25 Total		2	4	3	4	5	5	5	4	32

Non-Mnemonic Participants Response

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
N04	CV	2						2		4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					4				4
	PF						4			4
	TE							4		4
	VE			1		1			2	4
N04 Total		2	4	4	4	5	4	6	3	32
N05	CV	1			3					4
	GE		4							4
	IN	1			1			1	1	4
	OX						4			4
	PE			1		3				4
	PF			2		2				4
	TE				3		1			4
	VE			1				2	1	4
N05 Total		2	4	4	7	5	5	3	2	32
N08	CV		1		1		1	1		4
	GE	1	1			1	1			4
	IN			2		1			1	4
	OX		1	1	1			1		4
	PE						3	1		4
	PF			1		1	2			4
	TE	1	1		1	1				4
	VE				1		2		1	4
N08 Total		2	4	4	4	4	9	3	2	32
N09	CV	2			1			1		4
	GE		4							4
	IN			3		1				4
	OX	2			1				1	4
	PE	1		1		2				4
	PF						4			4
	TE	1			2			1		4
	VE				1	3				4
N09 Total		6	4	4	5	6	4	2	1	32
N10	CV	4								4

	GE		4							4
	IN			4						4
	OX				4					4
	PE					1	3			4
	PF			1	1		2			4
	TE							4		4
	VE					2	1		1	4
N10 Total		4	4	5	5	3	6	4	1	32
N16	CV	2			1		1			4
	GE		4							4
	IN	1				1	1	1		4
	OX							2	2	4
	PE			2			1		1	4
	PF	1					3			4
	TE	1			3					4
	VE			1		2		1		4
N16 Total		5	4	3	4	3	6	4	3	32
N20	CV					1		1	2	4
	GE		1	1			1	1		4
	IN					1	3			4
	OX	1					2	1		4
	PE			1		3				4
	PF	1		1			2			4
	TE							2	2	4
	VE							2	2	4
N20 Total		2	1	3		5	8	7	6	32
N21	CV	1	1				2			4
	GE		4							4
	IN	1		1	1				1	4
	OX				2			2		4
	PE		1			2			1	4
	PF	2					2			4
	TE	2				1			1	4
	VE						1	2	1	4
N21 Total		6	6	1	3	3	5	4	4	32
N22	CV	1					1	1	1	4
	GE		3				1			4
	IN	1			1	1		1		4
	OX	1			1	1	1			4
	PE	2				1	1			4
	PF			3	1					4
	TE	1	2						1	4
	VE		1			1	2			4
N22 Total		6	6	3	3	4	6	2	2	32
N23	CV	4								4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE					4				4

	PF						4			4
	TE				1			3		4
	VE			2					2	4
N23 Total		4	4	5	5	4	4	3	3	32
N26	CV	2				1			1	4
	GE		4							4
	IN			3					1	4
	OX				4					4
	PE							3	1	4
	PF						4			4
	TE	3						1		4
	VE			1		2		1		4
N26 Total		5	4	4	4	3	4	5	3	32

Appendix F.3 Confusion Tables Day 2 (Test)

Mnemonic
Participants Response

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
M01	CV	3		1	2	2		8		16
	GE		14				2			16
	IN	1		3		2	4	3	3	16
	OX	1		2	5	4	2	1	1	16
	PE	3		1	1	5	1	3	2	16
	PF	2		4	2	4	2	2		16
	TE	8			3		1	3	1	16
	VE	1		2	3	3	3	3	1	16
M01 Total		19	14	13	16	20	15	23	8	128
M02	CV	4			4	1	1	3	3	16
	GE		16							16
	IN	1		8	1		1	5		16
	OX	1		4	3		6		2	16
	PE		1		1	12	1		1	16
	PF			8	7		1			16
	TE	3			2	3	2	6		16
	VE	2		4	2		2	3	3	16
M02 Total		11	17	24	20	16	14	17	9	128
M03	CV	11						5		16
	GE		15				1			16
	IN			10	1	4			1	16
	OX			1	14			1		16
	PE				1	9	2		4	16
	PF			1	1	1	13			16
	TE				1			15		16
	VE			6		1			9	16
M03 Total		11	15	18	18	15	16	21	14	128
M06	CV	14						2		16
	GE		16							16
	IN			13					3	16
	OX				16					16
	PE					15			1	16
	PF						16			16
	TE	2						14		16
	VE					3			13	16
M06 Total		16	16	13	16	18	16	16	17	128
M07	CV	9		1	1			4	1	16
	GE		16							16
	IN			11		1	1		3	16
	OX				14			1	1	16
	PE	1		2	1	7			5	16
	PF			1			15			16
	TE	5			2		1	8		16
	VE	3		4		4			5	16

M07 Total		18	16	19	18	12	17	13	15	128
M11	CV	4		3	5	1		1		14
	GE		12				2			14
	IN	1		3	1	5			4	14
	OX	2	2	2	5	2		1		14
	PE	3		5	1	1	1	2	1	14
	PF	1		1			12			14
	TE	2		1	2	4		2	3	14
	VE	2		3	4	2			3	14
M11 Total		15	14	18	18	15	15	6	11	112
M12	CV	13				1		2		16
	GE		15				1			16
	IN			11	4				1	16
	OX				15			1		16
	PE		2			14				16
	PF						16			16
	TE	6				1		9		16
	VE	1		1		2			12	16
M12 Total		20	17	12	19	18	17	12	13	128
M14	CV	13						3		16
	GE		16							16
	IN			11					5	16
	OX				16					16
	PE			1		14	1			16
	PF						16			16
	TE	4						12		16
	VE			8		2			6	16
M14 Total		17	16	20	16	16	17	15	11	128
M19	CV	11						5		16
	GE		16							16
	IN			14					2	16
	OX				16					16
	PE					16				16
	PF						16			16
	TE	4						12		16
	VE								16	16
M19 Total		15	16	14	16	16	16	17	18	128
M24	CV	2		2	2	2		6	2	16
	GE		14						2	16
	IN	3		1		4	4	1	3	16
	OX	2		4	3	3		2	2	16
	PE			3	3	4	2		4	16
	PF	2		2	1	1	9		1	16
	TE	4	2	1	1	2		6		16
	VE	1	1	3	4	3	1	3		16
M24 Total		14	17	16	14	19	16	18	14	128
M25	CV	13						3		16
	GE		16							16

	IN			10		2			4	16
	OX				16					16
	PE			2		13			1	16
	PF						15		1	16
	TE	1						15		16
	VE			7		3			6	16
M25 Total		14	16	19	16	18	15	18	12	128

Non-Mnemonic Participants

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
		Response								
N04	CV	9						1		10
	GE		9				1			10
	IN			9					1	10
	OX				10					10
	PE					10				10
	PF						10			10
	TE							10		10
	VE			2					8	10
N04 Total		9	9	11	10	10	11	11	9	80
N05	CV	3		2	8	2		1		16
	GE		16							16
	IN	4		3	2			5	2	16
	OX	1			5		6	1	3	16
	PE			4		7	4		1	16
	PF			1		4	11			16
	TE	4		1	8			2	1	16
	VE	1		5	2			5	3	16
N05 Total		13	16	16	25	13	21	14	10	128
N08	CV	2	1	3	2	1	3	2		14
	GE	1	5	2	2		2	2		14
	IN		2	4	1	1	3	1	2	14
	OX	1	1	1	4	1	3	1	2	14
	PE	5	1	2		2	2		2	14
	PF		4		1	2	3		4	14
	TE	2	1	2	1	1	1	5	1	14
	VE	1	1	2	2	2	1	2	3	14
N08 Total		12	16	16	13	10	18	13	14	112
N09	CV	10			4			2		16
	GE		16							16
	IN	1		1	1	5			8	16
	OX				14	1			1	16
	PE			5		11				16
	PF				2		12		2	16
	TE	4			2			10		16
	VE	1				11		1	3	16
N09 Total		16	16	6	23	28	12	13	14	128
N10	CV	13						3		16

	GE		16							16
	IN			16						16
	OX				16					16
	PE					12		1	3	16
	PF			1			15			16
	TE							16		16
	VE					4			12	16
N10 Total		13	16	17	16	16	15	20	15	128
N16	CV	10		1	1			4		16
	GE		16							16
	IN			14			1		1	16
	OX				16					16
	PE					12			4	16
	PF			2			14			16
	TE	4			1			10	1	16
	VE			5		1		2	8	16
N16 Total		14	16	22	18	13	15	16	14	128
N20	CV			1			1	5	7	14
	GE	1	5	1	4		2		1	14
	IN			6	1		3	4		14
	OX	1			1		2	4	6	14
	PE			3		4		3	4	14
	PF	3		6			4	1		14
	TE			3			1	4	6	14
	VE			2		1		5	6	14
N20 Total		5	5	22	6	5	13	26	30	112
N21	CV	3	2		1	5	2		3	16
	GE	1	14			1				16
	IN	1		4		1	3	3	4	16
	OX				8	3		3	2	16
	PE	4	2		3	1	1		5	16
	PF	4			1	2	5	1	3	16
	TE	3	5	1	2	3	1		1	16
	VE	3		2	4		1	3	3	16
N21 Total		19	23	7	19	16	13	10	21	128
N22	CV	7		1	3	1	2		2	16
	GE		11				5			16
	IN	1		1	3		2	4	5	16
	OX	3	2		7		1		3	16
	PE		3	3	2	3	5			16
	PF	1	2	2	4	1	3	2	1	16
	TE	3	1	3	5	1	3			16
	VE	2	3	1	2	1	2	2	3	16
N22 Total		17	22	11	26	7	23	8	14	128
N23	CV	11						5		16
	GE		16							16
	IN			15			1			16
	OX				16					16
	PE					15			1	16

	PF						16			16
	TE							16		16
	VE			5					11	16
N23 Total		11	16	20	16	15	17	21	12	128
N26	CV	13						3		16
	GE		16							16
	IN			9		3			4	16
	OX				16					16
	PE			1		8		2	5	16
	PF						16			16
	TE	3						13		16
	VE			1		6		1	8	16
N26 Total		16	16	11	16	17	16	19	17	128

Appendix F.4 Confusion Tables Distracter Task

Mnemonic
Participants Response

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
M01	CV	5				1	1			7
	GE		4			1		1		6
	IN						2	3	1	6
	OX	2		1				2		5
	PE	1		1	1		2	1		6
	PF	1		2	1	1	1			6
	TE	1			3		1	1		6
	VE	1		1	3	1				6
M01 Total		11	4	5	8	4	7	8	1	48
M02	CV	1		1	2	1		2		7
	GE		5			1				6
	IN	1		1			3		1	6
	OX	1		2	1		1			5
	PE		1		1	3	1			6
	PF	1		2	1			2		6
	TE	1				1	1	3		6
	VE	1		1	1	2		1		6
M02 Total		6	6	7	6	8	6	8	1	48
M03	CV	4		1				2		7
	GE		6							6
	IN			1	1	1	1		2	6
	OX				5					5
	PE					4	2			6
	PF						6			6
	TE	1						5		6
	VE			1		1	2		2	6
M03 Total		5	6	3	6	6	11	7	4	48
M06	CV	5		1				1		7
	GE		6							6
	IN			5	1					6
	OX				5					5
	PE					4			2	6
	PF						6			6
	TE							6		6
	VE					2			4	6
M06 Total		5	6	6	6	6	6	7	6	48
M07	CV	3		1				3		7
	GE		6							6
	IN	1		5						6
	OX				4			1		5
	PE	1				3			2	6
	PF						6			6
	TE					1		5		6
	VE			1		1		1	3	6

M07 Total		5	6	7	4	5	6	10	5	48
M11	CV	2			1	1		3		7
	GE		6							6
	IN	1		1	1			1	1	5
	OX	2		1		1			1	5
	PE	1			1	1		1	2	6
	PF			1			5			6
	TE	4			1			1		6
	VE	1		1		1			3	6
M11 Total		11	6	4	4	4	5	6	7	47
M12	CV	3			1			2	1	7
	GE		6							6
	IN			2	1	1	1	1		6
	OX			2	3					5
	PE					4	2			6
	PF						6			6
	TE	2	1					3		6
	VE					1			5	6
M12 Total		5	7	4	5	6	9	6	6	48
M14	CV	3		1				3		7
	GE		6							6
	IN			1	2				3	6
	OX				5					5
	PE					6				6
	PF						6			6
	TE	2						4		6
	VE			4					2	6
M14 Total		5	6	6	7	6	6	7	5	48
M19	CV	5		1				1		7
	GE		6							6
	IN			4	1		1			6
	OX				5					5
	PE					6				6
	PF					1	5			6
	TE	1						5		6
	VE								6	6
M19 Total		6	6	5	6	7	6	6	6	48
M24	CV				1	1		4	1	7
	GE		6							6
	IN	3			1	1	1			6
	OX	1			1			2	1	5
	PE	1				4		1		6
	PF				1	1	2		2	6
	TE				1	2		2		5
	VE	2			1	1			2	6
M24 Total		7	6		6	10	3	9	6	47
M25	CV	5		1				1		7
	GE		6							6

	IN			4	1	1				6
	OX				5					5
	PE			1		5				6
	PF			2			4			6
	TE							6		6
	VE	1				1			4	6
M25 Total		6	6	8	6	7	4	7	4	48

Non-Mnemonic Participants

Subject	R Event	Response CV	GE	IN	OX	PE	PF	TE	VE	Total
N04	CV	6		1						7
	GE		6							6
	IN			4	1				1	6
	OX				5					5
	PE					6				6
	PF						6			6
	TE	3						3		6
	VE								6	6
N04 Total		9	6	5	6	6	6	3	7	48
N05	CV			1	4	1		1		7
	GE		6							6
	IN			3			1	1	1	6
	OX						3		2	5
	PE			3		2			1	6
	PF	2					4			6
	TE	2		2	2					6
	VE			3				3		6
N05 Total		4	6	12	6	3	8	5	4	48
N08	CV	1			1		4	1		7
	GE	2	2	1					1	6
	IN				2	1	3			6
	OX				1		4			5
	PE		1	1		1	1	2		6
	PF	1			3		1	1		6
	TE	3		1			2			6
	VE	1	1	1		1		2		6
N08 Total		8	4	4	7	3	15	6	1	48
N09	CV	2			1	1		3		7
	GE		6							6
	IN			1	1	3		1		6
	OX				5					5
	PE			1		4	1			6
	PF					1	5			6
	TE				1			5		6
	VE				1	4			1	6
N09 Total		2	6	2	9	13	6	9	1	48
N10	CV	5		1				1		7

	GE		6							6
	IN			4	1				1	6
	OX				5					5
	PE					5	1			6
	PF						6			6
	TE	1						5		6
	VE			1					5	6
N10 Total		6	6	6	6	5	7	6	6	48
N16	CV	4		1	1			1		7
	GE		6							6
	IN			4	1				1	6
	OX				5					5
	PE					3			3	6
	PF			1			5			6
	TE							6		6
	VE			3	1				2	6
N16 Total		4	6	9	8	3	5	7	6	48
N20	CV		1		1		1	3	1	7
	GE		6							6
	IN			3			3			6
	OX	1					4			5
	PE					5			1	6
	PF	2		1			3			6
	TE						1	3	2	6
	VE							3	3	6
N20 Total		3	7	4	1	5	12	9	7	48
N21	CV	1		2		2			2	7
	GE	1	4		1					6
	IN	1			3		1		1	6
	OX			1	4					5
	PE	1			1	1	2		1	6
	PF	2				1	3			6
	TE	3						1	2	6
	VE			1		1	1	1	2	6
N21 Total		9	4	4	9	5	7	2	8	48
N22	CV	3		2	1				1	7
	GE		6							6
	IN	1		1	2				2	6
	OX				4				1	5
	PE			2	1		2		1	6
	PF	1		2			3			6
	TE	1		2			2		1	6
	VE	2		1					3	6
N22 Total		8	6	10	8		7		9	48
N23	CV	6		1						7
	GE		6							6
	IN			4	1				1	6
	OX				5					5
	PE					6				6

	PF						6			6
	TE							6		6
	VE			1					5	6
N23 Total		6	6	6	6	6	6	6	6	48
N26	CV	6		1						7
	GE		6							6
	IN			4	1		1			6
	OX				5					5
	PE					2			4	6
	PF						6			6
	TE	2						4		6
	VE			1		2			3	6
N26 Total		8	6	6	6	4	7	4	7	48

Appendix F.5 Experiment 1 vs. Experiment 2 –Three Factor ANOVA Descriptive Statistics.

			Count	Mean
Musical			32	0.729163178
Non-Musical			28	0.443195644
	Mnemonic		30	0.60755081
	Non-Mnemonic		30	0.564808011
		Experiment 1	38	0.644845713
		Experiment 2	22	0.527513109
Musical	Mnemonic		16	0.759931345
Musical	Non-Mnemonic		16	0.698395011
Non-Musical	Mnemonic		14	0.455170276
Non-Musical	Non-Mnemonic		14	0.431221011
Musical		Experiment 1	24	0.779930923
Musical		Experiment 2	8	0.678395433
Non-Musical		Experiment 1	14	0.509760502
Non-Musical		Experiment 2	14	0.376630785
	Mnemonic	Experiment 1	19	0.670817743
	Mnemonic	Experiment 2	11	0.544283878
	Non-Mnemonic	Experiment 1	19	0.618873683
	Non-Mnemonic	Experiment 2	11	0.510742339
Musical	Mnemonic	Experiment 1	12	0.826633523
Musical	Mnemonic	Experiment 2	4	0.693229167
Musical	Non-Mnemonic	Experiment 1	12	0.733228324
Musical	Non-Mnemonic	Experiment 2	4	0.663561699
Non-Musical	Mnemonic	Experiment 1	7	0.515001962
Non-Musical	Mnemonic	Experiment 2	7	0.39533859
Non-Musical	Non-Mnemonic	Experiment 1	7	0.504519042
Non-Musical	Non-Mnemonic	Experiment 2	7	0.35792298
			60	0.586179411
			SST	2.942934905

APPENDIX G Experiment 3 Data

Appendix G.1 Confusion Tables

Subject	R Event	GE	IN	OX	PE	PF	TE	VE	Total
P02	GE	10							10
	IN		9		1				10
	OX			10					10
	PE	1			6	2		1	10
	PF					10			10
	TE						10		10
	VE							10	10
P02 Total		11	9	10	7	12	10	11	70
P03	GE	40							40
	IN		21	3	2			14	40
	OX		3	30			3	4	40
	PE		4	1	30	1	2	2	40
	PF		1		1	36	1	1	40
	TE		1	4	1	1	32	1	40
	VE		13	1				26	40
P03 Total		40	43	39	34	38	38	48	280
P04	GE	27				1			28
	IN		10	6		5	2	5	28
	OX		2	13				13	28
	PE	4	4	1	15	3	1		28
	PF	1	4	2	2	18		1	28
	TE			1	1		26		28
	VE		3	2	9	1	7	6	28
P04 Total		32	23	25	27	28	36	25	196
P05	GE	19				1			20
	IN		12	2	5			1	20
	OX		1	19					20
	PE		2		10	3		5	20
	PF		1	1		18			20
	TE			2			18		20
	VE			2			1	17	20
P05 Total		19	16	26	15	22	19	23	140
P06	GE	10							10
	IN		7		2	1			10
	OX		1	9					10
	PE				7			3	10
	PF		1			9			10
	TE						10		10
	VE					1		9	10
P06 Total		10	9	9	9	11	10	12	70
P07	GE	16				1	2	1	20
	IN		8	2	2	2	1	5	20
	OX		4	10		3	1	2	20
	PE		3		3	6	5	3	20
	PF		8	2	2	5		3	20
	TE	2	2			1	13	2	20

	VE		3	1	3	2	7	4	20
P07 Total		18	28	15	10	20	29	20	140
P08	GE	22							22
	IN		7	3	4	2	1	5	22
	OX			21				1	22
	PE		2	2	5	6	3	4	22
	PF		1	2	5	14			22
	TE		1		1		17	3	22
	VE		1	2	2	2	3	12	22
P08 Total		22	12	30	17	24	24	25	154
P09	GE	13			2		1		16
	IN		5	5			5	1	16
	OX		2	10		1	2	1	16
	PE	1	1		11		1	2	16
	PF		4		1	9	1	1	16
	TE		3	2	4		4	3	16
	VE		3	4	1	1	5	2	16
P09 Total		14	18	21	19	11	19	10	112
P10	GE	26			1	1			28
	IN		13	2	3	5	2	3	28
	OX	1	13	4	5	1	2	2	28
	PE	1	2	1	14	7	1	2	28
	PF		1	2	7	16	2		28
	TE	2	2		3	4	13	4	28
	VE		6	1	4	4	3	10	28
P10 Total		30	37	10	37	38	23	21	196
P11	GE	16							16
	IN		11		2			3	16
	OX			14				2	16
	PE				15	1			16
	PF					16			16
	TE						16		16
	VE		1		7			8	16
P11 Total		16	12	14	24	17	16	13	112
P12	GE	10	1		2		1		14
	IN	1	5	2	2	3	1		14
	OX		3	7	1	1		2	14
	PE	2	1	1	6		1	3	14
	PF	2		2		8		2	14
	TE		3		1		7	3	14
	VE	2	6		2	1	2	1	14
P12 Total		17	19	12	14	13	12	11	98
P13	GE	8							8
	IN		7			1			8
	OX			6	1			1	8
	PE				5	2		1	8
	PF				1	7			8
	TE						8		8
	VE				1	1		6	8
P13 Total		8	7	6	8	11	8	8	56
P14	GE	22				1	1		24
	IN		5	3	7	4	2	3	24
	OX	3	4	8	1	3	1	4	24

	PE	4	1		9	9	1		24
	PF		1		3	16	1	3	24
	TE	4	3	4	4	2	5	2	24
	VE	1	5	2	8		7	1	24
P14 Total		34	19	17	32	35	18	13	168
P15	GE	20							20
	IN		6		9	1		4	20
	OX			19			1		20
	PE		11		4			5	20
	PF	1	2		2	13		2	20
	TE			1			19		20
	VE		4		2	4		10	20
P15 Total		21	23	20	17	18	20	21	140
P16	GE	26							26
	IN		10	1	2	2		11	26
	OX		1	23	1		1		26
	PE	1		2	18	1		4	26
	PF	1	1			16		8	26
	TE				1		24	1	26
	VE		4	1	8	1		12	26
P16 Total		28	16	27	30	20	25	36	182
101	GE	4							4
	IN		4						4
	OX			4					4
	PE				4				4
	PF					4			4
	TE						4		4
	VE							4	4
101 Total		4	4	4	4	4	4	4	28
102	GE	4							4
	IN		4						4
	OX			4					4
	PE				4				4
	PF					4			4
	TE						4		4
	VE							4	4
102 Total		4	4	4	4	4	4	4	28
103	GE	24							24
	IN		8	5	3	3		5	24
	OX			22		1		1	24
	PE	1	1		15	3		4	24
	PF		8	1	1	13		1	24
	TE						24		24
	VE		5		4	1	1	13	24
103 Total		25	22	28	23	21	25	24	168
104	GE	24							24
	IN		19			1		4	24
	OX			24					24
	PE		3		16	1		4	24
	PF		5			18		1	24
	TE						24		24
	VE		1		8			15	24
104 Total		24	28	24	24	20	24	24	168

105	GE	19	3	1	4	8	1		36
	IN	3	10	8	3	3	5	4	36
	OX	1	6	5	5	8	7	4	36
	PE	4	4	6	4	11	6	1	36
	PF	6	5	7	5	5	5	3	36
	TE	6	5	10	5	3	5	2	36
	VE	4	6	6	8	2	5	5	36
105 Total		43	39	43	34	40	34	19	252
106	GE	36							36
	IN		32	2	1			1	36
	OX			33		2	1		36
	PE				30	2		4	36
	PF			2		34			36
	TE		2	1			32	1	36
	VE		9		3			24	36
106 Total		36	43	38	34	38	33	30	252
107	GE	28							28
	IN		12		5			11	28
	OX			28					28
	PE				27	1			28
	PF				2	26			28
	TE				1		27		28
	VE				5			23	28
107 Total		28	12	28	40	27	27	34	196
108	GE	34					2		36
	IN		27			5		4	36
	OX			34		1		1	36
	PE		1		26		3	6	36
	PF					32	2	2	36
	TE	2	1		3		27	3	36
	VE		4		5	5	1	21	36
108 Total		36	33	34	34	43	35	37	252
		520	476	484	497	515	493	473	3458

APPENDIX H Experiment 3 Data

Appendix H.1 Confusion Tables

Subject	R Event	CV	GE	IN	OX	PE	PF	TE	VE	Total
1	CV	21						1		22
	GE		22							22
	IN			22						22
	OX	2			20					22
	PE					17			5	22
	PF						22			22
	TE	1			2			19		22
	VE			1		6			15	22
1 Total		24	22	23	22	23	22	20	20	176
2	CV	8								8
	GE		8							8
	IN			4				2	2	8
	OX			1	7					8
	PE					5			3	8
	PF						8			8
	TE	3						5		8
	VE	1		1				2	4	8
2 Total		12	8	6	7	5	8	9	9	64
3	CV	7	1		1			4	3	16
	GE		16							16
	IN	1		13	1			1		16
	OX				15			1		16
	PE			1		15				16
	PF	1		1	1		13			16
	TE	1		3		1		10	1	16
	VE			1		1	2	1	11	16
3 Total		10	17	19	18	17	15	17	15	128
4	CV	9	2		3			3	1	18
	GE		17			1				18
	IN	1		12	1		1		3	18
	OX				14		1	2	1	18
	PE					16	2			18
	PF	3			1		12	2		18
	TE				2	1		12	3	18
	VE	1	1	2		1	2		11	18
4 Total		14	20	14	21	19	18	19	19	144
5	CV	16			2			1	1	20
	GE		20							20
	IN	4		6	1	1	2	2	4	20
	OX	1	1	1	15	1			1	20
	PE	3	2	4	1	7			3	20
	PF		1			2	10	3	4	20
	TE	3		1			2	12	2	20
	VE	1		6	1	4	1	4	3	20
5 Total		28	24	18	20	15	15	22	18	160
6	CV	8			4	1		5		18

	GE		18							18
	IN	1	1	6	1	2		4	3	18
	OX		1		11	1	5			18
	PE		2	2		9	1		4	18
	PF			2	2	3	8		3	18
	TE	10				1		7		18
	VE			4		8	1		5	18
6 Total		19	22	14	18	25	15	16	15	144
7	CV	10			3			1		14
	GE		14							14
	IN			13	1					14
	OX	2			11		1			14
	PE			4		8		1	1	14
	PF						14			14
	TE	1						13		14
	VE			4		2			8	14
7 Total		13	14	21	15	10	15	15	9	112
9	CV	5		1	3	1	5	2	1	18
	GE		17	1						18
	IN	2		5	1	2	1	2	5	18
	OX	2		2	10		3	1		18
	PE	3		3	1	11				18
	PF			1	1	3	11	2		18
	TE	1		3	1	1	2	8	2	18
	VE	2	1	3	3	3		2	4	18
9 Total		15	18	19	20	21	22	17	12	144
10	CV	3	2	2	4	4	3	12		30
	GE	2	27		1					30
	IN	12		4	5		4	3	2	30
	OX	8		3	6	1	1	4	7	30
	PE	5	1	3	4	4	5	3	5	30
	PF	4	3	4	4	4	4		7	30
	TE	9	1	3	1	2	2	12		30
	VE	5	1	6	5	1	3	5	4	30
10 Total		48	35	25	30	16	22	39	25	240
RAE001	CV	12						8		20
	GE		20							20
	IN			15					5	20
	OX				20					20
	PE			1		16			3	20
	PF				1		19			20
	TE	2		1				17		20
	VE			2					18	20
RAE001 Total		14	20	19	21	16	19	25	26	160
RAE002	CV	18			1			10	1	30
	GE		30							30
	IN	1		19	1		5		4	30
	OX				30					30
	PE					29			1	30
	PF						30			30
	TE	18			1			11		30
	VE					16		2	12	30
RAE002 Total		37	30	19	33	45	35	23	18	240

RAE003	CV	15		1	9	7	2	4	2	40
	GE		37				1	2		40
	IN	5		8	8	9	7	3		40
	OX	6	2	7	11	3	5	3	3	40
	PE	4		6	2	3	6	5	14	40
	PF	3	1	4	6	2	23	1		40
	TE	11	3	4	5	10	3	3	1	40
	VE	7		10	4	6	3	2	8	40
RAE003 Total		51	43	40	45	40	50	23	28	320
RAE004	CV	8								8
	GE		8							8
	IN			6		1		1		8
	OX				8					8
	PE			3		5				8
	PF						8			8
	TE	1				1		6		8
	VE								8	8
RAE004 Total		9	8	9	8	7	8	7	8	64
RAE005	CV	18	1		3			8		30
	GE		29				1			30
	IN		1	17	1	4			7	30
	OX		1		27		1		1	30
	PE		1	4	1	19			5	30
	PF	2	1				25	1	1	30
	TE	7	2		1			20		30
	VE		1	6	1	5			17	30
RAE005 Total		27	37	27	34	28	27	29	31	240
RAE006	CV	17	1		4	2		6	4	34
	GE		32				1	1		34
	IN	2		14		11	3	1	3	34
	OX	2		2	29		1			34
	PE	1		1		30	1	1		34
	PF			3	1		28	1	1	34
	TE	6		1			1	11	15	34
	VE	9			1	7		4	13	34
RAE006 Total		37	33	21	35	50	35	25	36	272
RAE007	CV	8	2	1	2	3	1	4	1	22
	GE		21	1						22
	IN	3		12	1		2	1	3	22
	OX				17	2		1	2	22
	PE	2	2	4		6	1	4	3	22
	PF	2		3		1	13	2	1	22
	TE	7		1	1	1		11	1	22
	VE	5		4	1	2		5	5	22
RAE007 Total		27	25	26	22	15	17	28	16	176
RAE008	CV	21		1	9			2	1	34
	GE		31				3			34
	IN	4	1	17	1		1		10	34
	OX	6		1	25				2	34
	PE			1		28	2		3	34
	PF	1		3		2	25	1	2	34
	TE	1						33		34
	VE	2		4	1	1			26	34

RAE008 Total		35	32	27	36	31	31	36	44	272
RAE010	CV	12			1			1		14
	GE		14							14
	IN			13		1				14
	OX				13				1	14
	PE					9			5	14
	PF				1	6	7			14
	TE	4				1		9		14
	VE					3	2		9	14
RAE010 Total		16	14	13	15	20	9	10	15	112
RAE011	CV	22	1		1					24
	GE		24							24
	IN	1		19			1	1	2	24
	OX				23				1	24
	PE			1		22			1	24
	PF			1			22		1	24
	TE				1			23		24
	VE			4	1	2			17	24
RAE011 Total		23	25	25	26	24	23	24	22	192
RAE012	CV	12			1				1	14
	GE		14							14
	IN			13	1					14
	OX	2			12					14
	PE					11	1	1	1	14
	PF				4		9	1		14
	TE	1						12	1	14
	VE			2		2			10	14
RAE012 Total		15	14	15	18	13	10	14	13	112

APPENDIX I <u>Recruitment Brochure</u>

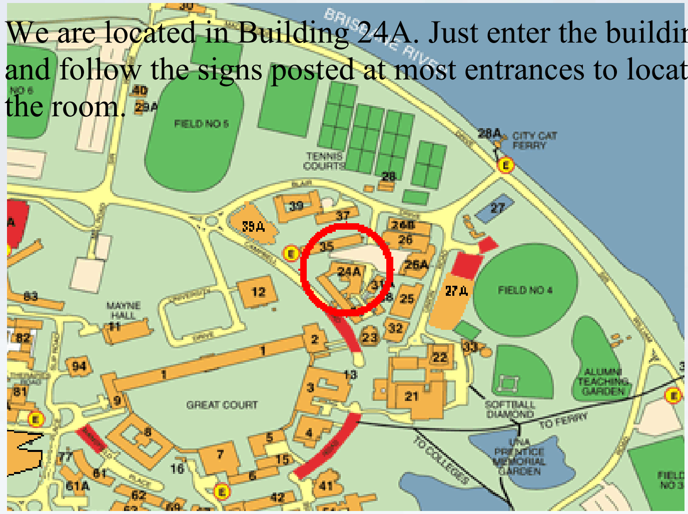
Getting to the University of Queensland

By car
The main entrance to the University is along Sir Fred Schonell Drive, located on Map 179 F2 of the *UBD Brisbane "Refidex" Street Directory*.

By ferry
The CityCat ferry is the most enjoyable and often the most convenient means of getting to the St Lucia campus. The campus CityCat terminal is located near the corner of Sir William MacGreggor Drive and Blair Drive.

By bus
From Brisbane's Central Business District, the Brisbane City Council bus routes which run directly to the University of Queensland are the 407 (Rocket), the 412 (Express), and the 411. Other BCC bus routes which stop here are the 402 (Toowong), the 428 (Indooroopilly), the 427 (Chapel Hill), the 432 (Kenmore) and the 414 (West Taringa).
The main bus stop on the St Lucia campus is located at Chancellor's Place near the J.D. Story building. For timetables and further information call TransInfo on 13 12 30.

We are located in Building 24A. Just enter the building and follow the signs posted at most entrances to locate the room.



Contacts

Alexandra Wee
School of Information Technology and Electrical Engineering
University of Queensland
Tel: 33651634
Mobile: 04 221 77840
Email: alexwee@itee.uq.edu.au

Professor Penelope Sanderson
Key Centre for Human Factors and Applied Cognitive Psychology
Tel: (07) 3365 3988
Email: psanderson@itee.uq.edu.au

Contact Address:
The University of Queensland
St Lucia, Australia
Qld, 4072

Paid Participation For Registered Nurses Earn

$40

for a 2.5 hr experiment

for research into the

Improvement of Monitoring Alarms

in Patient Monitoring Equipment

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Why should you participate?

IEC standards have proposed a new standard for alarms to be used in medical electrical equipment. While the alarm standard has now been in place for 2 years, the alarms have not been adopted in many clinical settings.

Given the highly safety critical environments in which the alarms will be used, it is essential that we test the alarms for memorability and discriminability.

This research has been cleared with the University of Queensland Ethics Committee and the Australian New Zealand Collage of Anesthetists (ANZCA) Ethics board.

What does this study this involve?

In this experiment we will be evaluating the 2003 standard of alarms which used in Medical Electrical equipment. We are interested in assessing the alarms for their ability to convey the correct information and whether they are easily distinguishable from each other. Our goal is to test the alarms design, **NOT** to test you or your knowledge. However we do ask that you try to do your best in the evaluations.

You will receive introduction information regarding the alarms, which will involve listening to knowing some background as to which machines in an operating theatre would emit these alarms.

Where

The experiment may be conducted at the University of Queensland Usability Lab (see campus map. Parking permits for the experiment days can be arranged) or at UQ facilities such as UQ libraries located on the Princess Alexandra Hospital and Royal Brisbane Hospital Campuses.

When

The experiment is spread over 2 sessions. This can be whenever is convenient for you. The sessions should be about 1 week apart.

Session 1 lasts a maximum of 50 mins
Session 2 lasts a maximum of 1h 30 mins

What's in it for you

You will be compensated for your time and participation in this study with $40 and a small gift which you may chose from a selection.

Contact for Appointment

Alexandra Wee

Mobile: 04 221 77840
Tel: 336 51634
Email: alexwee@itee.uq.edu.au