Reducing Fire Deaths in Older Adults: Optimizing the Smoke Alarm Signal Research Project

Investigation of Auditory Arousal with Different Alarm Signals in Sleeping Older Adults

THE
FIRE PROTECTION
RESEARCH FOUNDATION

FIRE RESEARCH

Reducing Fire Deaths in Older Adults: Optimizing the Smoke Alarm Signal
Research Project

Investigation of Auditory Arousal with Different Alarm Signals in Sleeping Older Adults

Prepared by

Dorothy Bruck
Ian Thomas
Ada Kritikos

Victoria University

THE
FIRE PROTECTION
RESEARCH FOUNDATION

FIRE RESEARCH

FOREWORD

Smoke alarm and signaling systems are a proven strategy for reduction of fire fatalities in the general population. However, studies have shown that the elderly do not fully benefit from conventional smoke alarm systems, particularly during the sleeping hours. In April of 2005, the Fire Protection Research Foundation was awarded a Fire Prevention and Safety Grant by the US Fire Administration for a new project to study this topic.

A portion of the study involved the conduct of human behavior studies to investigate the arousal thresholds from sleep in older adults to the current US smoke alarm and compare these thresholds to several alternative signals, and to investigate the performance abilities of older adults when awoken suddenly by an alarm. This report presents the results of this portion of the study.

The overall goal of the project is to optimize the performance requirements for alarm and signaling systems to meet the needs of an aging population. The balance of the study is presented in a companion report also published by the Foundation entitled "Reducing Fire Deaths in Older Adults: Optimizing the Fire Alarm Signal".

The Research Foundation expresses gratitude to: this report authors: Dorothy Bruck, Ian Thomas, and Ada Kritikos, Victoria University, Australia; the Project Technical Panel: Guylene Proulx, David Albert, Dana Mulvany, Arthur Lee, Donald Sievers, Rita Fahy, Wendy Gifford, Isaac Papier, Karen Boyce, Leonard Belliveau, Paul Patty, and Lee Richardson; and the project sponsors: US Fire Administration, BRK Brands/First Alert, Innovalarm, SimplexGrinnell, Siemens Building Technologies, National Electrical Manufacturers Association, GE Security, Honeywell, and Kidde.

The content, opinions and conclusions contained in this report are solely those of the authors.

Reducing Fire Deaths in Older Adults: Optimizing the Smoke Alarm Signal Research Project

Technical Panel

David Albert, InnovAlarm

Leonard Belliveau, Hughes Associates, Inc.

Karen Boyce, University of Ulster

Rita Fahy, NFPA

Wendy Gifford, Invensys Controls/Firex

Arthur Lee, U.S. Consumer Product Safety Commission

Dana Mulvany

Isaac Papier, Honeywell Life Safety

Paul Patty, Underwriters Laboratories Inc.

Guylene Proulx, National Research Council of Canada

Lee Richardson, NFPA

Donald Sievers, D.E. Sievers & Associates, Ltd.

Sponsors

U.S. Fire Adminsitration

BRK Brands/First Alert

GE Security

Honeywell

InnovAlarm

Kidde

National Electrical Manufacturers Association

Siemens Building Technologies

SimplexGrinnell

Investigation of auditory arousal with different alarm signals in sleeping older adults

Dorothy Bruck,* Ian Thomas,[#] & Ada Kritikos,*

*School of Psychology, Victoria University
[#]Centre for Environmental Safety and Risk Engineering (CESARE)

Victoria University, Australia

May 2006

VICTORIA UNIVERSITY | **A NEW SCHOOL OF THOUGHT**

Table of Contents

List of Tables

List of Figures

Acknowledgements

Executive Summary

Over the last decade research on which emergency signal will best awaken sleeping individuals has led to a recognition that more work is needed on the audibility of existing smoke alarms and the comparative waking effectiveness of alternative signals. This research focuses on these issues in a population known to have an elevated risk of dying in a fire, adults aged over 65 years. It investigates responsiveness to different signals in sleeping older adults as well as measuring performance upon awakening (sleep inertia). This comparison of arousal thresholds required a tightly controlled experimental design, with selection criteria and methodological requirements that increase the validity of such comparisons using a manageable sample size, but do not allow direct extrapolations to the field in terms of expected arousal thresholds in a real emergency or percentages of the population that may awaken to certain signals. These population and methodological factors probably result in the research to date *underestimating* the proportion of people who will not wake up to an alarm.

Aims and the relevant findings are set out below, followed by a discussion of the key conclusions and recommendations.

Responsiveness to signals:

Arousal thresholds to different sounds were determined by playing auditory signals to the participants (aged 65-85 years, n=42) when they were in deep sleep (slow wave sleep). Each signal was presented with a stepped increase in volume from 35 dBA to 95 dBA and a bedside button was pressed by the participant to indicate awakening. The same participants received all four signals over two nights.

Aim 1: To investigate the arousal thresholds from sleep in older adults (aged 65- 85 years) to the current US smoke alarm (a high frequency T-3) and compare these thresholds to several alternative signals. The three alternative signals were a mixed frequency T-3 signal, a male voice (saying Danger, Fire, Wake up) and a 500 Hz pure tone in a T-3 pattern.

7

The first hypothesis was that the older adult sample would have significantly higher auditory arousal thresholds to the high pitched T-3 signal than to the two signals of mixed frequency (the mixed T-3 and the male voice). This hypothesis was only partially supported, with the results showing that the volume needed to wake up to the high T-3 was significantly higher than that needed with the mixed T-3. The most important findings were that,

(a) the **older adults needed a lower volume to wake to the mixed frequency T-3 signal** (median = 45 dBA) than to the other three signals tested (male voice, 500 Hz T-3 and high T-3), and

(b) the **current high frequency T-3 needed the highest volume** (median= 65 dBA) to produce awakenings compared to the other signals.

The second hypothesis was that the older adult sample would have significantly lower arousal thresholds to all signals than a young adult sample tested under similar conditions. Mean values showed differences in the predicted direction for both the mixed and high T-3 signals but only for the mixed T-3 was the difference across age groups significant. Surprisingly, for the male voice signal the young and older adults woke to similar volumes. Individual responses from three participants of non-English speaking background (NESB) suggested that a voice alarm with English text would not be suitable for them, although the inclusion of such NESB people did not cause the overall poor performance of the voice alarm with the older adults. Overall, these results indicate that **for older adults a male voice alarm would not be a suitable alternative.**

Sleep Inertia: This study was the first to assess older adults on several cognitive and physical tasks after awakening, and compare such performance to pre–sleep (baseline) levels.

Aim 2: To investigate the performance abilities of older adults when awoken suddenly by an alarm. This sleep inertia was assessed in terms of their simple and complex cognitive functioning and physical performance (with the latter involving a psychomotor task plus getting out of bed and walking 15 metres).

8

The results suggest that a **decrement in physical functioning of around 10-17%** may be expected across the first five minutes after awakening. **No important effects on simple or complex cognitive functioning were evident.** There was a wide variation in performance across individuals, with perfomance under baseline conditions strongly predicting performance under sleep inertia conditions.

Conclusions and Recommendations:
The present study, using a rigorous design and sufficient sample size of sleeping adults aged over 65 years, has found a substantial difference in the median auditory arousal threshold of 20 dBA between the current high frequency T-3 and the best performing alternative signal tested. Thus all the available data testing the waking performance of smoke alarm signals shows that a high frequency alarm signal[1] performs the *most poorly* of the alternatives tested for waking all the different population groups tested so far (i.e. children, sober and alcohol intoxicated young adults, older adults aged over 65 years). The evidence is sufficient to lead to the following recommendation:

Key Recommendation: The high frequency alarm signal currently found in smoke alarms should be replaced by an alternative signal that performs significantly better in awakening most of the adult population, once the nature of the best signal has been determined.

The findings of the current study, together with previous literature, indicate that a mixed frequency T-3 signal has *performed significantly better* than a high frequency signal in its ability to awaken sleepers in every sample group tested so far. This includes children, young adults (sober and alcohol intoxicated) and older adults. Voice signals appear to be as effective as the mixed T-3 in the children and young adult groups, but are less effective than the mixed T-3 in the older adults.

---

[1] A high frequency signal is typically used in all smoke alarms, the literature reported here has variously tested both a high frequency T-3 signal or continuous pulsing high pitched beeps.

The implications of introducing a signal frequency recommendation into the standards for smoke alarm notifications are considerable, involving a retooling of the entire industry. In view of this, any signal change that is mandated must be done on the basis of rigorous evidence that the best signal has in fact been found. The research is not yet at this point. A brief outline of suggestions for future research is set out below. These may take two to three years to complete.

In the meantime there are some recommendations that can increase the chance of sleeping individuals waking to a fire.

(a) Encourage interconnected alarms. Interconnected alarms that include an alarm in each bedroom will mean that the volume at the pillow is likely to be above 85 dBA.

(b) Consider the special emergency awakening needs of "normal hearing" older adults. Given the hearing thresholds for high frequencies of older adults it is inadequate to require their current high frequency smoke alarm to be a minimum level of 75 dBA at the pillow. The current study shows that those aged over 75 were particularly poor at waking to the current high T-3 (median of 70 dBA for high T-3 compared with 40 dBA for the mixed T-3). One possibility would be to recommend that older adults should have interconnected alarms, or at the very least stand alone alarms (with the current signal) in their bedroom. An additional, more satisfactory, possibility is for smoke alarm manufacturers to market special alarms for this age group that emit a mixed T-3 signal and suggest placement, as a minimum, in the bedroom.[2]

The *future research* that should be completed prior to the mandating of a specific signal encompasses a variety of issues.

(a) Research is needed to determine the optimal pitch and pattern of an alternative signal to wake people up, using a single convenient population, such as young

---

[2] Such a mixed frequency alarm would also be beneficial for individuals of any age who know they have high frequency hearing loss.

adults. The option of a voice alarm should no longer be considered for adult populations. Alternative pitches and pitch patterns should be investigated within the T-3 temporal pattern, at least in the first instance.

(b) Once several signals have been shown to have the lowest auditory arousal thresholds in the one population tested, they need to be tested in other sleeping populations, especially those most at risk of dying in a fire or of sleeping through an alarm signal. The signals should also be tested for salience and/or urgency as an emergency notification signal requiring action in awake individuals.

(c) Because of the inability to generalise data from the current study to field estimates, further research is needed using large numbers of non-primed, unselected groups to yield population based estimates of waking effectiveness. It seems most likely that the research to date may be significantly *underestimating* the proportion of people who will not wake up to an alarm. This arises from a range of factors, including the important fact that almost all of the participants in the relevant empirical studies on alarms and sleep have been primed to expect that a signal will go off on one of several nights.

(d) A study characterising the spectral characteristics of the background noises in a range of "typical" bedrooms would be informative and relevant. The extent of possible masking can be determined by combining this information with the acoustical characteristics of the signal that is most likely to awaken sleepers.

1 Introduction

Around the Western world the number one priority for residential fire safety has been promotion of the installation of smoke alarms. However, when residential smoke alarms were first developed and widely distributed in the 1970s the focus was on the technology to detect heat and/or smoke and little attention was paid to the nature of the audible signal. A high frequency signal was easily generated by a small piezo device and this was included as the standard alarm signal. As noted by Berry (1978), the issue of the audibility of fire warning equipment was relegated to an Appendix of the NFPA (74-1975) and the assurances about the ability of the signal to awaken people that were provided in the Appendix were at variance with the published auditory threshold data available at the time. Fire code standards include specifications of the volume that the alarm must emit, typically as a range of volumes which are above the ambient sound pressure level (e.g. 10 dBA above ambient, and within the range of 65-105 dBA, AS1670.1-2004). Recommendations about the volume that the alarm must be received inside a bedroom were added and these are generally 75 dBA (e.g. USA, Canada and Australia) at the pillow. A caution that this level may not be adequate to awaken all sleepers is often included (e.g. AS1670.1-2004). ISO 8201 "Acoustics- Audible Emergency Signal" defined a temporal three pattern (T-3) in 1987 and this was adopted by the NFPA in July 1996 (and later by many other countries) as the required fire notification signal, including in smoke alarms. No recommendation as to a frequency level of the sound is included.

The U.S. Consumer Product Safety Commission initiated a project in 2003 (Lee, Midgett, & White, 2004) to review the sound effectiveness of residential smoke alarms, with a focus on children (who had been shown to not reliably awaken to a smoke alarm) and older adults (who have death rates in residential fires of more than twice the national average). Among the recommendations was the need for further research examining what deficiencies exist regarding the **audibility of current smoke alarms**. Furthermore, previous research has raised the possibility that an alarm of a **different frequency and/or different sound** may be more effective for waking sleeping individuals.

This project empirically investigates both issues with regard to sleeping individuals aged 65 to 85 years. The results may have implications for the development of a more effective alarm signal for smoke alarms. The study also examines increased cognitive confusion and performance impairment (sleep inertia) that may influence effective and timely evacuation behaviour upon awakening in an older adult population.

2   Review of Literature

2.1   Signal significance and characteristics

Contrary to popular belief the brain does not "shut down" during sleep. During sleep we continue to monitor the environment and selectively respond. Discrimination between different signals clearly occurs during sleep, showing that the arousability of an auditory signal is not simply a function of how loud it is. Because cortical analysis of the meaningfulness of a signal precedes arousal, people respond selectively to signals, depending on the level of significance to them. An early study found that sleeping participants responded more often to their own name than to other names (Oswald, Taylor & Treisman, 1960). Significance can be added to a signal by "priming" the person to respond to some signals (e.g., a doorbell), but not to others (e.g., a telephone). When participants were primed to respond to a certain signal presented during the deepest stage of sleep, awakenings increased from 25% to 90% (Wilson & Zung, 1966). Clearly, signal significance and interpretation will affect arousal likelihood and thus it is important that any emergency signal has a unique sound quality that allows it to be readily identified and easily discriminated from other electronic beeping sounds in our environment (car alarms, mobile phones, microwave ovens, etc.).

It has been found, using functional MRI technology (Portas, Krakow, Allen, Josephs, Armony & Frith, 2000), that sounds that have an emotional significance have lower arousal thresholds and an increased probability of waking up a person. The involvement of a central nervous system "pathway of learned fear" has been suggested, with a key implication being that during sleep the emotional content of a signal may be processed independently of cortical input about the meaning of the signal. Thus the use

13

of sounds which arouse our emotions, such as a voice conveying an urgent message, may be an important consideration in emergency signals.

There is now an important body of literature about auditory alarms signals and their interpretation by individuals when awake (Edworthy, Loxley & Dennis, 1991; Edworthy and Stanton, 1995) and this has lead to design criteria suggestions to improve the effectiveness of emergency notifications in awake populations. It has been reported that signals that produce the highest ratings of perceived urgency were those with a higher frequency, a fast speed (tested across 0-500 msec), and a high level of loudness (Haas and Edworthy, 1996). The frequencies tested were across the range of fundamental frequencies from 200 Hz to 800 Hz, where each had higher component frequencies. The one that was perceived as most urgent had a fundamental frequency of 800 Hz with components of 800, 1600, 2400, 3200 and 4000 Hz .

A few studies have evaluated the alerting capabilities of alarms that are not auditory, specifically strobe lights and vibrating tactile devices located on the bed (Bowman, Jamieson & Ogilvie, 1995; Ashley, Du Bois, Klassen & Roby, 2005) especially in the context of emergency arousal for the hearing impaired. These devices are beyond the scope of the current literature review and research, which will focus exclusively on different auditory emergency devices. One reason for this selectivity is that auditory alarm devices are likely to be much lower in cost. Four types of alarm signals will be considered in this review; the high frequency beeping alarm, the Temporal 3 pattern, voice alarms and naturalistic sounds. Note that the literature evaluating their differential waking capabilities will be reviewed in Section 2.3.

A **high frequency beeping noise** is the most widely available smoke alarm signal and was most likely chosen for residential smoke alarms as high frequencies are rare in the normal environment, so they are likely to be more easily differentiated from other sounds. In addition they are subjectively piercing, not easily ignored and small battery operated devices can easily generate such sounds. Most residential smoke alarms emit beeps of a single high frequency which may be between 3000 Hz and 5000 Hz (Nober, Peirce & Well, 1981a; Ball and Bruck, 2004a; Ashley, Dubois, Klassen and Roby, 2005)

14

with a sound intensity in the vicinity of 85 dBA at 10 feet (the latter is a requirement in the US per UL217). Earlier smoke alarms sometimes combined two modulating signals peaking at 2000 Hz and 4000 Hz (Kahn, 1984).

A high frequency signal, however, appears to have several drawbacks. The most obvious disadvantage is that those with high frequency hearing loss (a part of normal aging) will have more trouble hearing the signal (see Section 2.3.3). A further disadvantage of a high frequency signal is that high frequencies are more easily reduced by doors and walls than frequencies below 500 Hz. This reduction occurs because walls reflect the energy from high frequencies rather than transmit it. For low frequencies more energy is transferred through the wall rather than being reflected. Thus, sound reduction is lower at low frequencies and higher at high frequencies (e.g. above 2000 Hz). Figure 2.1 shows transmission losses in dB as a function of the frequency of the sound and the surface mass of the material (e.g. a wall). It can be seen that transmission losses vary by about 20 dB for material of the same surface mass, depending on whether the frequency of the sound is low or high.
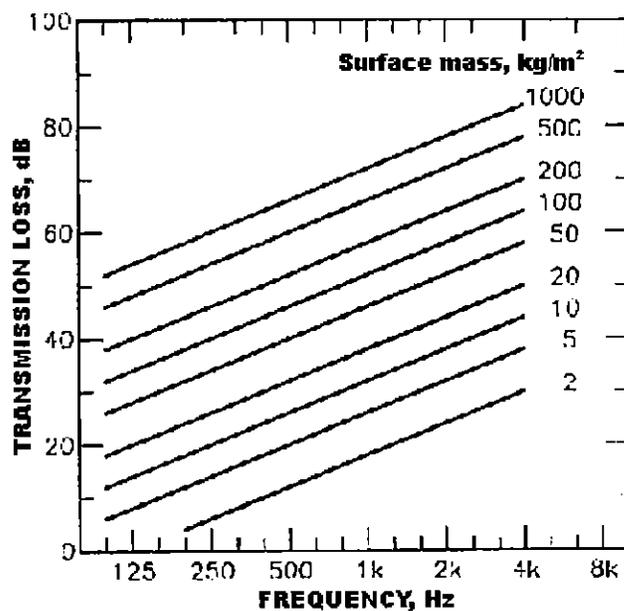
Figure 2.1: Sound transmission losses as a function of frequency of the sound and surface mass of the material through which the sound is being transmitted (from Quirt, 1985).

Robinson (1986) reported that the sound loss from the corridor to the room with the door open was about 12 dB for all frequencies above 500 Hz, with the closure of a door typically contributing another 15 dB, increasing to 20 dB if the door was edge sealed. This data suggests it would be impossible for a 90 dB smoke alarm located in the hallway to reach the pillow at 75 dB if the door was closed. Similarly, others have reported that a hallway smoke alarm will penetrate a closed bedroom door with a resulting bedside volume of between 51 and 68 dBA, depending on the room configuration and materials (Nober, Peirce & Well, 1981b). More recently, Lee (2005) completed a study on the audibility of smoke alarms and noted that bedroom doors attenuate a smoke alarm signal by about 10 dBA, while each home level attenuates the signal by about 20 dBA.

Clearly an alarm signal needs to be louder to awaken a sleeper if significant background noises, such as air conditioners exist (see Section 2.3.1). Masking occurs when the presence of one sound inhibits the perception of another. The greatest masking occurs when two sounds are similar in frequency. Importantly, a signal with multiple frequency components is less likely to be masked than one with fewer frequency components (Lawrence, 1970).

The unimpaired human ear is not equally sensitive to sounds at all frequencies and it is especially sensitive to frequencies between 1000 Hz and 3000 Hz when awake. However, as the change in sensitivity with frequency is most notable at reduced sound intensities, especially below 55 dBA (Lawrence, 1970), this may not be a major issue in determining the optimal frequency for an alarm signal. (Where industry recommendations and standards are for a minimum alarm sound intensity of 75 dBA at the pillow.)

In various Western countries (including the US and Australia, but not Canada) smoke alarms are now being sold which emit the **Temporal- Three (T-3)** pattern. The International Standard ISO 8201 -1987 (Acoustics – Audible Emergency Evacuation Signal) defines the T-3 signal and the specific temporal pattern of the T-3 is as shown in Table 2.1. The International Standard does not limit the smoke alarm signal to any one

sound, so signals of different frequencies and acoustic characteristics can be used within the T-3 pattern. The aim is that people will recognise the specific timing pattern as the signal to evacuate immediately.

Table 2.1: One cycle of the temporal pattern of the T-3 evacuation signal.

SIGNAL ON	0.5 sec
SIGNAL OFF	0.5 sec
SIGNAL ON	0.5 sec
SIGNAL OFF	0.5 sec
SIGNAL ON	0.5 sec
SIGNAL OFF	1.5 sec

One study (Proulx & Laroche, 2003) set out to assess people's recollection and identification of the T-3, as well as how urgent the signal was perceived to be. Results showed the T-3 was rarely identified as a smoke alarm or evacuation signal and was not judged as conveying urgency. The T-3 was usually judged to be a domestic signal, such as a busy phone tone.

There is a considerable body of literature about the possible use of the **human voice** in alarm signals. The appeal lies in the fact that a voice message can directly convey both meaning and emotional significance. Individuals hearing voice messages can successfully identify the emotions intended (Banse & Scherer, 1996). Moreover, the words used and the manner in which the words are spoken can influence their believability, appropriateness and sense of urgency (Edworthy, Clift-Matthews, & Crowther, 1998; Hellier, Edworthy, Weedon, Walters & Adams, 2002). It has been argued that humans have a particular cognitive specialisation for speech perception (Liberman & Mattingly, 1989). Phonetic perception may be immediate, with no translation of patterns of pitch, loudness and timbre being necessary. Language, unlike other forms of communication, may operate at a level that is precognitive. If this is the case when awake, then humans may also be particularly tuned to speech sounds during sleep.

A higher pitch is associated with a more intense emotion (Bachorowski & Owren, 1995), and the female voice is correspondingly assessed as more urgent than a male voice (Hellier *et al.,* 2002). Infants have been found to be selectively more responsive to tones at lower frequencies (Weir, 1976), perhaps because these are associated with human speech. The parameters of pitch of human speech show it to be a complex sound, generally below 2500 Hz . While prerecorded voice messages have been found to be helpful in encouraging people to evacuate, studies of warnings in large public spaces such as train stations (Proulx & Sime, 1991) show that a live directive voice announcement is highly effective. Clearly, such an announcement overcomes people's concern that it might be a false alarm. The key disadvantage of a voice alarm is that the signal must be designed to meet standards for both audibility and intelligibility (Grace, Woodger & Olsson, 2001). In addition, the speakers required to produce a quality, loud voice may not be able to be housed in the current small smoke alarm units.

Innovative research has used Gibson's theory of perception (Gibson, 1979) and information processing to test whether alarms that closely match their **naturalistic** intention or meaning are more effective than the more usual beeping signals. In an intensive care ward within a hospital, alarm signals were developed that closely matched the emergency situation they were aiming to alert staff about (Stanton & Edworthy, 1998). It was found that the naturalistic alarm signals were more effective than the standard signals in alerting novice medical staff who had little or no training of the standard signals. Building on this research, Ball and Bruck (2004b) set out to design a more meaningful, perhaps also emotional, signal. The first stage of this was to ask people which sounds would (i) make them feel a negative emotion, (ii) draw their attention when sleeping, and (iii) make them feel the need to investigate upon awakening. Collating 1447 responses showed that for all three questions people overwhelmingly nominated sounds within three categories; expressions of *human emotion* such as a baby crying or a person screaming, *manufactured alerting sounds* such as a smoke alarm, and other sounds that may *naturalistically* alert them to the possibility of danger, such as the sound of footsteps. Two new sounds (conveying either emotional and naturalistic signals) were developed with the aim of testing their ability to awaken sleeping people in a fire emergency. As the *naturalistic* sound needed to be

situationally congruent and indicate a fire, a signal consisting of house fire sounds (fire crackling, roaring and popping, together with glass breaking) was developed. For a signal conveying human *emotion*, ethical considerations ruled out using genuine sounds of human distress. The second signal developed was a female actor's voice conveying human emotion through an urgent voice tone and choice of words (danger, fire etc). The testing of these signals is described in Section 2.3.1 below.

Naturalistic fire cues were also used in a study (Bruck & Brennan, 2001) with the aim of determining whether adults would awaken to low level fire cues, including two auditory cues. Both the crackling sound of a fire and a "shuffling" sound (as reported by fire survivors) were presented to sleeping individuals at very low levels (received at 38 to 48 dBA) and a relatively high rate of arousal was found (91% to crackling and 83% to shuffling).

It is not unusual for smoke alarms in buildings to move through a **signal shift**, or a series of different signals, such as beeping tones with different temporal and frequency patterns and whooping tones. Although it has not previously been investigated, anecdotally such shifting makes sense, as a signal that is constantly changing is likely to attract attention (when awake or asleep). We know that sometimes people can sleep while a TV is on, only to wake up when it is turned off. The change in auditory signal, even to silence, may induce arousal.   Moreover, studies of auditory arousal thresholds (see below) consistently note major individual differences in thresholds and it is possible (but not established) that different people may respond better to different signals and shifting signals increase the chance that one of the signals will be perceived more easily by some people and acted upon. To date only one study (Ball & Bruck, 2004b) has tested the efficacy of a signal shift pattern in sleeping individuals and this will be discussed below in Section 2.3.1

2.2 Human characteristics

There are a wide range of factors that affect the auditory threshold of a person while asleep. These have been discussed in some detail in two earlier review papers (Bonnet, 1982; Bruck, 2001) and only the most relevant and important points will be

19

summarised here. In this section discussion will focus on research using signals that are not emergency alarms, such as pure tones. Alarm research and sleep will be reviewed in Section 2.3 below. The literature shows that the issue of what will wake different people under different circumstances is complex.

Of all the possible variables it seems that **individual differences** account for the most variability in auditory threshold. One study examined responsiveness to a 5 second 800 Hz tone during sleep (Zepelin, McDonald & Zammit, 1984) in people in various adult age categories, across three different stages of sleep (REM, stage 2 and stage 4). It was found that the thresholds varied for each age and sleep stage data point by at least 54 dBA with the largest range being 82 dBA (i.e., range from 39 dBA to 121 dBA for people in their 40's being awoken from stage 2, see Table 2.2). It is known that people's individual susceptibility to being awoken is quite consistent from night to night and within a night and that those who tend to sleep more deeply will do so in every stage of sleep, relative to those who sleep more lightly in all stages of sleep (Bonnet, Johnson & Webb, 1978). Moreover, once an individual is asleep, the issues of whether they are a good or poor sleeper (i.e., awaken frequently) do not appear to be an important variable (Johnson, Church, Seales & Rossiter, 1979).

Age is likely to be the next most critical variable, with major differences between the arousal thresholds of children, middle-aged adults and elderly individuals. Older people are likely to awaken more easily than younger people and children are generally the hardest to arouse (Busby, Mercier, & Pivik, 1994; Zepelin et al., 1984). Table 2.2 shows the gradual reduction in arousal thresholds across three different adult age groups, in both stage 4 and stage 2. Zepelin et al. (1984) found that the decline was sharpest in stage 4 sleep, but occurred in stage 2 and REM as well. The authors concluded that age was not as influential as individual differences in accounting for the auditory arousal threshold levels, but age differences were nevertheless substantial, with the decline becoming evident by the 40s.

20

Table 2.2: Auditory awakening thresholds (dB) to a 5 second 800 Hz tone at three different age levels by stage of sleep (n=52). Data from Zepelin *et al.* (1984).

		Stage 4	Stage 2
18-25 yrs	mean	101	82
	standard deviation	17	22
	range	49-116	45-121
40-48 yrs	mean	87	71
	standard deviation	20	22
	range	59-116	39-121
52-71 yrs	mean	71	61
	standard deviation	20	13
	range	39-116	44-98

There may be several factors operating that mean arousal thresholds decline with advancing age. Perhaps the most important is the age related change in electroencephalogram (EEG) energy levels (based on power spectrum density) within sleep. Adult EEG energy levels (documented across ages 18 to 43 years) show a decline with increasing age (Astrom & Trjaborg, 1992). Secondly, the duration of the deeper parts of sleep (slow wave sleep, SWS, consisting of both stage 3 and 4 sleep) reduces with age so that younger adults spend more time in SWS than older adults. The decrease is especially evident in the amount of stage 4 sleep in the older individuals and more so in men than women. In some cases stage 4 may disappear in people over the age of 60 (Carskadon and Dement 2000). A recent meta-analysis concluded that the minutes of SWS decline with age such that at age 65, 75 and 85 we could expect 67 min, 50 min and 25 min respectively of SWS (Ohayon, Carskadon, Guilleminault, Vitiello, 2004).

The ability to be awoken in different **sleep stages** varies. Stages 3 and 4 are subjectively the deepest part of sleep and predominate in the first third of a night of sleep. Most studies show (see Bonnett, 1982 and Section 2.3 below) that it is harder to

21

arouse a person from stage 4 compared to all other sleep stages and that arousal thresholds are approximately equal in stage 2 and REM. However, the average difference in decibel level needed to awaken an adult in different stages may not be substantial. For example, Table 2.2 shows that Zepelin *et al.* (1984) found mean differences across nine 52-71 year olds of only 10 dB between stage 2 and stage 4 sleep, while individual differences, as shown by the range values, are much greater (50-80 dB). Time of night differences, independent of sleep stage, do not appear to be robust (Bonnett, 1982).

Several studies have considered how **sleep deprivation** affects people's ability to respond to auditory signals when asleep. In some cases the experimental design relies on successful tone discrimination, or reaction time, rather than considering thresholds specifically. Performance is consistently reduced by sleep deprivation across a range of studies (Williams, Hammack, Daly, Dement & Lubin, 1964), even after just one night of partial sleep restriction to four hours (Synder & Scott, 1972). An early study (Lindsley, 1957) found that after 38 hours of sleep deprivation sleeping participants reacted to a tone less frequently than on control nights (only 600 times compared to 1500 times), suggesting increased thresholds. It is well known that sleep deprivation changes the architecture of sleep on the recovery nights, with considerably more stage 4 sleep in the first third of the night. It also seems likely that EEG energy levels increase across all sleep stages in recovery sleep, presumably making it harder to arouse the sleeper.

Most early studies found no significant **sex differences** in arousal thresholds. However, there were some exceptions. Wilson & Zung (1966) found more responsiveness among sleeping women than men to sounds they were motivated (by a reward) to respond to, while Zepelin *et al.* (1984) found a trend for older women to have higher thresholds than older men. The strongest evidence of a sex difference in arousability comes from the statistical modelling of arousal to low level fire cues (Hasofer & Bruck, 2004). Involving a total of 53 adults and using four different fire cues (crackling sound, shuffling sound, flickering light and smell) a statistically significant difference was found, with females showing a higher probability of waking to each cue than males. A trend was also noted for the mean response time to awakening to be

shorter for females. A subsequent study involving smoke alarm signals and alcohol consumption also found similar significant sex differences (see Section 2.3.3 below).

One study has considered the effect that a dose of **hypnotics** (flurazepam 30 mg) may exert on arousal to pure tones (Johnson, Church, Seales & Rossiter, 1979). When the drug was exerting its maximum effect (some two to three hours after ingestion) the auditory threshold was approximately 30 dBA higher on drug nights compared to placebo nights. There are no published studies available on arousal thresholds to sounds that are not alarms after consuming other drugs, such as alcohol or marijuana. Studies testing responsiveness to smoke alarms after intake of different soporific substances, including alcohol, are described in the next section.

2.3 Awakenings with various alarm signals

Within the published literature there are a comparatively small number of studies considering arousal from sleep to an auditory emergency signal and most of these have involved the high frequency smoke alarm signal (continuous beeps rather than the T-3 unless otherwise specified). Several recent studies have compared this high frequency signal with a small range of different signals. These studies will all be reviewed here, in three categories;

- adults (where the studies have used samples of unimpaired adults or where any factors which may have impaired their arousal, such previous late nights or drinking, were not systematically manipulated);
- children;
- adults impaired by hypnotics, alcohol or hearing difficulties.

2.3.1 Adults

The first study to consider the issue of whether people would wake up to a smoke alarm was by Nober *et al.* (1981b). It was found that all 30 of the 18 to 29 year old male participants were able to wake up quickly (within 21 seconds) to a high frequency alarm presented in their homes at levels ranging from 55 to 85 dBA at the pillow. All the men even woke up when a 70 dBA signal was presented with a 53 dBA air conditioner noise in the background, although this took them up to 85 seconds. However, at the volume

23

of a hallway alarm (55 dBA) only 70% of the men awoke when the air conditioner was on. In a subsequent, similar investigation 12 males were tested in a laboratory (Kahn, 1984) using smoke alarms of 44, 54 and 78 dBA at the pillow, against a background noise level of 44 dBA. The percentage who awoke were 25%, 50% and 100% respectively. Both studies clearly showed the detrimental effect of background noise (causing masking of the alarm signal) and suggested the importance of placing the smoke alarm within the bedroom itself to facilitate awakening.

A decade later Bruck and Horasan (1995) exposed 24 young adults (18-24 years) twice to a 60 dBA alarm. The percentage who awoke to both alarm presentations varied slightly according to the sleep stage at the time of signal presentation, with 87%, 75% and 75% awakening consistently across stage 4, stage 2 and REM sleep respectively. Latency to awakening was longer in stage 4 than in the other two stages (79 seconds compared to 20 seconds or less). It was found that those participants who slept through one or both signals were sleep deprived, due to significant exam-pressure sleep-restriction on the night before the experiment. Thus all the participants were not 'unimpaired' and this introduced a confound into the study. Studies of adolescent and young adult sleep patterns (Carskadon, Harvey & Dement, 1981) show that it is not at all unusual for individuals in this age group to have highly irregular sleep patterns, alternating nights of restricted sleep hours with nights of recovery sleep.

In a subsequent study, Bruck (1999) set out to more thoroughly investigate the waking likelihood of adults (across a wider age range) and children in the setting of their family home. A high frequency beeping alarm was set up in the hallway of selected homes such that it reached the pillows of both parents and children at 60 dBA. The 16 parents involved were aged from 30 to 59 years and the equipment was in their homes for five nights. Individuals who participated in the study were screened carefully and asked to abstain from any alcohol consumption and keep regular sleep/wake hours. They were told the smoke alarm would be activated on two of the five nights but did not know more specific details. It was always activated in the middle third of the night (1 to 4.30 am). Impressively, all parents awoke on both nights within 32 seconds.

In a recent study (Ashley, Du Bois, Klassen & Roby, 2005) 32 people with established normal hearing were tested in a sleep laboratory across the sleep stages of slow wave sleep, stage 2 and REM. A high frequency smoke alarm (3100 Hz) in the T-3 pattern was presented for two minutes at 75 dBA and it was found that 96% of participants awoke.

A large scale study involving 621 sleeping Disaster Protection trainees staying in a hotel (Nakano & Hagiwara, 2000), found that 90% evacuated within 120 seconds, where 74% reportedly awoke to the 50-53 dBA hotel emergency bell, a further 9% awoke to the subsequent 60-67 dBA siren, 2% to the final 48-55 dBA voice broadcast and 8% were awoken by others. The degree to which these young men were unimpaired is hard to judge as 193 reported that they had "drunk very much" during the evening, while 70 "got dead drunk". Nevertheless, the reported rate of responding to the signals is high.

To date the only controlled studies of the response of sleeping adults to different alarm signals are by Ball and Bruck (2004a, 2004b). These studies adapted the method of limits procedure, whereby a continuous signal was presented via a bedside speaker, starting at the whisper volume of 35 dBA and increasing in 5 dBA steps to a maximum of 95 dBA. Signals at each volume were presented for 30 seconds and moved on to a higher volume if there was no response. The main variables of interest were the time to the pressing of a bedside button and the decibel level when the person awoke (auditory arousal threshold, AAT). Three signals were presented each night during stage 4 sleep. The participants were self reported deep sleepers aged 18 to 25 years and a repeated measures design was used to minimise the variability due to individual differences. Their first study was a pilot study (n=8) to determine the relative effectiveness of three newly developed signals in waking up participants. In Section 2.1 above the development of two signals presenting the naturalistic house fire sounds and the female actor's voice (conveying emotion) was described (Phase 1 of Ball & Bruck, 2004b). The third signal tested in the pilot study (Phase 2) combined these two signals, continuously presenting each for 5 seconds (i.e., a signal shift). In this small sample it was found that the female voice signal was significantly more effective than either the naturalistic house fire sounds or the signal shift in waking the participants up.

25

In a further similar subsequent study, the comparative effectiveness of the female voice (300 to 2500 Hz), high pitch alarm (4000 to 5000 Hz) and a mixed frequency T-3 alarm signal (500 to 2500 Hz) were compared using 12 young adults (Ball & Bruck, 2004a). Based on the literature suggesting that signal significance was an important component in facilitating arousal, the researchers were expecting the human voice to be the most effective in waking participants up. However, it was found that the AATs for the female voice and the mixed T-3 alarm were similar and significantly lower than for the high pitch alarm (see Figure 2.2 in Section 2.3.3 below - sober condition).

A subsequent pilot study specifically compared responsiveness to a male voice with a female voice in a small sample of 10 young adult participants using a repeated measures design. (M. Ball & D.Bruck, 2005, unpublished data). The mean AAT for the female voice was 61.0 dBA (S.D.=18.1) and to the male voice, 52.5 dBA (18.3). Due to the small sample size this difference did not achieve statistical significance but six of the subjects found the male voice more alerting than the female voice at a lower volume, three equal and only one person was more easily alerted to the female voice. It was concluded that with an increased sample size it was likely that the male voice would yield significantly lower AATs than the female voice.

2.3.2 Children

The first study to suggest that children may not be effectively aroused by a smoke alarm assessed awakening using a hallway high pitched beeping alarm, which reached the pillow at 60 dBA (Bruck, 1999). Of the 20 children aged from 6 to 15 years, only 6% awoke on both nights when the alarm was presented. When the volume of the signal was increased to 89dBA at the pillow, the percentage who reliably awoke increased to 50% (Bruck & Bliss, 2000). However, the younger children (aged 6-10 years) were clearly more at risk, with only 29% within this age subset reliably awakening to 89 dBA. The researchers went on to consider the ability of this 6 to 10 year old age group to awaken to different signals, all presented at the volume of an alarm installed above their bed (89 dBA). Across several studies using a similar methodology Bruck and Ball

(2004) found that significantly fewer children awoke to the high frequency alarm compared to two voice alarms or the mixed frequency T-3 (see Table 2.3). [3]

Table 2.3. Number of children who awoke within different time categories to different alarm signals (from Bruck & Ball, 2004).

	N	Valid alarm present-ations[4]	0 - 30 sec	31 - 60 sec	60-180 sec	Awoke within 180 seconds but exact time not known[5]	Did not wake	% Total awake
mother's voice	20	19	15	4	0	0	0	100%
female voice	20	19	12	5	0	1	1	94%
high pitch alarm	14	28	10	1	4	1	12	57 %
mixed T-3	14	28	14	7	0	6	1	96%

The voice alarms consisted of either the child's own mother's voice (saying their name about once every 6 seconds) or a female actor's voice (as used in Ball & Bruck, 2004a and 2004b). Table 2.3 shows that significantly more children awoke to both the voice alarms and mixed T-3, compared to the high pitch alarm. In addition, the children awoke more promptly to the voice alarm and T-3 signal compared to the high pitch alarm and this difference was also significant.

2.3.3 Impaired adults

It is not surprising that the intake of hypnotics substantially reduces the ability to wake to a smoke alarm. Only one study has examined this effect experimentally (Johnson,

---

[3] The comparisons for this age group between the high frequency signal and the mixed T-3 were not repeated measures on the same children.
[4] i.e. the child reported retrospectively that they were asleep before the alarm was sounded
[5] This was due to technical difficulties with the wrist actigraphs.

Spinweber, Webb & Muzet, 1987) and found that 50% of the adults receiving the hypnotic, triazolam (0.25 or 0.5 mg), did not awaken to three one minute 78 dBA alarms, presented during deep sleep when the drug was exerting its maximum effect (2 hours post ingestion). This compared to 100% awakening with the placebo. With 35 million prescriptions for sleeping medications in the US in 2004, the arousal thresholds of many individuals are regularly substantially impaired, with the elderly disproportionately likely to take hypnotics (Medco Health Solutions 2005).

Despite the strong association between fire fatality and **alcohol** consumption (Sekizawa, 1991, Brennan 1998) the ability of intoxicated people to awaken to a smoke alarm has only recently been investigated. Arousal thresholds to three different alarm signals were explored in 12 young adults under three different levels of alcohol intoxication: sober, 0.05 Blood Alcohol Content (BAC) and 0.08 BAC (Ball & Bruck, 2004a).

Figure 2.2 shows that responsiveness to both the female voice and the mixed T-3 were very closely matched, and both signals aroused individuals at a mean sound intensity that was lower than the high pitched signal. It also shows the substantial increase in magnitude required for all signals when alcohol was administered. The research followed the modified method of limits procedure described earlier, so the time taken from the first 30 seconds, 35 dBA signal presentation to when the participant responded with a button press was a key dependent variable. Analyses showed that both the difference between the sounds and the difference between the three alcohol conditions were statistically significant (MANOVA).
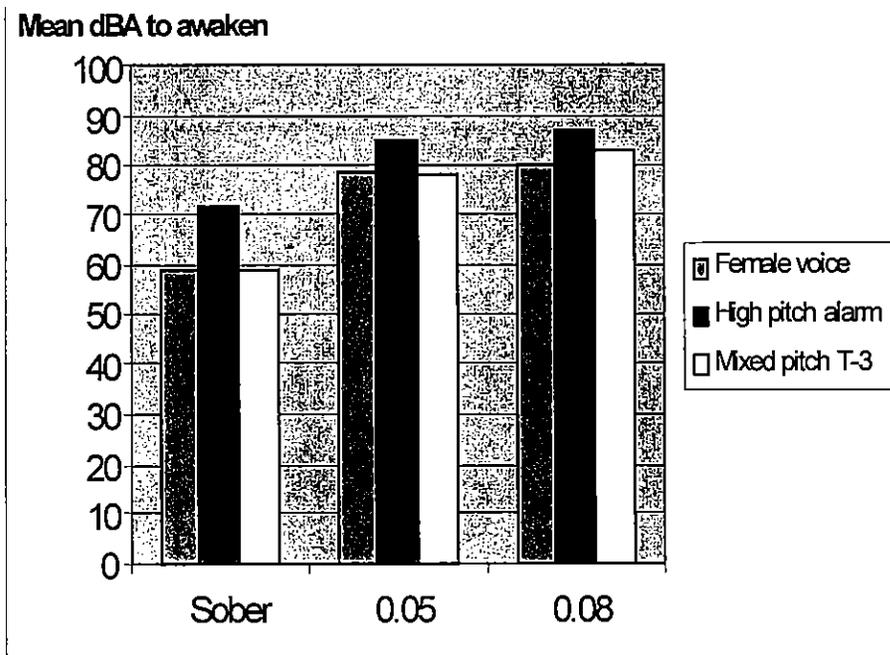
Figure 2.2: Comparison of auditory arousal thresholds (AATs, mean dBA levels of different alarms required for waking) of young adults under different blood alcohol conditions (n=12) (from Ball and Bruck, 2004a).

Further analyses of the above data, applying a sophisticated stochastic random walk model (Hasofer, Thomas, Bruck & Ball, 2005) enabled predictions to be made about arousal, given a certain signal and certain individual characteristics. The modelling showed that both the estimated recognition probability and estimated waking up threshold of the various alarm signals is consistently different for females than for males, indicating greater sensitivity to the signal in sleeping females than males when both have the same BAC.

As auditory smoke alarms are by far the most commonly installed smoke alarm, and are compulsory in many countries of the world, the issue of which type of signal is most likely to be heard by those with the most common types of **hearing impairment** arises. It is not simply a case of an increased volume being more effective. The most common type of hearing loss is that associated with advancing age, with US census data (Lucas, Schiller & Benson, 2004) suggesting that 14% of the population is hard of hearing. Considering only an older group, 46% of 48-92 year olds (n=3753) were found to have some hearing loss (Cruickshanks, Wiley, Tweed, Klein, Mares-Perlman, & Nondahl,

1998) with older people most likely to lose their sensitivities to higher frequencies first (and males more so than females). Figure 2.3 shows that hearing thresholds (when awake) for a tone at 3000 Hz are much higher than for a 500 Hz tone. Thus in order for a 70 year old man to hear a 3000 Hz signal it would need to be over 30 dBA louder than a 500 Hz signal.
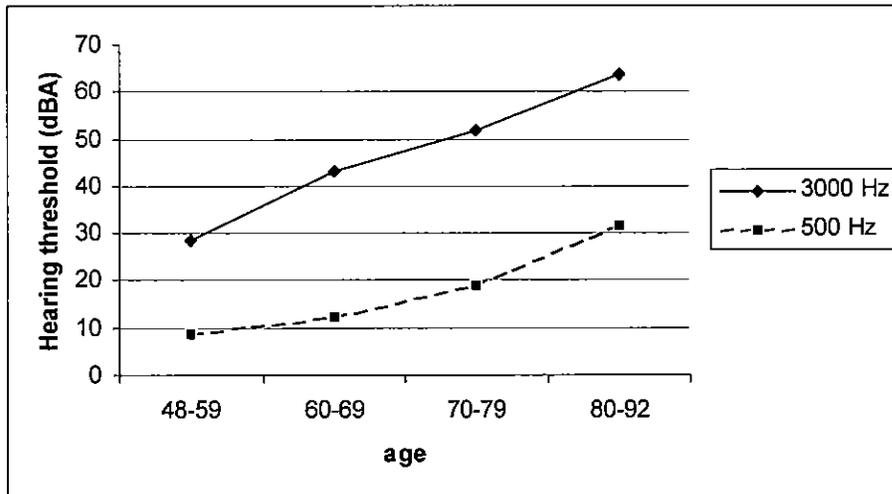
Figure 2.3: Hearing threshold values (dBA) for tones at two different frequencies for males of different ages (right ear) when awake (data from Cruickshanks *et al.*,1998).

In order to estimate the percentage of those aged 60-69 years who would not awaken to a hallway high pitched alarm (55-60 dBA alarm of 2000-4000 Hz), Bruck (2001) extrapolated from ISO 7029-1984 data on hearing threshold values. Using a derived 41 dBA difference between awake and asleep thresholds it was estimated that at least 25% of people in their 60s would not be awoken to such a hallway alarm. Many people are not aware that their ability to hear high pitched sounds is impaired with advancing age and assume that they will hear such a signal. In a study testing the waking ability of the hard of hearing, 39 hearing impaired individuals were exposed to an alarm during different stages of sleep (Ashley *et al.*, 2005). The hearing ability of these individuals was reduced by between 20 and 90 dBA over the frequency range of 250 to 8000 Hz . Across this group only 57% awoke to a 75 dBA 3100 Hz signal.

Some studies have considered the ability of individuals of different ages to hear sounds

(when awake) encountered in medical environments and as ringers for the home telephone. Wallace, Ashman and Matjasko (1994) tested the ability of anesthesiologists across ages 25 to 74 years to hear alarms in an operating room. They found that the inability to hear alarms occurred only with those alarms that had a frequency of 4,000 Hz or more and concluded that high frequency alarms may go undetected by the ageing human ear. Three acoustically different electronic telephone ringers were compared across 20-30 year olds and participants over 70 years of age (Berkowitz & Casali, 1990). For the older group it was found that signals with prominent low to mid range frequency components (1000-1600 Hz) could be more easily heard than higher frequency ringers (with peaks at 3150 and 20,000 Hz). The authors cite an early conference paper by Hunt (1970) which noted that the most effective ringers have at least two spectral components between 500 and 4500 Hz with a prominent component below 2000 Hz.

2.3.4 Summary of risk factors

Studies on auditory arousal from sleep have shown that most unimpaired adults will awaken quickly to quite low volume noises, including hallway smoke alarms. One conclusion is that sleep in "normal" populations is *not in itself* the major risk factor for fire fatality but that additional risk factors need to be present to substantially increase the chance of not waking to an alarm. The literature from the studies of smoke alarms and sleep tells us that significant risk factors include being a child, being under the influence of hypnotics, being alcohol intoxicated, being hearing impaired, being aged over 60 (for high frequency signals), being sleep deprived and having high levels of background noise. Females tend to wake slightly more easily than males but this difference appears to be subtle and overshadowed by major individual differences in auditory thresholds.

Importantly, it is not known whether there is consistency in which signal is most effective across different populations. The research so far has found that the lower frequency signals were more effective for children, sober adults and alcohol intoxicated adults. What is not yet known is whether the best signal for these groups is also the best signal for other groups, such as the elderly. No studies have been conducted to date to

31

investigate the extent to which older individuals will waken to the current smoke alarm signal, or how their responsiveness to other signals may compare. The first step should be to investigate such questions in a group of unimpaired older people, who are within the normal hearing limits for their age.

2.4 Sleep inertia

Sleep inertia effects operate as soon as a person awakes and lead to a decrease in performance. This decrease may be modest or considerable, with the person being very sleepy, confused or disorientated. Its manifestation is most dramatic when awakening from sleep has been abrupt, regardless of whether the sleep occurs at night or during a daytime nap (Dinges, Orne, Evans & Orne, 1981; Dinges 1989). The documented duration of sleep inertia varies with the performance tasks used to measure it (Akerstedt, Torsvall & Gillberg, 1989). Most studies of sleep inertia have used simple motor, automatic or attentional tasks (such as reaction time, arithmetic or vigilance tasks). No studies have been published investigating the sleep inertia of older adults.

From the perspective of the behaviours and cognitions required if awakening in an emergency, the most relevant tasks involve complex cognitive functioning, such as decision making and physical functioning. Bruck and Pisani (1999) found that sleep inertia reduced decision making performance for at least 30 minutes in young adults, with the greatest impairments (in terms of both performance and subjective ratings) being within the first 3 minutes after abrupt awakening. Decision making performance was reduced by 51% during these first few minutes, compared to baseline. During the first nine minutes the decrements were significantly greater if the person had been awoken from deepest sleep (slow wave sleep, stages 3 and 4) compared to REM sleep.

It has been argued that sleep inertia is not qualitatively different from sleepiness (Balkin and Badia, 1988) and both may reflect an incomplete disengagement from sleep processes. A recent study suggested that the Trail Making Task (TMT) may be an effective measure of sleepiness in the aged, with performance on the TMT differentiating between those who regularly napped during the day and those who did

32

not (Bliwise & Swan, 2005). The TMT has been used consistently in both clinical and experimental contexts over four or so decades, since its original inclusion in the Halstead-Reitan Neuropsychological Battery (see Reitan & Wolfson, 1985). The TMT is considered to be, overall, a test of executive functioning. It consists of two parts, A and B, which are completed consecutively in that order. Part A of the TMT is thought to measure psychomotor speed, whereas part B has been variously postulated to measure shifting of cognitive set, sustained attention and sequencing (Lezak, 1995). Because the TMT B is postulated to assess cognitive shifting as well as speed of processing, it is in ideal measure of ability to progress from one step (or idea) to the next under conditions of time pressure – such as those of an emergency. In other words, in situations where participants are required to progress, in strict sequence, and in a timely manner, from one step to the next to achieve a goal.

3   Research Aims

The study had two research aims. The **first aim** was to investigate the arousal thresholds from sleep in older adults (aged 65- 85 years) to the current US smoke alarm emitting the high pitched T-3 and compare these thresholds to several alternative signals. The signals and their rationale are set out in Table 3.1. The spectral analyses of all four signals are shown in Appendix A.

Table 3.1: Description and rationale for the four signals delivered in this study.

Pattern	Dominant pitch	Frequency (Hz)	Rationale
T-3	High	3000	As currently in smoke alarms sold in the US
T-3	Low	500	Help define optimal lower frequency
T-3	Mixed	500-2500	Similar frequency range as voice alarm & quite effective in previous research (see Section 2.3)
Male voice	Mixed	200-2500	Recent pilot work (see Section 2.3.1) suggested it was more alerting than the female voice alarm

A review of the literature informed the development of two hypotheses with regard to the comparative assessment of arousal thresholds.

Hypothesis 1: The older adult sample would have significantly higher arousal thresholds to the high pitched T-3 signal than to the two signals of mixed frequency (the mixed T-3 and the male voice).

Note: The inclusion of the 500 Hz pure tone was exploratory as there had been no previous research using such a tone. However, it was felt to be particularly valuable to determine whether a pure low frequency tone performed equally well, or better, than mixed frequency signals that incorporated 500 Hz levels.

Hypothesis 2: The older adult sample would have significantly lower arousal thresholds to all signals than a young adult sample tested under comparable conditions. (This applies particularly to the mixed T-3 and male voice signals where identical

comparisons are available.)

The **second aim** of the study was to investigate the performance abilities of older adults when awoken suddenly by an alarm. Their sleep inertia would be assessed in terms of their simple and complex cognitive functioning and physical mobility, with the performance assessments designed to have some face validity in terms of the skills and behaviours that may be used in an emergency.

Hypothesis 3: Compared to baseline levels, a complex performance task (Trail Making Task B[6]) completed under sleep inertia conditions would require an increased time to complete and include more errors.

Hypothesis 4: Compared to baseline levels, physical performance tasks (Trail Making Task A, getting out of bed and walking 15 metres) and a simple cognitive task (completing a phone call) assessed under sleep inertia conditions would require an increased time to complete.

---

[6] Trail Making Tasks are included in Appendix L

4 Methodology

4.1 Participants

Forty five adults aged 65 to 83 years were involved in the sleep research. The overall mean age was 73.1 years (standard deviation = 5.6). Table 4.1 shows the age and sex distribution of participants who completed the study. Not all 45 participants completed all aspects of the study. A total of 42 completed the section involving presentation of sounds. The distribution of those who did complete all four signals is shown in brackets in Table 4.1. Because some participants had difficulty with Part B of the Trail Making Task only 39 completed all trials, while 41 completed the other physical sleep inertia performance tasks on both nights. Further details of the age and sex distribution for different tasks can be found in Appendix B.

Table 4.1: Numbers of participants as a function of their age and sex. (Numbers in brackets refer to those who completed all four signals.)

Age	N males	N females	Total
65-74	13(12)	14(13)	27 (25)
75-85	9 (8)	9 (9)	18 (17)
Total	22 (20)	23 (22)	45 (42)

In the course of conducting the study five people passed their hearing test but then did not commence the study. Three people dropped out after completing two or three signals and their reasons are described in Appendix C.   Three participants had insufficient English to follow instructions and a member of their family volunteered to provide translations. They all spoke Arabic and two had sufficient knowledge of the Latin alphabet to complete the Trail Making Task, while the third was unable to do so.

Recruitment was conducted by a graduate psychology student, predominately through social groups. A report on the recruitment process is contained in Appendix D. All participants were paid $200 (Australian) for their involvement. Where recruitment was

from a social group, the group was paid $150 for each person who completed the study.

Inclusion criteria for participants were that they would
- be independently mobile (although use of a walking stick or walking frame was permitted),
- not be taking medication affecting their sleep,
- be cognitively capable (screened using the Mini Mental State if doubts existed)
- report that they normally do not have significant difficulties falling asleep, and,
- report that they considered their hearing to be average or above average for someone their age.

A total of 59 potential participants underwent the hearing screening test and nine failed (15%). The hearing of each potential participant was screened across five auditory frequencies (500, 1000, 2000, 3000 and 4000 Hz) by a professional audiologist from HEAR Service Victoria. The criteria was that they perform within, or better than, one standard deviation of the mean age and sex-matched normative threshold at each of the five frequencies in each ear (Cruickshanks, Wiley, Tweed, Klein, Klein, Mares-Perlman, & Nondahl, 1998). The criteria levels are shown in Appendix E. This meant that those who perform in the lowest 15.9% for their age and sex at any of the five frequencies in either ear were not included. A comparison of the mean hearing thresholds of the participants (when awake) with the mean values as found by Cruickshank *et al.* (1998) is also contained in Appendix E. Age and sex details and the different reasons why people failed their hearing test are contained in Appendix F.

4.2 Apparatus and Materials

Signals: Two sets of equipment were used. Each set consisted of portable sleep stage monitoring equipment (Compumedics Siesta), a laptop computer, two speakers to deliver the alarm signals and a hand held sound meters. The latter were professionally recalibrated immediately prior to the study. Full details of the process by which the multiple sound files were created for delivery during sleep, as well as information on sound measurement, calibration, and delivery can be found in Appendix G.

The origin of the four sounds are as below.

- Mixed T-3 was from Simplex 1996, 4100 Fire Alarm Audio Demonstration CD.[7]

- Male voice was recorded in a radio studio with a male actor, chosen for his particularly deep voice (see Appendix H for text).[8]

- High T-3 was recorded from a current US smoke alarm (Kidde).

- Low 500 Hz T-3 was generated by a computer program.

A spectral analysis of all four sounds can be found in Appendix A. Their frequency details are as follows.

- Mixed T-3 had a fundamental frequency of around 520 Hz (+/- 4 Hz) with odd harmonics (3rd, 5th etc.).

- Male voice had dominant frequencies in the range from 500 Hz to 2,500 Hz, with some additional frequencies from 2,500 to 4,000 Hz.

- High T-3 had a fundamental frequency just above 3000 Hz.

- Low 500 Hz T-3 was a pure tone of just below 500 Hz.

General Forms: The consent form and information sheet are in Appendix I. The demographics and screening form are in Appendix J. The questionnaire about prior sleep and alcohol consumption is in Appendix K.

Sleep Inertia: The apparatus for this aspect of the testing consisted of a 15 metre rope to follow when walking, a 'Stable Table' to write on during the Trail Making Task (TMT) and the participant's own phone and an answering machine.

The performance test used to measure *complex cognitive* sleep inertia was the TMT, which had both an A and B task. Examples of these materials are contained in Appendix L. The TMT was originally included in the Halstead-Reitan

---

[7] The use of the mixed T-3 in these sleep studies started out somewhat serendipitously, with a demonstration CD of a T-3 signal being obtained from Canada, so that the same signal was used in these sleep studies as in the Proulx and Larouche (2003) study.

[8] This recording was the same as used in a pilot study by Ball and Bruck, discussed elsewhere in this report.

Neuropsychological Battery (see Reitan & Wolfson, 1985). Apart from a substantial corpus of normative data, the TMT has been administered to clinical populations, including traumatic brain injury, dementia and early dementia, stroke, depression and psychosis. It consists of two parts, A and B, which are completed consecutively in that order. Both parts begin with a short practice sample. Part A consists of joining consecutively 25 dots, numbered clearly from 1 to 25. Part B consists of *alternating* between numbers (1 to 13) and letters (A to L); that is, 1, A, 2, B etc. Time to completion and number and type of errors are recorded. In this study, participants were tested on the TMT several times. To avoid practice effects, alternate forms were generated by mirror reversed and inversion versions of the original (canonical) arrays for both parts A and B (the originals are in Appendix L). This had the advantage of retaining all spatial relationships (and sequences) between the points intact.

4.3 Procedure

Volunteers meeting the self report selection criteria underwent a screening hearing test at a professional hearing clinic. Each selected participant had their sleep monitored on two separate nights in their own homes. Two different signals were presented each night. Tests were normally one week apart to allow for recovery from any sleep deprivation, with the minimum being three nights. The participant was required to sleep on their own with the bedroom door closed.

All participants were told they needed to have an average or above average sleep the night before testing and that only a very moderate quantity of alcohol, if any, could be had earlier in the day and that it was important that both days of testing were as similar as possible. A questionnaire (see Appendix K) was completed each testing night to check these requirements. (In all cases these requirements were met.) The sleep technician (ST) arrived at the participant's home about 1.5 hours prior to their usual bedtime. After setting up the equipment the ST measured the level of background noise in decibels (using an average reading with the meter on a slow response). They then calibrated the sounds in the bedroom. The speakers were placed approximately one metre from the pillow. A file of the mixed T-3 sound was played which had previously been recorded to be received at 60 dBA when 1 metre from the speakers and the

delivery levels calibrated (see Appendix G). The sounds to be delivered that night were played to the participant without comment. The sleep technician applied the electrodes and then the practice trials and baseline measures of the performance tasks were completed (TMT, walk to phone from sitting in bed, phone call).

All participants completed the TMT both at baseline and as soon as they woke up from their second awakening for the night (called the sleep inertia condition). In both conditions the bedside light was turned on, they sat up and were given a "Stable Table" for their lap with the TMT form on it and a pen. They were timed from when they began the task. All participants had previously completed a shorter version (eight dots) of both TMT forms A and B, so they knew what was expected of them. Participants were instructed to join the numbers (part A) or the numbers and letters (part B) as fast and as accurately as possible and not to lift their pencil from the page. Time to completion was recorded, as were the number and type of errors. In analysing the data, these raw scores (time, number of errors) were used, as well as difference scores (part B minus part A). The difference scores, by using the participant as his / her own control, arrive at a more sensitive measure of individual cognitive (dys)function than raw scores (for example, Arbuthnott & Frank, 2000).

The performance task used to measure *physical* sleep inertia consisted of the time taken to get out of bed and complete a 15 metre walk from the bed to the phone, following a 15 m green rope. This task was timed from sitting in bed until the phone was reached. During the study it was decided that it may also be of interest to know how long it took each participant to do each of the two parts of the task, that is, (i) getting out of bed and (ii) walking the 15 m. Thus after the first 13 participants the two components were timed and recorded separately (using the lap function on the stopwatch to still obtain an accurate and comparable overall time).

The *simple cognitive* task that was required of the participant, after walking to the phone, was to dial a certain number and repeat a message to the answering machine. This consisted of their name and address and what night and condition of testing it was (i.e. night 1 baseline condition). The complete phone call was timed. It was initially

40

expected that the number of errors made in the message would also be analysed but it was found that very few errors were made, so this aspect did not proceed. The involvement of the ST in supervising the out of bed tasks was minimal (primarily preventing falls and ensuring compliance in following the 15 m trail).

After baseline measures were recorded the participant went to bed to sleep. The ST was in the hallway outside the bedroom monitoring their sleep patterns on a laptop computer. Signals were delivered in slow wave sleep, either stage 3 or 4.[9] When the participant entered stage 3 sleep the ST waited 90 seconds before delivering the sound. If the participant moved to a lighter sleep (e.g. stage 2 or 1) then the ST waited till they again reached stage 3 sleep and maintained it for 90 seconds (or went into stage 4 sleep). They then commenced the automatic sound delivery program, set to play the required auditory signal. Each signal was presented at each volume level for 30 seconds at a time, beginning at a low pillow volume level (35 dBA) and increasing by 5 dBA until awakening occurred. The loudest signal was 95 dBA and this continued for a total of three and a half minutes, or until awakening occurred, whichever occurred first.

The order of the presentation of the signals was counterbalanced across both subjects and nights. The participant was instructed that for the first signal each night all that was required was for them to press the button by their bedside three times when they first woke up and then try to go back to sleep. With the second (and final) signal presentation each night they were required to do a series of tasks to test their sleep inertia once they wake up. (If they did not awaken with the second signal the ST would gently shake them awake.) After the sleep inertia tasks were completed the electrodes were removed and the ST departed. All participants had the same ST on both nights.

This research was approved by the Victoria University Human Experimentation Ethics Committee.

---

[9] The option of either stage 3 or 4 sleep was chosen in contrast to only stage 4 sleep, which has been used with younger adults, because of concern that not all participants would enter stage 4 sleep sufficiently to consistently present all signals in that stage. See relevant section of Section 2.2.

4.4 Data analysis

All data was analysed using SPSS for Windows 11.0 and the alpha level required for significance was set at 0.05.

Responsiveness to Signals: For each of the four signals the following dependent variables were available for analysis across the whole group and as a function of age (above and below 75 years) and sex:

- Auditory arousal threshold (AAT), or the mean decibel level at which participants awoke - mean, standard deviation, range and median.
- Behavioural response time - mean, standard deviation, range and median. The behavioural response time is the accumulated time to press the bedside button from when the signals are presented at incremental volumes (i.e increasing from 35 dBA every 30 seconds to maximum of 95 dBA).

For the purposes of data analysis it was important to be able to incorporate the data of those who slept through the presented signals in such a way as to allow statistical comparisons. This was operationalised as in previous studies (e.g. Ball and Bruck 2004a). Specifically, if a participant slept through the full three and a half minutes of signal presentation at 95 dBA the volume at which they awoke was arbitrarily assigned as 105 dBA and their behavioural response time as 600 seconds. In considering mean values the effect of such an assignation must be kept in mind as it may *underestimate* the mean values of those signals where people are more likely to sleep through very loud volumes. This is because it assumes that everyone will wake up at 105 dBA, but this in fact may not be the case at all. It may also mean that statistical comparisons may fail to find a difference when there really is a difference. Median values are, of course, unaffected by this.

Comparisons were made across the four signals using repeated measures analysis of variance and across the categories of age and sex (using independent t-tests). Some descriptive frequency analyses were also conducted on the AAT in terms of how many people woke at different decibel levels.

The data from this study was directly compared to data collected by M. Ball and D. Bruck (partially published in Ball and Bruck 2004a) in sober 18-26 year olds (n=14 or 10 for different signals, as discussed below).

Sleep Inertia: Any decrements due to sleep inertia were objectively determined for each participant on each of the two nights of testing, by comparing the baseline and sleep inertia (alarm awakening) conditions. Data are normally presented for each night. Two way analysis of variance calculations were made with 'nights' as one factor and 'condition' (baseline and sleep inertia) as the other.   Variables include number correct and time taken for the Trail Making Task (A and B), time taken to get out of bed, time for the 15 m walk, time for the phone call. Because the order of the signals needed to be counterbalanced it was not possible to compare sleep inertia with awakening to different alarm signals.

5   Results

5.1   Responsiveness to Signals

5.1.1   Differences across the four signals

There was a highly significant difference between all four signals presented for both of the dependent variables measured (behavioural response time[10] and auditory arousal threshold, AAT). Table 5.1 presents the relevant data and analyses results. Participants awoke most readily to the mixed T-3 signal, while the highest AAT was to the high T-3 (the current US smoke alarm signal). Consideration of the median AATs shows a 20 dBA difference, from 45 to 65 dBA, between the mixed T-3 and the high T-3 respectively.

Table 5.1 also shows the percentage of participants who slept through the 75 dBA level (the minimum recommended level at the pillow in the US). Between 14 and 18% slept through the three signals that performed most poorly (high T-3, 500 Hz T-3 and the male voice), while 5% slept through the mixed T-3 at 75 dBA.

It can be seen in Table 5.1 that three of the older adult group did not awaken at all to the male voice (at 95 dBA). On closer inspection of the raw data it was determined that two of these people were from a non-English speaking background (NESB, Arabic) and had participated in the study with the help of a translator. They had not slept through any other signal presented. It was decided to re-run the key analyses across the four signals omitting the three NESB participants.[11] This re-analysis changed the mean AAT to the male voice from 55.9 (S.D.=19.2) to 53.6 (16.4) but did not change the level of significances of the overall analyses, including the pair-wise comparisons shown in Table 5.2. Thus the results for the male voice signal were not confounded in any important or significant way by the inclusion of the three NESB participants. (Although they certainly raise an issue to be researched further if a voice alarm is being considered.)

---

[10] to press the bedside button indicating awakening
[11] the one NESB participant who did awaken to the voice alarm had their highest AAT to this signal

Table 5.1: Summary of descriptive statistics and repeated measures ANOVA analyses for auditory arousal threshold (AAT) and behavioural response time for the four signals presented (n=42).[12]

		Mixed T-3	Male voice	High T-3	500 Hz T-3	ANOVA F (df=3,39)	p
AAT (dBA)	mean	48.0	55.9	63.7	52.6	9.9	.000
	S.D.	13.3	19.2	15.3	18.1		
	range	35-85	35-105	35-105	35-105		
	median	45	50	65	45		
N (%) slept thru 75 dBA		2 (4.6%)	6 (14.0%)	8 (18.3%)	7 (15.5%)		
N (%) slept thru 85 dBA		1 (2.3%)	4 (9.3%)	2 (4.6%)	3 (6.6%)		
N (%) slept thru 95 dBA		0 (0%)	3 (7.0%)	1 (2.3%)	1 (2.3%)		
Behavioural response time (seconds)	mean	93.3	153.9	192.1	124.5	9.9	.000
	S.D.	77.9	147.7	105.2	121.8		
	range	6-324	19-600	11-600	8-600		
	median	75	91	197.5	83		

Table 5.2 shows the pair-wise comparisons between all four signals (using the Least Significant Difference statistic). The same pattern of differences was found for most comparisons whether the dependent variable was mean AAT or mean behavioural response time. Importantly, the mixed T-3 fared better than ALL other signals presented, with either a significant difference being found between comparisons, or a trend.

---

[12] The partial eta squared statistic was 0.43, while the observed power was 0.99. The derived effect size (Cohen's d) for the mixed T-3 versus the high T-3 was 0.97.

Table 5.2: Matrix showing the level of significance for pair-wise comparisons across the four signals (using Least Significant Difference statistic) (n=42). (Unless otherwise specified results were the same for both AAT and behavioural response variable.)

	Mixed T-3	Male Voice	High T-3	500 Hz T-3
Mixed T-3	/////	/////	/////	/////
Male Voice	**	/////	/////	/////
High T-3	***	0.03/ns[13]	/////	/////
500 Hz T-3	#	ns	***	/////

ns indicates not significant with p>.10;   # p<.10;   *p<.05;   **p<.01;   ***p<.001.
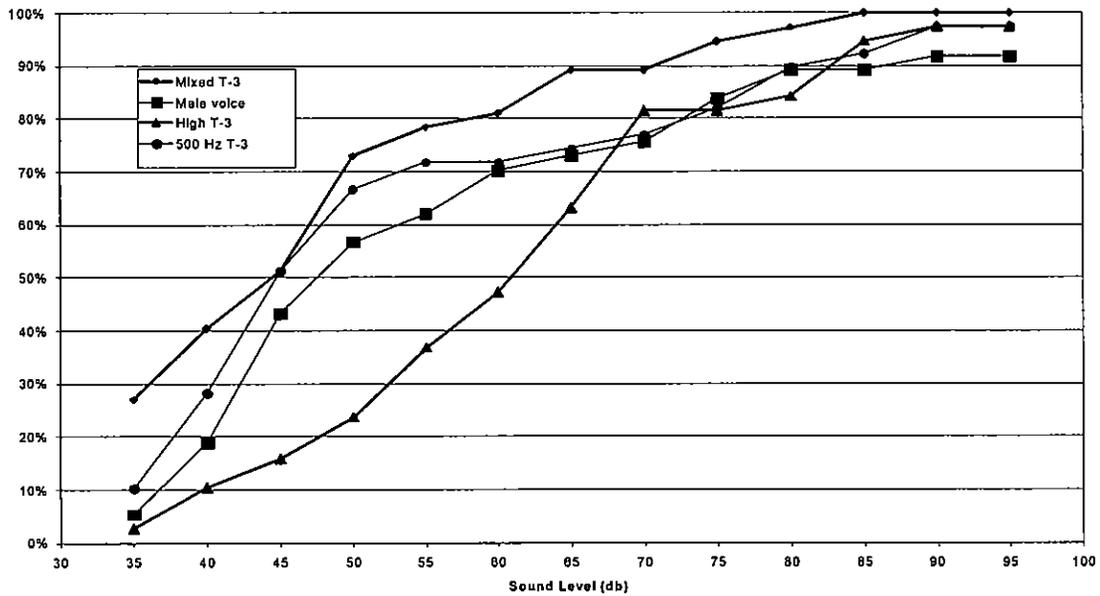
Figure 5.1 Cumulative frequency (percentage) for the four signals as a function of auditory arousal threshold. (Where the cumulative percentage did not attain 100%, not all participants awoke.)[14]

---

[13] The first item represents the level of significant difference for the AAT variable and the second for the behavioural response variable.

[14] This chart does not represent the overall population because only participants meeting the selection criteria were tested.

Figure 5.1 shows the cumulative frequencies for AATs for each signal. The main differences between the best performing signal (mixed T-3) and the poorest performing signal (high T-3) are at volumes below 70 dBA. However, cumulative graphs, by their very nature, tend to cluster at the upper levels. The important figures at any sound level are the proportion of people who do not wake, rather than the proportion who do wake. Thus at 75 dBA 5% of this sample had not woken for the mixed T-3, whereas for the high T-3 it is about 18% (over three times less effective), and for the 500 Hz T-3 and male voice it is 14-16%, about three times less effective. The objective is to awaken as many people as possible and at 75 dBA we may *theoretically* be comparing 5 deaths per hundred to 18 deaths. Similar interpretations can be made at each subsequent level (see Table 5.1 for the percentages). For reasons spelt out elsewhere (see Discussion Section 6.1.7), these AAT levels and percentages cannot be generalised to the general population.

5.1.2 Sex and Age differences

Further analyses were conducted to consider sex differences and differences between the 65-74 year olds and 75-85 year olds. Table 5.3 shows that no significant differences were found between males and females for AATs to any of the four signals. Table 5.4 shows that there was a significant difference between the 65-74 and 75-85 age group on AATs to the high T-3, with the 75-85 year old adults having higher AATs. For the older group the median was 70, compared to a median of 60 for the 65-74 year olds. It was found that 5/18 (28%) of the 75-85 year old participants slept through the high T-3 at 75 dBA, while 1/18 (6%) slept through the 95 dBA high T-3. The AATs of this 75-85 year old age group (including sex differences) are further explored in Section 5.1.3.

To further investigate the relationship between age and AAT a correlation was performed comparing age with AAT for each of the four signals presented. A moderate correlation was found between the high T-3 AAT and age ($r=.47$, $p=.001$, $n=44$). All other correlations were less than 0.4. This is consistent with the findings in Table 5.4.

Table 5.3: Summary of descriptive statistics and independent t-test analyses for AATs for the four signals presented for males versus females.

	Mixed T-3		Male Voice		High T-3		500 Hz T-3	
Sex	m	f	m	f	m	f	m	f
n	21	22	21	22	21	23	21	23
mean	51.2	45.0	54.5	57.3	65.7	61.9	54.7	50.4
S.D.	14.4	11.6	15.6	22.5	15.4	15.3	19.9	16.3
median	47.5	42.5	50	47.5	65	62.5	45	45
t (df) p level	1.55 (41) ns		.46 (41) ns		0.81 (42) ns		.81 (43) ns	

Note: n and df numbers vary due to some missing data. *ns* indicates not significant at p<.05

Table 5.4: Summary of descriptive statistics and independent t-test analyses for AATs for the four signals presented for 65-74 year olds versus 75-85 year olds.

	Mixed T-3		Male Voice		High T-3		500 Hz T-3	
Age(yrs)	65-74	75-85	65-74	75-85	65-74	75-85	65-74	75-85
n	27	16	26	17	26	18	27	18
mean	49.8	45.0	54.4	58.2	58.9	70.8	54.8	49.2
S.D.	13.9	11.9	17.9	21.5	14.2	14.5	17.9	18.3
median	45	40	47.5	57.5	60	70	45	42.5
t (df) p level	1.15 (41) ns		.63 (41) ns		2.7 (42) .009		1.1 (43) ns	

Note: n and df numbers vary due to some missing data. *ns* indicates not significant at p<.05

5.1.3 Hearing levels: awake versus asleep

The hearing threshold data when awake (decibel hearing level, dBHL) for all participants was tabulated across all five frequencies tested and categorized by age and sex. The data is shown in Appendix E (part 2). A comparison between the values

found for the study participants and the normative values shows that the study participants typically had lower thresholds, indicating better hearing. This is to be expected given that the study group did not include the lowest 16% of hearing thresholds.

The mean dBHL values for the participants were then correlated with the mean auditory arousal thresholds (AATs) for each of the four signals tested. Some moderate correlations (i.e. r > 0.4) were found between the higher frequencies (i.e 3000 and 4000 Hz) and the AATs for the high T-3 (see Table 5.5).[15] No other correlations reached a moderate level. A scattergram of the relationship for the best correlation is shown in Figure 5.2.

Table 5.5: Matrix of correlation values (Pearson's r) between dBHL (decibel hearing level) at 3000 and 4000 Hz (when awake) and AAT to the high T-3 (when asleep) (n=41).

	High T-3 AAT
Right ear dBHL 3000 Hz	0.57, p=.000
Left ear dBHL 3000 Hz	0.55, p=.000
Right ear dBHL 4000 Hz	0.47, p=.002
Left ear dBHL 4000 Hz	0.46, p=.002

---

[15] It should be remembered that the nature of the high frequency signals is different, although their frequencies are similar. When awake a single pulse was tested, while when asleep the signals was in the T-3 pattern.

auditory threshold for 3000 Hz (right ear) when awake
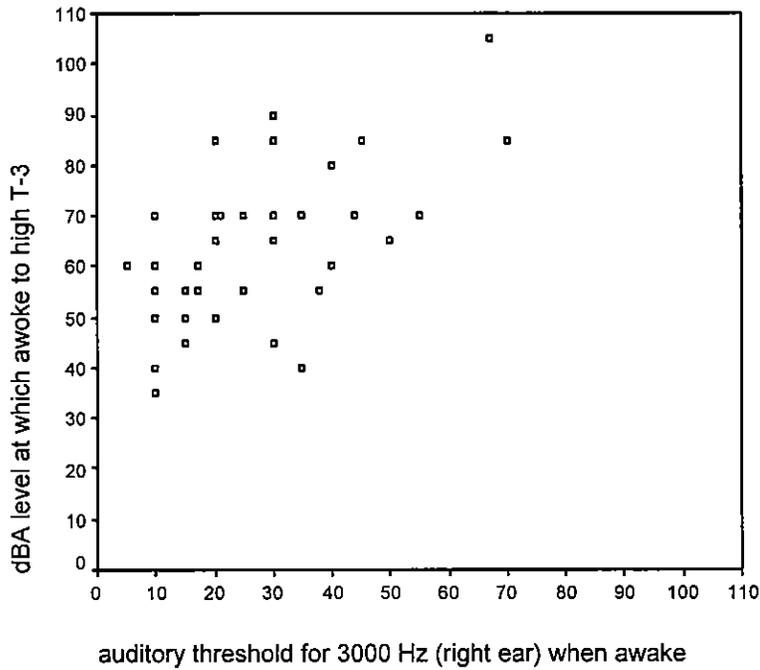
Figure 5.2: Scattergram comparing arousal to 3000 Hz high T-3 signal (from sleep) to auditory threshold (dBHL) to 3000 Hz when awake (n=41).

Given the above findings and the fact that hearing for higher frequencies declines with increasing age more for males than females, the thresholds for higher frequencies for the 75-85 year olds were further explored as a function of sex. Results are shown in Table 5.6.

Table 5.6: Comparison of hearing thresholds at 3000 Hz when asleep and awake for the 75-85 year old participants, by sex. dBHL= decibel hearing level as determined during a screening test when awake. Norms are taken from Cruickshank *et al.* (1998) and relevant details of these and dBHL for the study participants can be found in Appendix E (Part 2).

		Males	Females
Asleep	n	9	9
High T-3 AATs (dBA)	mean	72.2	69.4
for 75-85 yr olds	S.D.	14.6	15.1
	median	70	70
	range	55-105	40-85
	Slept thru 75 dBA	2	3
	Slept thru 95 dBA	1	0
Awake	75-85 yrs: study:		
dBHL for 3000 Hz (left ear)	mean	44.3	35.0
	S.D.	20.1	14.4
	median	44	30
	range	15-70	22-55
	70-79 yrs: norms mean	56.1	34.9
	80-89 yrs norms mean	63.4	48.3

It can seen from Table 5.6 that the 75-85 year old males and females performed at similar levels when asleep. That is, females are sleeping through similar volumes to males although their hearing at the upper frequencies is better. Caution in interpretation is necessary due to the small numbers of participants in these subgroups.

The ability to hear high frequency signals when asleep versus awake was further explored. A variable was calculated that was the difference between the high T-3 AAT and dbHL for the 3000 Hz signal (left ear-worst).   The results are shown in Table 5.7 and it can be seen that for some participants the difference was very small (5 dBA), while for others it was large (65 dBA). No significant differences were found between the 65-74 year and 75-85 year age groups (t=.56, df=39, p>,05).

Table 5.7: Descriptive statistics for the difference between auditory thresholds when awake (dBHL for 3000 Hz, left ear) and asleep (AAT for high T-3) for 3000 Hz.

n	41
Mean dBA difference	33.2
S.D.	15.6
Median	30
Minimum dBA difference	5
Maximum dBA difference	65

5.1.4 Comparisons between older adults and young adults

In this study several of the signals presented were the same as presented to a group of 18 to 26 year olds and a similar methodology was used.[16] Thus comparisons could be made across age groups, and they are shown in Figure 5.3. The study is the same as reported in Ball and Bruck (2004a) except that more participant data had become available. The young adult data is based on n=14 for all signals except the male voice, where n=10.
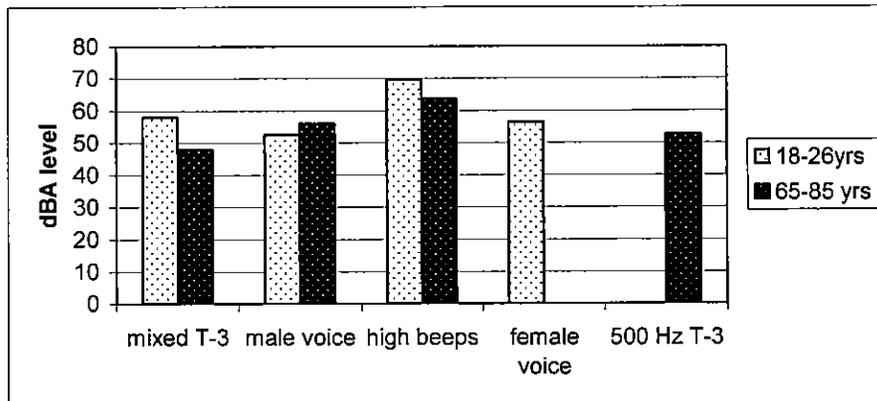
Figure 5.3: Comparison of AAT (dBA level at which awoke) for the older adult sample with a sample of young adults (see text).

Independent T-tests were conducted comparing the young adult sample with the older adult sample for the three signals where data was available for both age groups. For the mixed T-3 signal a significant difference was found (t=2.31, df=55, p=.03), indicating that the young adults had a significantly higher mean AAT (57.9 dBA, S.D.=13.9) than the older adult sample (48.0 dBA; as in Table 5.1). Comparisons between the two age groups for the male voice and the high pitched alarms[17], showed no significant differences. The expectation that the older adults would have lower AATs to all signals compared to the young adults was not fulfilled. If the NESB participants are excluded from the male voice data for the older adults, the means are very similar to the male voice data for the young adult group.

5.2 Sleep Inertia

5.2.1   Trail Making Task (TMT)

The time taken to complete the *TMT A task* was analysed using a two way repeated measures ANOVA where test condition (baseline versus sleep inertia) was one factor and test night (N1, N2) was the other factor. The main effect for condition was significant (F=7.65, p=0.009), such that performance in TMT A was 17.4% slower in the sleep inertia condition compared with the baseline (46 vs 54 sec respectively), averaged across test nights (see Table 5.8 for descriptive statistics). There was no significant difference between the two nights (F=0.055, p<0.05). The interaction between test night and condition was not significant. Strong correlations were found for the time taken to complete TMT A across baseline and sleep inertia conditions (Pearson's r >.85, p=.000, for both N1 and N2).

---

[16] However, for the older adults signals were presented in either stage 3 or stage 4, while for the young adults all were presented in stage 4. Nevertheless, for both groups this sleep represents their dominant deepest sleep stage (as stage 4 declines considerably in older adults, see section 2.2)

[17] Caution must be used in comparing the data for "high beeps" across the two age groups as, although they are both in the range 3000-4000 Hz, with the young adult group the sound was a continuous fast beeping (as found in the older US alarms) and in the older adult group the beeps were in a T-3 sequence.

Table 5.8: Descriptive statistics for Trail Making Task A and B.

	Mean	S.D.
baseline time on trail making A N1	60.3	39.9
sleep inertia for trail making A N1	69.4	49.8
baseline time on trail making A N2	58.6	41.1
sleep inertia for trail making A N2	66.3	55.0
baseline time on trail making B N1	128.6	81.7
sleep inertia for trail making B N1	148.4	97.2
baseline time on trail making B N2	128.6	81.7
sleep inertia for trail making B N2	133.4	83.3

Similarly, the time taken to complete the more complex *TMT B task* was analysed using a two way repeated measures ANOVA where condition (baseline versus sleep inertia) was one factor and test night (N1, N2) was the other factor (see Table 5.8). The main effect for test night was not significant, and neither was the main effect of condition, or the test night by condition interaction (F=1.19, p>.05; F=0.001, p>0.05 and F=0.69, p>0.05 respectively).

The difference between performance for *TMT B minus TMT A* was calculated as a measure of cognitive switching or processing 'efficiency'. Scores were submitted to a two way ANOVA, as above. The main effects for test night and test condition were both non-significant (F=1.08, p>0.05 and F=0.07, p>0.05 respectively) indicating that there were no alterations (either improvement or deterioration) in processing as assessed by the TMT across test nights or test occasions.

Further exploration of the data was conducted by considering *TMT errors*, specifically the frequency of participants who coped well or poorly with the TMT tasks. A cut off of

three errors was used as the threshold for the two categories.[18] The data for N2 was used because by the second night the participants had become more familiar with the task. Table 5.9 suggests stability in the number of errors across the N2 baseline and sleep inertia conditions for both TMT A and TMT B. It also indicates that TMT B was a very difficult task, even under optimal conditions (baseline on N2), with 44% making more than 3 errors at baseline.

Table 5.9: Frequency of participants having 0-3 or >3 errors on TMT A and TMT B on night 2, across baseline and sleep inertia conditions. (Total n= 45)

		Baseline	Sleep inertia
TMT A	0-3 errors	41	38
	>3 errors	2 missing: 2	4 missing: 3
TMT B	0-3 errors	21	24
	>3 errors	20 missing: 4	16 missing: 5

Thus, speed of processing was affected adversely by the sleep inertia condition, as assessed by performance on TMT A. Cognitive switching, however, did not differ between baseline and sleep inertia conditions. Participants performed comparably across the two nights and across the two sleep conditions on TMT B.

5.2.2 Simple physical task

The overall time taken for participants to get out bed and walk 15 metres to the phone was assessed under baseline and sleep inertia conditions. The descriptive and statistical data for this variable was available for 41 participants and are shown in Table 5.10. A two way ANOVA was performed with condition (sleep inertia) as one factor and nights (N1 versus N2) as the other factor. A significant difference was found for the condition effect, with an increased time taken under the sleep inertia condition

[18] This criteria is applied clinically where the clinician is advised to terminate the test as being too difficult if the patient makes more than three errors (Groth-Marnat, 2000).

(F(1,40)=13.2, p=.001). No significant difference was found across test nights or asan interaction between the two factors.

This above variable of overall time was broken down into two variables for 28 of the participants (see Table 5.10). The first variable was the time taken to get out of bed and stand, for which a significant difference was found for condition (sleep inertia versus baseline, F(1,27)= 5.2, p=.03). The second variable was the time to walk 15 metres from the bed to the phone and a significant difference was found between the sleep inertia and baseline condition (F(1,27)=8.12, p=.008). There was no significant difference between the two nights or the interaction between nights and condition for either of these variables.

The mean values for the total time to get out of bed and walk 15 metres (averaged across N1 and N2, Table 5.10) showed a 9.6% decline from the baseline to sleep inertia condition. For all variables large standard deviations and range values are evident. In all the comparisons where significant differences were found across baseline and sleep inertia conditions there were significant correlations across the two conditions, indicating that the time an individual took completing the physical task under baseline conditions strongly predicted the time taken to complete the task under sleep inertia conditions. A scattergram for the overall time variable (see Figure 5.4) shows a strong linear relationship. The correlation was high and significant (Pearson's r=.98, p=.000, N=44).

Table 5.10: Descriptive statistics (in seconds) and ANOVA p levels for sleep inertia vs baseline condition for simple physical tasks across night 1 (N1) and night 2 (N2).

Performance task	Condition and night	Mean (S. D.)	Range	ANOVA p level
Getting out of bed and walking 15m to phone (n=41)	baseline N1	30.7 (21.6)	11-144	
	sleep inertia N1	34.4 (28.0)	12-183	.001
	baseline N2	29.5 (21.3)	14-148	
	sleep inertia N2	34.4 (26.9)	13-184	
Getting out of bed only (n=28)	baseline N1	4.5 (3.3)	1-16	
	sleep inertia N1	7.3 (7.6)	1-41	(.08)
	baseline N2	5.0 (3.1)	1-18	
	sleep inertia N2	7.5 (7.3)	1-32	
Walking 15 m only (n=28)	baseline N1	29.1 (24.6)	8-139	
	sleep inertia N1	30.4 (26.6)	11-143	.008
	baseline N2	26.2 (23.4)	12-138	
	sleep inertia N2	29.6 (25.5)	14-152	

average of N1 & N2 for baseline total walk
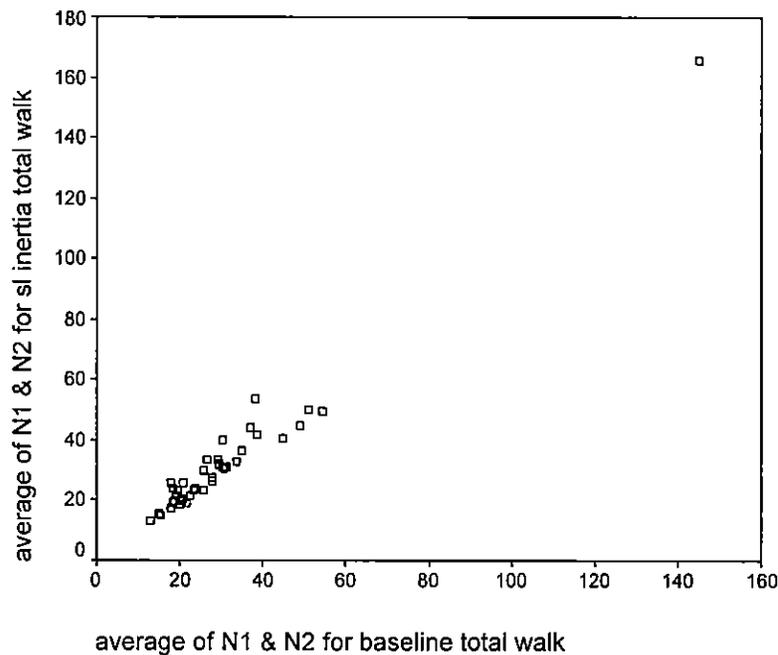
Figure 5.4: Scattergram of time to get out of bed and walk 15 metres, comparing baseline and sleep inertia conditions (average over both nights) (n=44).

No significant differences as a function of sex or age group were evident. To determine this the physical task data (for variables as shown in Table 5.10) were averaged across both nights, and difference scores calculated (i.e. difference between sleep inertia times and baseline scores). Each difference score was analysed for possible sex differences and differences across the 65-74 year versus 75-85 year old groups using independent t-tests.

5.2.3 Simple cognitive task:

The simple cognitive task was the time taken for the participant to make a phone call and record a simple message stating their name and address to an answering machine. This was completed after the physical tasks discussed above. Thus it was undertaken, on average, about four minutes after starting the Trail Making Task, perhaps about an average of five minutes after awakening. A two way ANOVA found no significant differences for condition (baseline versus sleep inertia), night (N1 vs N2) or an interaction. Table 5.11 shows the descriptive data.[19]

Table 5.11: Descriptive data for "time to make phone call" (in seconds) under baseline and sleep inertia conditions across N1 and N2.

		Baseline	**Sleep inertia**
N1	Mean	55.7	59.9
	S.D.	11.9	18.2
	range	22-87	23-138
	n	44	44
N2	Mean	56.2	56.6
	S.D.	9.5	10.9
	Range	29-77	28-90
	n	43	41

5.3 Other data

5.3.1 Presentation order

It is possible that the AATs may have been influenced by the order of presentation of the signals. In other words participants may have found it easier, or harder, to wake up quickly with the first signal presented compared to the fourth (and final) signal. To check for the possible existence of such a confound a repeated measures ANOVA was conducted where the data was organised by order of presentation. No significant order effect was found (F(3,39) = 1.16, p=.34)

5.3.2 Optimal number of participants

An investigation was conducted of the number of participants that were required to obtain stability in the data. Too few subjects and the results are subject to a great deal of uncertainty, too many and there is little or no value added by additional subjects. Appendix M contains graphs that enable an evaluation of the appropriate number of subjects for this research. The figures in this appendix show the variation in the mean and standard deviation of the time and sound level at awakening as the number of subjects increased through this research. It can be seen that there was a great deal of variation in both the mean and standard deviation while the number of subjects was low (say below ten), that the variation of both decreased as the number of subjects increased, and that once the number of subjects was above about twenty both the mean and standard deviation varied little. These observations hold for all of the sounds, though there is some variation between the sounds. On the basis of these observations it is proposed that similar experimental programs in the future be limited to 25 subjects. This number seems appropriate in that it appears that the results may be expected to be quite representative without testing of an excessive number of subjects for little increase in accuracy resulting from additional subjects.

---

[19] In listening to the answering machine messages, the occurrence of errors was extremely rare, so analyses of this variable was not undertaken.

5.3.3 Time to fall asleep

Participants fell asleep at night without too much difficulty. The mean time taken to fall asleep for the first opportunity to sleep each night (initial sleep onset latency) was 18.4 minutes (S.D.= 8.3) with a minimum of 5.5 minutes and a maximum of 39.5 minutes. Perusal of the intervals between the first and second awakening each night showed that in almost all cases the time needed to fall asleep after the first awakening was within the range of the initial sleep onset times.

5.3.4 Background sound levels

The average decibel levels measured in the participant's bedroom (with the door closed) was 39.4 dBA (S.D. = 4.6) with a range from 33.5 to 50.5 dBA.

5.3.5 Stage of sleep / response time

The aim was to present all the signals in slow wave sleep. The signals were to be presented 90 seconds after the first appearance of stage 3 sleep, provided the person did not move into a lighter stage of sleep (e.g. stage 2). During this 90 second interval some participants would have moved into stage 4. Independent scoring of a representative sample of the sleep stage data was subsequently conducted and it was determined that approximately 92% of all signals were initially presented during stage 3 sleep and about 8% were presented during stage 4 sleep. No systematic differences in response time between these two groups were evident.

From this sample it was further determined that in approximately two thirds of all participants less than 15 seconds elapsed between the first moment of electroencephlagraphic (EEG) defined wakefulness to the time of first pressing the bedside button. About one third took between 16 and 60 seconds.

6 Discussion

6.1 Responsiveness to signals

6.1.1 Difference between signals

The first hypothesis stated that the older adult sample would have higher arousal thresholds (AATs) to the high pitched T-3 than to the two signals of mixed frequency (the mixed T-3 and the male voice). This hypothesis was only partially supported, with the results showing that mean AATs to the high T-3 were significantly higher than to the mixed T-3. In other words responsiveness during sleep to the high T-3 was significantly worse than to the mixed T-3. The three poorer performing signals were somewhat clustered, with no significant difference between AATs to the high T-3 and 500 Hz signal or between the male voice and the 500 Hz signal. The most important finding was that the mean AAT for the mixed T-3 was lower than for the other three signals tested. The difference between the median levels was 20 dBA, with the mixed T-3 having a median AAT of 45 dBA and the high T-3 median AAT being 65 dBA. This is a substantial difference.

6.1.2 Sex, age and hearing analyses

No sex differences were evident in the AAT data. A significant difference was found between the 65-74 year olds and 75-85 year olds on AATs to the high T-3, with the older group needing increased volumes to wake up. This older group had a median AAT of 65 dBA for the high T-3, compared to a median AAT of 40 dBA for the mixed T-3. This substantial difference is most likely related to the normal age related decline in ability to hear high frequencies. In this context it should be noted that the hearing screen criteria (as set out in Appendix E) had threshold criteria at 3000 Hz of 70 dBA for those males aged 70-79 and a threshold of 50 dBA for females in the same age group. This sex difference in auditory thresholds when awake did not translate to a sex difference in AATs to the high T-3 when asleep, even for the 75-85 year old group. There was not even a trend for a difference between the sexes when they were asleep.

Consistent with this, further explorations of the relationship between hearing thresholds when awake and AATs when asleep revealed no simple relationship. While moderate

correlations between high T-3 AATs and awake hearing thresholds for 3000 Hz and 4000 Hz signals were found, the data showed that for some people the difference between the high frequencies they could hear when awake and the high T-3 volume needed to wake them up was 5 decibels, while for others it was up to 65 decibels, with the average being a 33 dBA difference for a 3000 Hz signal. The norms (see Appendix E) show that a volume of 70 dBA is required for a 3000 Hz signal to be heard by 84% of 70 year old males when awake. This increases to a threshold of 80 dBA for males over 80 years. Adding the 33 dBA mean value to these awake volumes means that, on average, very high volumes of the high T-3 would be required for awakening for older males.

6.1.3 Older adults versus young adults

The second hypothesis predicted that the older adults would have lower AATs to all signals than a young adult sample tested under comparable conditions. Identical signals (under similar presentation methods) were available for the mixed T-3 and the male voice. The hypothesis was supported for the mixed T-3 signal, where the young adults were found to need significantly higher volumes to wake up compared to the older adults. It was not supported for the male voice signal, where the young adults had similar AATs to the older adults. Mean values showed a minor non-significant difference in the predicted direction for the high frequency signals.

6.1.4 The mixed T-3

In this sample of older adults the mixed T-3 signal clearly woke individuals more readily than the alternatives. The mixed T-3 has now been found to perform significantly better in awakening sleepers at lower volumes than the current high pitched T-3 for various population groups. These groups are children (aged 6-10 years, Bruck and Ball, 2004), sober young adults and alcohol intoxicated young adults (Ball and Bruck, 2004a) and, from the current study, older adults.

The fact that the participants in the current study did not wake as readily to the 500 Hz T-3 signal shows that this advantage does not simply arise from the inclusion of low pitched frequencies in the mixed T-3. Comparison of the spectral analyses of the mixed

T-3 and the 500 Hz signal (see Appendix A) show that both have the same dominant frequency at around 500 Hz, but the mixed T-3 also includes the 3rd, 5th etc. harmonics, with the most dominant harmonic being at just over 1500 Hz. It would seem that the inclusion of these harmonics is important in helping people wake up at lower volumes. It needs to be determined if other mixed frequency signals, with perhaps higher dominant frequencies and higher harmonics are even more effective at waking people up.

It is not immediately obvious why the mixed T-3 signal should be the most effective signal tested so far for waking individuals up. Perhaps it is because human responsiveness to sounds when asleep is best when the signal includes a range of frequencies. However, because the male voice did not wake people as easily as the mixed T-3 it is not the case that *any* pattern with a mixed frequency profile between 500 Hz and 2500 Hz will give similar AATs. The fact that the best signal for waking up tested so far includes a range of spectral components in the low to mid range is consistent with some data testing responsiveness to different ringers when awake (Section 2.3.2).

6.1.5 Voice signal

The research literature, using comparatively small sample sizes, suggests that voice alarms performed as well as the mixed T-3 with the children and young adult samples, while the current study suggests voice alarms perform more poorly than the mixed T-3 with older adults. The reduced responsiveness to the male voice in these older adults was a somewhat unexpected finding which did not arise from the inclusion of three non-English speaking (NESB) participants. However, the fact that two of the NESB participants did not wake *at all* to the male voice suggests that voice alarms with an English text may be unsuitable for use in a population that includes NESB people.

While it is possible that the increased AATs to the male voice are related to the reduced ability to discriminate speech with increasing age (Kim, Frisina & Frisina, 2002; Frisina & Frisina, 1997), this seems unlikely to be the full explanation, as the task did not require *understanding* the words of the signal (which they had heard just prior to sleep),

only waking up when it occurred. One possibility is that older people may have developed more of a capacity to screen out irrelevant (and/or familiar?) noises occurring during sleep, and many irrelevant noises relate to people talking (e.g. background TV or radio sounds, conversations). Alternatively, perhaps the higher arousal thresholds to the voice signal was a function of increased dream incorporation of the voice compared to beeps, and older people may be more likely to do this than younger adults. In the absence of any evidence this remains speculative.

6.1.6 Signal continuity
Comparisons of this data with the AATs found by Zepelin et al. 1984 (where a 5 sec 800 Hz signal was presented at 2 minute intervals), show much lower AAT levels in the current data compared to Zepelin et al. and a much smaller gap in AATs between the different age groups (44 dBA gap between AATs across the two studies for the young adults versus 22 dBA gap across the two studies for the older group). These differences suggest there is an arousal advantage in a signal presenting continuous sounds (compared to lone signals) and this advantage may be especially important for the younger adults.

6.1.7 Experimental versus field settings
The sleep study was designed with one main intention. This was to compare arousal thresholds to different signals in older adults and to do so in a well controlled manner to reduce heterogeneity and therefore obtain the most valid result with a manageable sample size.

Extrapolating this data to actual field situations is highly problematic. This is the case for both arousal thresholds and estimates of what percentage of the population might wake up at what volume. There are multiple reasons for this. One set of limitations relates to the population used, while another relate to the methodology.

The selection criteria make it clear that this sleep study was conducted in a highly selected population of older adults. In many ways they represent the most highly functioning adults in the over 65 year age range. All older adults were omitted from the

study who were taking medication affecting sleep, cognitively not fully capable, not independently mobile, reported sometimes having difficulties falling asleep, and considered their hearing to be below average for someone their age. It was also required that alcohol intake was minimal or none during the study. In addition all those whose hearing at any of the five selected frequencies in either ear was in the lowest 16% of the population were excluded. Furthermore, the fact that recruitment was conducted through social groups meant that those older adults who were more isolated were not given the opportunity to be involved.

The methodology of the signal presentation in the study is not the same as may be expected in an emergency awakening situation in the home and three points are especially relevant. Firstly, the modified methods of limits methodology meant that each volume level of a signal was *not presented from silence*, as would be the case when a smoke alarm signal sounds. Examination of the literature on awakenings to alarms (see Sections 2.1 and 2.3) suggests that people may have lower AATs (i.e. wake up more readily at a certain decibel level) when that sound cuts in from silence. Thus the present AAT levels may be slightly inflated, with the sleepers becoming habituated to the gradually increasing volume of a signal up until a certain level, which then arouses them. The advantage of not including silences between the signals at different decibel levels is that it more readily allows statistical comparisons across signals, without the confounding effects of dealing with arousals that may happen in the intervening silence. This confound was discussed as a problem in an earlier paper (Ball and Bruck 2004b). An additional reason why generalisations are problematic is that all these participants were awoken from their *deepest sleep*, slow wave sleep (SWS). Previous literature suggests that AATs from other stages of sleep would be reduced (i.e. woken more easily) from other parts of sleep, with one study suggesting this would be by an average of 10 dBA in older adults (Zepelin *et al.* 1984, Section 2.2), In older adults most of their sleep consists of lighter sleep, with the minutes of SWS declining from an average of 67 minutes per night for 65 year olds to 25 minutes in 86 year olds (Ohayon *et al.* 2004, Section 2.2). Another consideration, however, is that most fatal fires occur earlier in the sleeping period, which is when SWS is most likely to occur (Thomas & Brennan, 2002). The third, and most important, methodological difficulty in generalising AATs from this

study to real field situations is that in this study all the participants were *primed*, that is, they knew they would be awoken by a signal, and knew which signals would be played on a testing night. Priming will significantly increase the likelihood of awakening, for example from 25% to 90%, (Wilson & Zung 1966, see Section 2.1), and this would decrease AATs.

Overall, the *population* factor acts to decrease the AATs (in the study, relative to the field), while within the *methodological* factor, the lack of silence and the use of the deepest sleep stage probably both act to increase the AATs. However, priming within the sleep study would decrease AATs. The differential effect of all these factors is not known for certain but the literature would suggest that the issues exerting the greatest effect are the population and priming. Thus in an unprimed, field population that includes people with a range of risk factors (e.g. intoxicated, sleep deprived, poorer hearing, being a child, etc., see Section 2.3.4) the responsiveness to signals would be reduced, compared to what has been found in this study. In other words, the AATs and percentage of those who slept through sounds in the study would be *underestimates* compared to a field setting.

Thus it is not valid to form any conclusions or recommendations on the notion that this sleep study suggests a certain percentage of older adults would wake to any particular alarm in the bedroom. In reality it is not known what percentage of adults over 65 would wake up in an emergency situation in their home. We do know, however, that of the alternatives tested the current high T-3 signal performs the worst, and under the conditions of the sleep study a mixed T-3 signal had a median AAT that was substantially lower (20 dBA). What is really needed are studies in a large sample of the general, unscreened population where the alarm is in their homes for a sufficient period such that they are no longer primed to wake up easily. This will allow meaningful predictions to be made about what percentage of the population may awaken to a certain signal at a certain volume in the home in an unprimed situation.

6.2 Sleep inertia

The Trail Making Task (TMT) provided good discrimination across the sample, with a

wide range of performance levels evident. The TMT data suggested that sleep inertia is associated with some quantitative changes in speed of processing, with a 17% increase in time (over baseline) to complete TMT A in the sleep inertia condition (with TMT A being essentially a psychomotor task of connecting numbers).

However, sleep inertia was not associated with any qualitative changes in the complex cognitive processing (cognitive switching, sustained attention and sequencing skills) as assessed by the TMT B. There are several possible explanations here. Firstly, that complex cognitive functioning is not reduced by sleep inertia in older adults. This would be in contrast to the effects of sleep inertia in young adults, where a 51% decrement in decision making performance has been documented within the first three minutes after awakening (Bruck & Pisani, 1999). Perhaps older adults were able to rally themselves to perform equally well in terms of complex cognitive performance (shown by TMT B accuracy) whether they have just woken up or not. This explanantion would be consistent with the reduction in EEG power with age (Astrom & Owren, 1995), indicative of overall lighter sleep in older adults. The results suggest it may be possible to predict which individuals would have trouble with the type of complex cognitive processing (cognitive switching, executive functioning) needed in some emergency situations by testing them under normal conditions while awake.

An alternative, but probably less likely, explanation for the current TMT B findings is that sleep inertia does not affect all aspects of cognitive functioning, and while it may affect decision making, it does not affect the type of cognitive functioning as assessed by TMT B. One piece of evidence arguing against this explanation is the study by Bliwise & Swan (2005) which suggested that the TMT was sensitive to sleepiness, which is held to be qualitatively the same as sleep inertia (Balkin & Badia, 1988). Further research is needed to explore both possible explanations.

Considering all the performance assessments made in this study, the main effect of sleep inertia was a slowing in speed of psychomotor task completion (shown by a 17% TMT A increase in completion time after awakening), and time decrement in physical performance tasks, where a 10 % slowing in the time taken to get out of bed and walk

67

15 metres (undertaken directly after the TMT) was found. For both these variables the percentage differences were statistically significant and the correlations showed that performance during baseline strongly predicted performance during the sleep inertia condition. In contrast, functioning on complex cognitive tasks (as assessed by the TMT B) was unaffected by sleep inertia. Simple cognitive tasks (as assessed by the time to complete the phone call) were also not affected by sleep inertia. This data has implications for fire protection engineers in estimating performance decrements arising from waking to a fire and, at an individual level, allow extrapolations from waking performance to likely performance when suddenly awoken.

6.3 Conclusions and Recommendations

The present study, using a rigorous design and sufficient sample size of sleeping adults aged over 65 years, has found a substantial difference in the median auditory arousal threshold of 20 dBA between the current high frequency T-3 and the best performing alternative signal tested. Thus all the available data testing the waking performance of smoke alarm signals shows that a high frequency alarm signal,[20] performs the *most poorly* of the alternatives tested for waking all the different population groups tested so far (i.e. children, sober and alcohol intoxicated young adults, older adults aged over 65 years). The evidence is sufficient to lead to the following recommendation:

Key Recommendation: The high frequency alarm signal currently found in smoke alarms should be replaced by an alternative signal that performs significantly better in awakening most of the adult population, once the nature of the best signal has been determined.

The findings of the current study, together with previous literature, indicate that a mixed frequency T-3 signal has *performed significantly better* than a high frequency signal in its ability to awaken sleepers in every sample group tested so far. This includes children, young adults (sober and alcohol intoxicated) and older adults. Voice signals

---

[20] A high frequency signal is typically used in all smoke alarms, the literature reported here has variously tested both a high frequency T-3 signal or a continuous pulsing high pitched beeps.

appear to be as effective as the mixed T-3 in the children and young adult groups, but are less effective than the mixed T-3 in the older adults.

Unfortunately, direct extrapolations from the present data to the field in terms of expected arousal thresholds in a real emergency or percentages of the population that may awaken to certain signals at certain volumes are not possible from this study due to the limitations imposed by the highly selected sample and methodology used. These population and methodological factors probably result in the research to date *underestimating* the proportion of people who will not wake up to an alarm.

The implications of introducing a signal frequency recommendation into the standards for smoke alarm notifications are considerable, involving a retooling of the entire industry. In view of this, any signal change that is mandated must be done on the basis of rigorous evidence that the best signal has in fact been found. The research is not yet at this point. Suggestions of future research issues are set out below. These may take two to three years to complete.

In the meantime there are some recommendations that can increase the chance of sleeping individuals waking to a fire.

(a) Encourage interconnected alarms. Interconnected alarms that include an alarm in the bedroom will mean that the volume at the pillow is likely to be above 85 dBA. Increased volume of any signal that can be perceived will increase the chance of waking up. In addition, the more people in a household that are exposed to alarm signals, the more chance that one of them will wake up.

(b) Consider the special hearing needs of "normal hearing" older adults. It is inadequate to require the smoke alarm (of 3000 Hz) for older adults to be a minimum level of 75 dBA at the pillow. The current study shows that both males and females aged over 75 were particularly poor at waking to the current high T-3 (median AAT of 70 dBA). (If their smoke alarm was a mixed T-3 signal the median AAT value would be 40 dBA.) One possibility would be to recommend

that older adults should have interconnected alarms, or at the very least stand alone alarms (with the current signal) in their bedroom. An additional, more satisfactory, possibility is for smoke alarm manufacturers to market special alarms for this age group that emit a mixed T-3 signal and suggest placement, at a minimum, in the bedroom.[21]

The future research that should be completed prior to the mandating of a specific signal encompasses a variety of issues.

(a) Determine the optimal pitch and pattern of an alternative signal to wake people up. Given the large individual differences in waking thresholds this should be done as a repeated measures design in an accessible population that sleeps well, such as young adults. From the results of the current study it is recommended that the option of a voice alarm be no longer considered for adult populations, both because it performed poorly with the older adults and because of the indicated problems with individuals who do not speak English. In view of the infinite number of different sound patterns that an alarm could take and the current dominance of the T-3 as the mandated temporal pattern it is suggested that alternative pitches (e.g. pure and mixed) and pitch patterns (e.g. changing frequencies across the 0.5 second tone) be investigated within the T-3 temporal pattern, at least in the first instance. On the basis of analyses undertaken with the current data it is recommended that sleep studies documenting comparative arousal to different signals using a similar methodology to this study, need only a maximum of 25 participants.[22]

---

[21] Such a mixed frequency alarm would also be beneficial for those who know they have high frequency hearing loss of any age. While bed vibrators are clearly an effective alternative for the hard of hearing (Ashley et al. 2005) their expense makes them unlikely to be widely adopted by older adults who consider their hearing to be average for their age.

[22] If analyses are to be performed across subgroups, or if major interpreations are made from cumulative percentages, more participants would be needed.

(b) Once several signals have been shown to have the lowest auditory arousal thresholds (AATs) in the one population tested, they need to be tested in other sleeping populations, especially those most at risk of dying in a fire or of sleeping through an alarm signal. These groups include prepubescent children, older adults (65+years), adults impaired with alcohol or hypnotic medication and people who are sleep deprived (such as is frequently the case with shift workers and adolescents). The signals should also be tested for salience and/or urgency as an emergency notification signal requiring action in awake individuals.

(c) Because of the inability to generalise data from the current study to field estimates, further research is needed using large numbers of non-primed, unselected groups to yield population based estimates of waking effectiveness. It seems most likely that the research to date may be *underestimating* the proportion of people who will not wake up to an alarm. This arises from a range of factors but especially because almost all of the participants in the relevant empirical studies on alarms and sleep have been primed to expect that a signal will go off on one of several nights. Studies are needed where there is a long time frame (e.g., one or two months) within which a test alarm may be activated and an unselected large sample is tested. This will yield valuable field estimates of alarm effectiveness where conditions influencing responsiveness are uncontrolled (e.g., alcohol intake, prior sleep deprivation, prior time in bed, sleep stage). A challenge would be the accurate monitoring of the latency to wake up.

(d) A study characterising the spectral characteristics of the background noises in a range of "typical" bedrooms would be informative and relevant. When this information is put together with the acoustical information about which signal is most likely to awaken sleepers, the extent of possible masking can be determined.

The current research is the first to consider the effects of sleep inertia in older adults. The results suggest that a decrement in physical functioning of around 10-17% may be expected across the first five minutes after awakening. No important effects on simple

71

or complex cognitive functioning were evident, although further research, using a variety of cognitive tasks, should be undertaken to verify this conclusion.

References

Akerstedt, T. Torsvall, L & Gillberg, M. (1989) Shift work and napping, In D.F. Dinges & R.J. Broughton (eds) Sleep and Alertness: Chronobiological, Behavioural and Medical Aspects of Napping. Raven Press, New York, 205-220.

Arbuthnott, K. and Frank, J. (2000) Trail Making Test part B as a measure of executive control: validation using a set-switching paradigm. Journal of Clinical and Experimental Neuropsychology, 22, 518-522.

Ashley, E., Du Bois, J., Klassen M & Roby, R. (2005). Waking effectiveness of audible, visual and vibratory emergency alarms across all hearing levels. Proceedings of the 8th International Symposium of the International Association for Fire Safety Science, Beijing, Sept (poster presentation).

Astrom, C., & Trjaborg, W. (1992). Relationship of age to power spectrum analysis of EEG during Sleep. Journal of Clinical Neurophysiology, 9(3), 424-430.

Bachorowski, J., & Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. Psychological Science, 6, 219 – 224.

Balkin, T. & Badia, P. (1988) Relationship between sleep inertia and sleepiness: cumulative effects of four nights of sleep distribution/restriction on performance upon abrupt nocturnal awakening. Biological Psychology, 27, 245-258.

Ball, M., & Bruck, D. (2004a). The effect of alcohol upon response to fire alarm signals in sleeping young adults. In Proceedings of the 3rd International Symposium on Human Behaviour in Fire Interscience Communications, Belfast, UK. 291-302.

Ball, M., & Bruck, D. (2004b). The salience of fire alarm signals for sleeping individuals: a novel approach to signal design. In Proceedings of the 3rd International Symposium on Human Behaviour in Fire Interscience Communications, Belfast, UK. 303-314.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, 70, 614 – 636.

Berry, C. H. (1978). Will your smoke detector wake you? Fire Journal, July, 105-108.

Berkowitz, J.P. & Casali, S.P. (1990) Influence of age on the ability to hear telephone ringers of different spectral content. Proceedings of the Human Factors Society 34th Annual Meeting, Santa Monica, CA: Human Factors Society, 132-136.

Bonnet, M.,(1982).Performance during sleep. in Biological Rhythms, Sleep and Performance. W.B. Webb (Ed.) J. Wiley & Sons, 205-237.

Bonnet, M. H., Johnson, L. C., & Webb, W. B. (1978). The reliability of arousal threshold during sleep. Psychophysiology, 14: 412-416.

Bowman, S.K., Jamieson D.G. & Ogilvie, R.D. (1995) Waking effectiveness of visual alerting signals. Journal of Rehabilitation Research and Development, 32 (1), 43-62.

Brennan, P. (1998). Victims and survivors in fatal residential building fires. *Human Behaviour in Fire- Proceedings of the First International Symposium,* Shields J (Ed.), Fire SERT Centre, University of Ulster, 157-166.

Bruck, D. (1999). Non-awakening in children in response to a smoke detector alarm. Fire Safety Journal, 32, 369-376.

Bruck, D. (2001). The who, what, where and why of waking to alarms: a review. Fire Safety Journal, 36: 623-639.

Bruck, D. & Ball, M. (2004). The effectiveness of different alarms in waking sleeping children. In Proceedings of the 3rd International Symposium on Human Behaviour in Fire Interscience Communications, Belfast, U.K. 279-290.

Bruck, D., & Bliss, R. A. (2000). Sleeping children and smoke alarms. . In Yamada T (ed) Proceedings of the Fourth Asia-Oceania Symposium on Fire Science and Technology, Asia-Oceania Association for Fire Science and Technology and Japan Association for Fire Science and Engineering, Tokyo, May, 603-612.

Bruck, D. & Brennan, P., (2001). Recognition of fire cues during sleep, In Shields J (Ed.) Human Behaviour in Fire- Proceedings of the Second International Symposium (Boston), Interscience Communications, London UK. 241-252.

Bruck, D. & Horasan, M. (1995). Non-arousal and non-action of normal sleepers in response to a smoke detector alarm. Fire Safety Journal, 25, 125-139.

Bruck D. & Pisani D. (1999). The effect of sleep inertia on decision making. Journal of Sleep Research, 8, 95-103.

Busby, K. A, Mercier, L., Pivik, R. T. (1994). Ontogenic variations in auditory arousal threshold during sleep. Psychophysiology, 31, 182-8.

Carskadon, M. A., & Dement, W. C. (2000). Normal human sleep: An overview. In M. H. Kryger, T. Roth, W. C. Dement (Eds.), Principles and practice of sleep medicine (pp. 15-25). Philidelphia: W. B. Saunders Company.

Carskadon, M.A., Harvey, K. & Dement, W.C. (1981) Sleep loss in young adults: an EEG study. Sleep, 4, 299-312.

Cruickshanks, K. J., Wiley, T., Tweed, T., Klein, B., Klein, R., Mares-Perlman, & Nondahl, D. (1998). Prevalence of hearing loss in older adults in Beaver dam, Wisconsin: the epidemiology of hearing loss. Am J Epidemiol, 148: 879-86.

Dinges, D.F. Napping patterns and effects in human adults. In D.F. Dinges & R.J. Broughton (eds) Sleep and Alertness: Chronobiological, Behavioural and Medical Aspects of Napping. Raven Press, New York, 171-204.

Dinges, D.F., Orne, E.C., Evans F.J. & Orne, M.T. (1981) Performance after naps in sleep inducing environments. In L.C. Johnson, D. I. Tepas, W.P. Coloquhoun & M.J. Colligan (eds) Biological Rhythms, Sleep and Shift Work. Spectrum: New York, 539-552.

Edworthy, J., & Stanton, H. (1995). A user-centered approach to the design and evaluation of auditory warning signals: 1. methodology, Ergonomics, 38,11, 2262-2280.

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving Auditory Warning Design: Relationship between warning sound parameters and perceived urgency, Human Factors, 33, 2312-2326.

Edworthy, J., Clift-Matthews, W. & Crowther, M. (1998). In Contemporary Ergonomics 1998 (Ed, Hanson, M. A.) Taylor & Francis, London, 316 - 320.

Ferrara, M. & De Gennaro, L. (2000) The sleep inertia phenomenon during the sleep-wake transition: Theoretical and operational Issues. Aviat Space Environ Med, 71:843-8.

Frisina, D.R. & Frisina, R.D. (1997) Speech recognition in noise and presbycusis; relations to possible neural mechanism. Hearing Research, 106(1),95-104 .

Gibson, J.J. (1979). The ecological approach to visual perception. Houghton Mifflin Company, Boston, MA.

Grace, T., Woodger, N., & Olsson, P. (2001). On the use of voice alarms. Human Behaviour in Fire- Proceedings of the Second International Symposium, Shields J (Ed.), Massachusetts Institute of Technology, 185-196.

Groth-Marnat, G. (2000). Neuropsychological Assessment in Clinical Practice: A Guide to Test Interpretation and Integration. Wiley and Sons, New York.

Hasofer A.M., & Bruck D. (2004). Statistical analysis of response to fire cues. J Fire Safety, 39, 663-688.

Hasofer, M., Thomas, I.R., Bruck, D., & Ball, M. (2005). Statistical modelling of the effect of alcohol and sound intensity on response to fire alarms. Proceedings of the 8th International Symposium of the International Association for Fire Safety Science, Beijing, Sept., 507-518.

Haas, E.C.& Edworthy, J., (1996) Designing urgency into auditory warnings using pitch, speed and loudness. Computing and Control Engineering Journal, August, 193-198.

Hellier, E., Edworthy, J., Weedon, B., Walters, K. & Adams, A. (2002). The perceived urgency of speech warnings' semantics versus acoustics. Human Factors, 44, 1 – 16.

ISO 7029 (1984). Acoustics –threshold of hearing by air conduction as a function of age and sex for ontologically normal persons. (1st Edition) International Organisation for Standardisation (ISO). Geneva, Switzerland.

ISO 8201(1987). Acoustics - audible emergency evacuation signal. International Organisation for Standardisation (ISO), Geneva, Switzerland.

Johnson, L. C., Church, M. W., Seales, D. M. & Rossiter, V. S. (1979). Auditory arousal thresholds of good and poor sleepers with and without flurazepam. Sleep, 1(3), 259-270.

Johnson, L. C., Spinweber, C. L., Webb, S. C. & Muzet, A. G.(1987). Dose level effects of triazolam on sleep and response to a smoke detector alarm. Psychopharmacology, 91, 397-402.

Kahn, M. J. (1984). Human awakening and subsequent identification of fire-related cues. Fire Technology, 20(1), 20-26.

Kim, S., Frisina R & Frisina R.D. (2002) Effects of age on contralateral suppression of distortion product octoacoustic emissions in human listeners with normal hearing. Audiology & Neurotology, 7:348-357.

Lawrence, A.B. (1970) Architectural acoustics. London: Elsevier Publishing Co. Ltd.

Lee, A., (2005) The audibility of smoke alarms in residential homes. CPSC-ES-0503. US Consumer Product Safety Commission, Washington D.C. 20207-0001. February 2005.

Lee, A., Midgett, J. & White, S., (2004) A review of the sound effectiveness of residential smoke alarms. CPSC-ES-0502. US Consumer Product Safety Commission, Washington D.C. 20207-0001. December 2004 (Revised).

Lezak, M.D. (1995) Neuropsychological Assessment (3rd Ed) Oxford University Press: New York.

Liberman, A.M. & Mattingly, I.G. (1989) A specialisation for speech perception. Science, 243, 4890, 489-494.

Lindsley, O.R. (1957). Operant behaviour during sleep: a measure of sleep depth. Science 126,1290-1291.

Lucas, J. W., Schiller, J. S. & Benson, V. (2004). Summary health statistics for US adults. Vital Health Statistics 10 (218), 5, 34-37.

Medco Health Solutions, Inc (Latest News October 17th 2005) http://www.medco.com/medco/corporate/home.jsp accessed 5th April 2006.

Mulvany, D. (2004). in Report of the committee on signalling systems for the protection of life and property. (National Fire Protection Association 72 National Fire Alarm Code, A2006 Report on proposals), p.112. Downloaded from http://www.nfpa.org/itemDetail.asp?categoryID=817&itemID=23682 on 12th August 2005.

Nakano, M. & Hagiwara, I. (2000). Experimental study on starting time of evacuation in sleeping condition. In Yamada T (ed) Proceedings of the Fourth Asia-Oceania Symposium on Fire Science and Technology, Asia-Oceania Association for Fire Science and Technology and Japan Association for Fire Science and Engineering, Tokyo, May, 263-274.

Nober, E. H., Peirce, H. & Well, A. D. (1981a). Acoustic spectral characteristics of household smoke detector alarms. Fire Journal, May, 94-98 (+144).

Nober, E. H., Peirce, H. & Well, A. (1981b). Waking effectiveness of household smoke and fire detection devices. Fire Journal. July, 86-91(+130).

Ohayon MM; Carskadon MA; Guilleminault C; Vitiello MV, (2004) Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. Sleep, 2004, 27 (7), 1255-73.

Portas, C., Krakow, K., Allen, P., Josephs, O., Armony, J. L. & Frith, C. D. (2000). Auditory processing across the sleep-wake cycle: simultaneous EEG and fMRI monitoring in humans, Neuron, 28, 991-999.

Proulx, G., & Laroche, C. (2003). Recollection, identification and perceived urgency of the temporal-three evacuation signal. Journal of Fire Protection Engineering, 13, 67-82.

Proulx, G., & Sime, J.D. (1991). To prevent panic in an underground emergency; why not tell people the truth? In G. Cox & B. Langford (Eds.), Fire Safety Science-Proceedings of the Third International Symposium. Elsevier, London, 843-852.

Quirt, J.D. (1985) Sound transmission through building components. Building science Insight '85 Noise Control in Buildings, Institute for Research in Construction, Research Council Canada. Downloaded from http://irc.nrc-cnrc.gc.ca/bsi/85-3_E.html on 15th August 2005.

Reitan, R.M. and Wolfson, D. (1985). The Halstead-Reitan neuropsychological test battery: Therapy and clinical interpretation. Tucson, AZ: Neuropsychology Press.

Robinson, D. A. (1986). Sound transmission loss from corridors to rooms: implications for locating fire alarm sounders. Fire Technology, 22(2), 122-135.

Sekizawa, A. (1991). Statistical analyses on fatalities characteristics of residential fires. Fire Safety Science- Proceedings of the Third International Symposium. G. Cox & B. Langford (Eds.) London: Elsevier Science, 475-484.

Snyder, F. & Scott, J. (1972). The psychophysiology of sleep. In Handbook of Psychophysiology. N.S. Greenfield & R.A. Sternback (eds). New York, Holt, Rinehart and Winston.

Stanton, N. & Edworthy, J. (1998). Auditory affordances in the intensive treatment unit . Applied Ergonomics, 29(5), 389-394.

Thomas, I. & Brennan P. (2002). Injuries and fatalities in apartment building fires. Fire

Safety Science- Proceedings of the Seventh International Symposium. D.D.
Evans (Ed.) London: Elsevier Science, 1085-1096.

Thomas, I., & Brennan, P. (2004). Occupants, ignition and fire outcomes. In
Proceedings of the 3rd International Symposium on Human Behaviour in Fire.
Interscience Communications, Belfast, UK. 45-56.

Wallace, M.S., Ashman, M.N. & Matjasko, J. (1994). Hearing acuity of anesthesiologists
and alarm detection. Anesthesiology, 81: 13-28.

Weir, C. (1976). Auditory frequency sensitivity in the neonate: a signal detection
analysis. Journal of Experimental Child Psychology, 21, 219-225.

Williams, H. L., Hammack, J. T., Daly, R. L., Dement, W. C. & Lubin, A.L. (1964).
Responses to auditory stimulation, sleep loss and the EEG stages of sleep.
Electroencephalography and Clinical Neurophysiology, 16, 269-279.

Wilson, W. P. & Zung, W. K. (1966). Attention, discrimination, and arousal during sleep.
Archives of General Psychiatry, 15: 523-528.

Zepelin, H., McDonald, C. S. & Zammit, G. K. (1984). Effects of age on auditory
awakening. Journal of Gerontology, 39(3), 294-300.

Appendix A: Spectral analyses of four signals tested
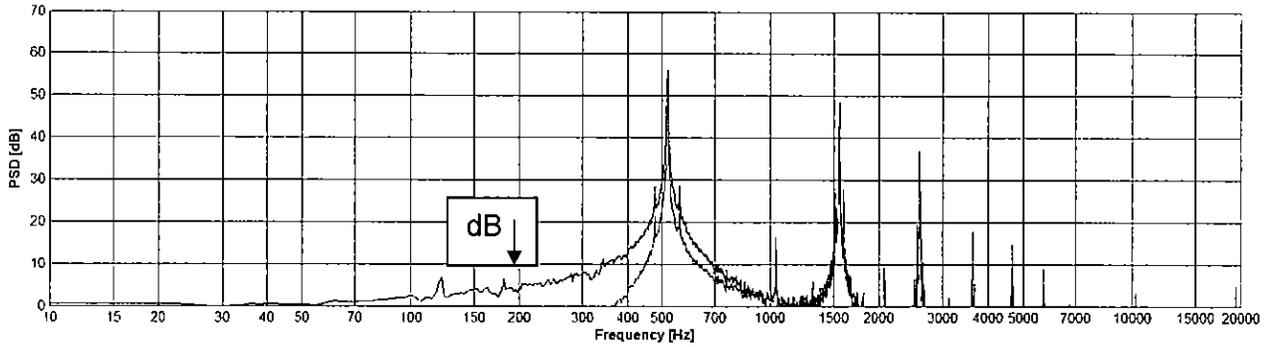
Figure A.1: Spectral analysis of the 60 dBA mixed T-3 sound file where the blue line indicates dB and the black line dBA.[23]
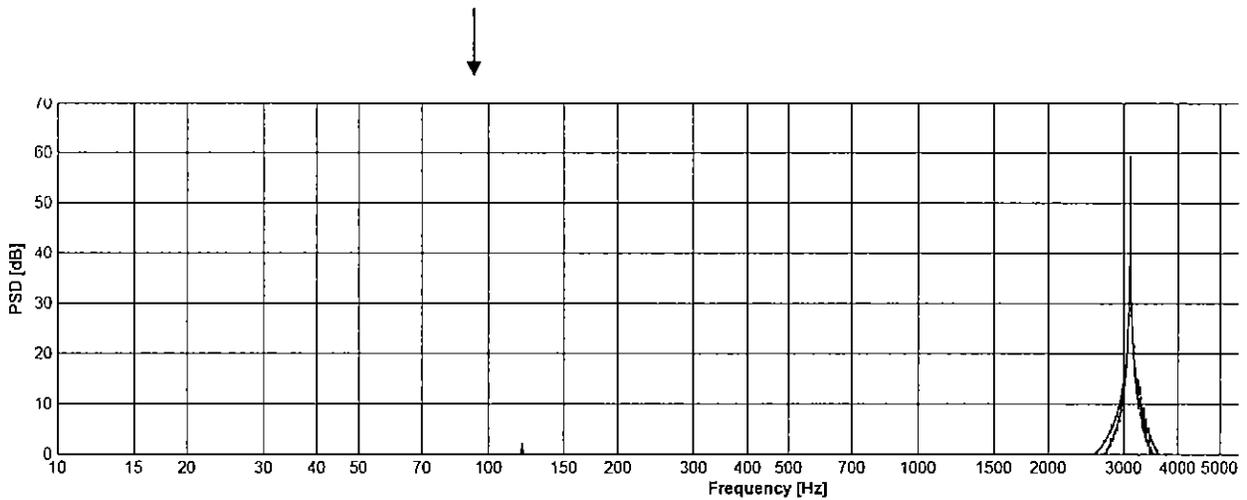
Figure A.2: Spectral analysis of the 60 dBA high T-3 sound file where the blue line indicates dB and the black line dBA.

---

[23] The dBA scale is a weighting of the dB scale which discriminates against low frequencies (below about 550 Hz) (Lawrence, 1970). It has become the most widely used scale across the audible range. In the graphs the blue line (dB) has more volume at lower frequencies.
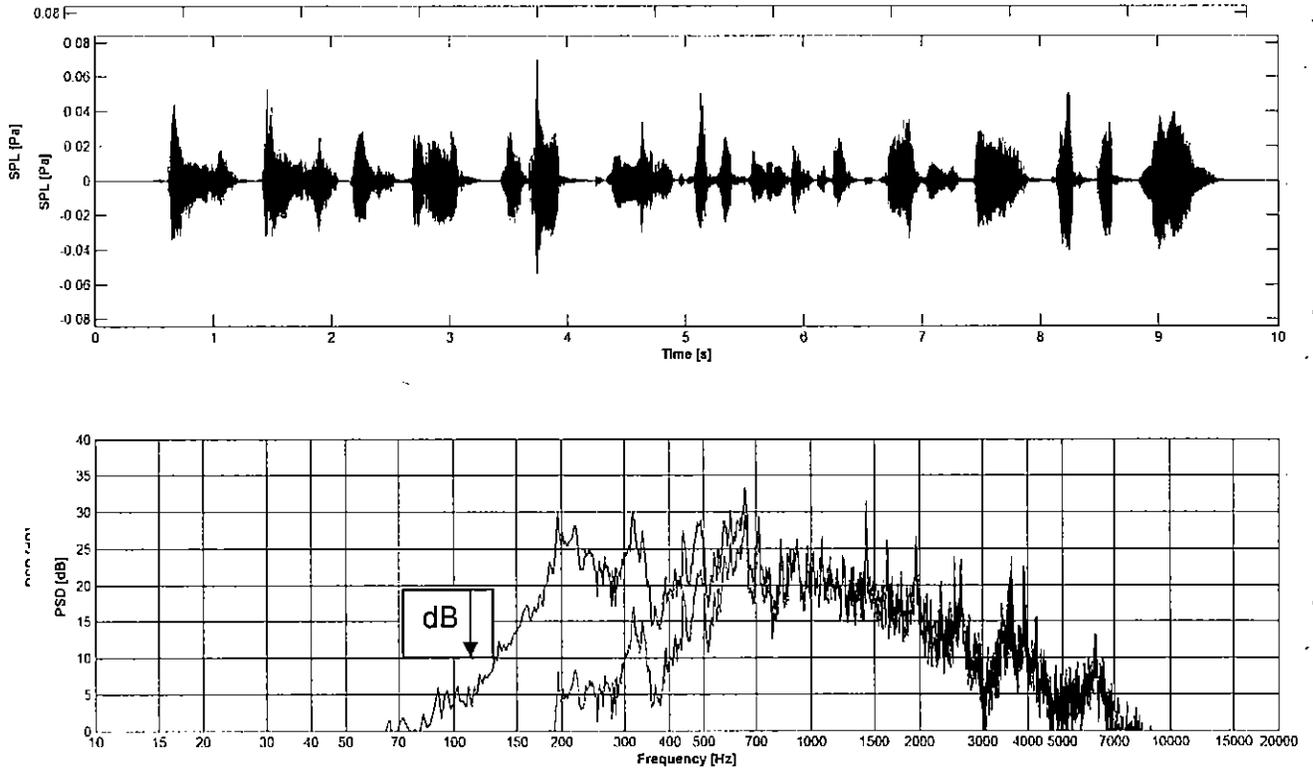
Figure A.3: *Top graph*: Sound pressure level as a function of time for the 60 dBA male voice sound graph. *Lower graph*: Spectral analysis of the 60 dBA male voice sound file where the blue line indicates dB and the black line dBA.
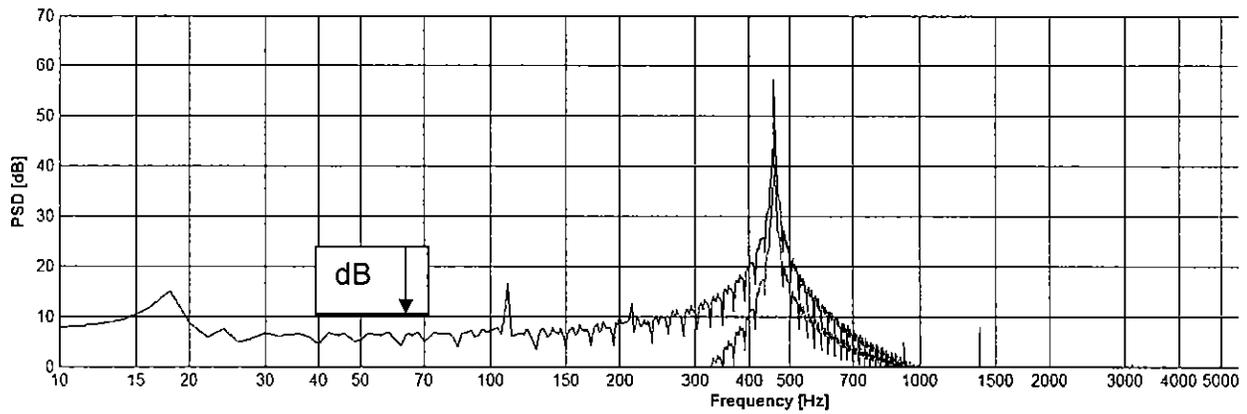
Figure A.4: Spectral analysis of the 500 Hz T-3 where the blue line indicates dB and the black line dBA.

Appendix B: The breakdown of the age and sex of subgroups

Table B.1: Number, age and sex details of participants who completed the Trail Making Task A & B. Number in brackets indicates number unable to complete task.

Age	N males	N females	Total
65-74	12(1)	12(2)	24 (3)
75-85	9 (0)	6 (3)	15 (3)
Total	21 (1)	18 (5)	39 (6)

Table B.2: Number, age and sex details of participants who completed all the physical performance tasks after second awakening each night. Number in brackets indicates number unable to complete task.

Age	N males	N females	Total
65-74	12(1)	13 (1)	25 (2)
75-85	8 (1)	8 (1)	16 (2)
Total	20 (2)	21 (2)	41 (4)

Appendix C: Details of those who dropped out of the study

1. Five people (ID 14, 16, 41, 45, 59) passed their hearing test and then decided they did not wish to participate, or were no longer available for other reasons, before they started the study.

2. In the course of conducting the study three people dropped out, not completing all signals. The reasons were as follows:
   - 1X developed a skin rash from the application of the electrodes (ID 28)
   - 1X had difficulty falling asleep after the first awakening on the second night and was no longer willing.(ID 25)
   - 1X participant (ID 40) inadvertently received the same two signals on both nights and when participated for a third night was unable to go to sleep. (This data showed some variability as indicated below)

Table C.1: Decibel level at which awoke to the same signal across two nights.

	High T-3	500 Hz
Night 1	85	50
Night 2	70	55

Appendix D: Report on recruitment of participants for the project
By Belinda Gibson, Project Officer (Recruitment)

Recruitment process

The recruitment process can be broken down into two segments, the first is the identification of existing contacts or the creation of new ones and the second segment consisted of approaching such contacts. The solitary way in that the recruitment process began and was sustained was through networking.

In the initial stages of the recruitment process networking consisted of identifying any existing contacts and then following them up. By identifying and existing contacts we were able to then build a contact list that consisted of various organisations and people who could connect us with an elderly group of some sort. The first contact made was with the Susan Feldman, the head of the AURA department at Victoria University. AURA's principle task is to conduct research in the area of the elderly population. Meeting with Susan was essential in the recruitment process and as a result helped create the foundation needed to recruit participants. Susan was able to connect us with a number of organisations and people within the elderly community. Having direct contact details and Susan's permission to mention her name made the recruitment process much more successful.

In particular, Susan was able to connect me with the Multicultural Service Officer at Centrelink in Footscray and Newport, Melbourne. As this officer has a strong presence within the local community she was able to pass on information about the project to a number of organisations, she was also able to set up appointments to meet with managers and coordinators who had service members who fit our criteria. This technique facilitated the inclusion of a variety of cultures.

In addition to contacting existing contacts and building on those, an Internet search was also conducted. The Internet search involved the investigation of names and numbers of organisations such as bowling clubs and senior citizen groups in and around

metropolitan Melbourne.

From the Internet search a list of 40 potential groups and organisations were recorded. The next step was to contact these groups and organisations and introduce the study. The desired form of contact was via the telephone, if this wasn't possible, then information packs were sent to the organisation (information packs were also sent to organisations or groups who were interested in participating in the study). The information packs outlined clearly and precisely what would be involved for anyone who participated in the study. The pack described in full the criteria and each and every step involved in the study as well as the financial benefit that the organisation and participant would receive if they chose to participate. Following the delivery of information packs each organisation was phoned and the study and their participation was discussed.

After my initial meetings with various groups and those with Susan Feldman and the Multicultural Officer the recruitment process took on a snowball effect. A meeting was organised with one person and then that person would connect me to someone else and so on.

Steps of recruitment

Below is a basic flowchart that simplistically outlines the steps taken in the recruitment process. This process is continual, once step three was reached step one was readdressed.

Step 1 - Research
-Locating possible clubs/ groups with members over the age of 65 years.

↓

Step 2 - Networking
-Talking to existing contacts or alternatively
-Connecting to the organisations found through research.

↓

Step 3 - Presentations/Meetings
-Presenting the project to managers/co-ordinators
-Presenting project to service users.
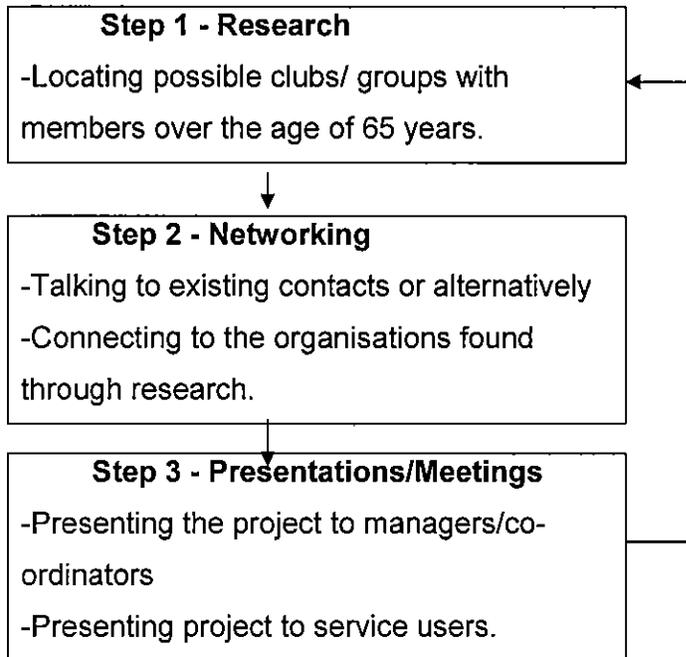
Appendix E: Hearing criteria and comparison with norms

Part 1: Hearing criteria guidelines

Mean pure tone air-conduction decibel hearing level (dBHL) by frequency, ear, sex and age, with the addition of one standard deviation. Normative values and standard deviations from data collected in 1993-1995 from Beaver, Wisconsin USA; n=3,753 aged 48-92years. (Cruickshanks, Wiley, Tweed, Klein, Klein, Mares-Perlman & Nondah, 1998).

Pass = air conduction thresholds in dBHL ≤ the values below for each ear

Fail = air conduction thresholds in dBHL > the values below for either ear

Table E.1: Minimum pure tone air conduction dBHL thresholds required to pass screening criteria for _females_ across different frequencies.

Age (years)	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz
60 – 69	25	30	30	35	45
70 – 79	35	40	45	50	55
80 - 92	45	50	55	65	70

Table E.2: Minimum pure tone air conduction dBHL thresholds required to pass screening criteria for _males_ across different frequencies.

Age (years)	500 Hz	1000 Hz	2000 Hz	3000 Hz	4000 Hz
60 – 69	25	25	45	65	75
70 – 79	35	40	55	70	80
80 - 92	50	55	70	80	85

Part 2: Comparison of mean auditory thresholds, when awake, for the participants in the current study with normative data from Cruickshank et al. (1998).

Table E.3: Comparison of mean thresholds (and standard deviation) for the participants in the current study with normative data from Cruickshank et al. (1998) for ages 60-69.

	Study n	500 Hz		1,000 Hz		2,000 Hz		3,000 Hz		4,000 Hz	
		Norms	Study	Norms	Study	Norms	Study	Norms	Study	Norms	Study
Females right ear	10	12.9 (14.1)	12.0 (4.7)	15.7 (14.4)	11.2 (5.7)	18.0 (16.0)	15.5 (6.8)	21.8 (16.7)	16.7 (7.4)	27.5 (17.6)	19.4 (8.9)
Males right ear	7	12.2 (13.6)	12.3 (6.8)	15.8 (14.6)	14.3 (11.3)	26.2 (20.9)	20.8 (11.8)	43.3 (23.6)	27.1 (17.0)	54.0 (23.7)	35.0 (14.4)
Females left ear	10	13.5 (13.1)	10.0 (4.6)	16.1 (14.0)	9.7 (3.9)	18.9 (15.3)	14.4 (7.9)	23.9 (17.2)	18.0 (7.5)	28.7 (17.6)	21.2 (8.8)
Males left ear	7	12.1 (11.6)	10.0 (5.8)	16.4 (12.8)	12.0 (8.5)	28.6 (19.3)	18.4 (11.6)	46.6 (21.6)	33.5 (17.7)	55.1 (21.7)	38.8 (16.7)

Table E.4: Comparison of mean thresholds (and standard deviation) for the participants in the current study with normative data from Cruickshank et al. (1998) for ages 70-79.

	Study n	500 Hz		1,000 Hz		2,000 Hz		3,000 Hz		4,000 Hz	
		Norms	Study	Norms	Study	Norms	Study	Norms	Study	Norms	Study
Females right ear	8	19.7 (16.6)	15.0 (10.3)	22.9 (17.4)	16.0 (4.7)	27.6 (18.6)	17.2 (9.0)	32.5 (19)	26.5 (11.0)	39.0 (19.8)	35.9 (13.9)
Males right ear	9	18.8 (14.5)	15.8 (6.1)	23.6 (17.5)	17.4 (6.6)	35.5 (21.1)	25.7 (14.7)	51.7 (20)	30.1 (14.9)	62.0 (19.0)	42.2 (13.0)
Females left ear	8	19.1 (16.5)	15.0 (9.0)	22.1 (17.6)	13.4 (7.5)	28.3 (18.7)	16.6 (8.4)	34.9 (19.1)	26.1 (12.4)	40.9 (19.8)	40.4 (14.2)
Males left ear	9	20.3 (17.1)	18.2 (9.8)	25.3 (19.4)	16.1 (7.4)	38.8 (21.5)	23.0 (18.2)	56.1 (19.1)	30.5 (17.8)	64.6 (19.3)	44.2 (17.8)

Table E.5: Comparison of mean thresholds (and standard deviation) for the participants in the current study with normative data from Cruickshank et al. (1998) for ages 80-89.

	Study n	500 Hz		1,000 Hz		2,000 Hz		3,000 Hz		4,000 Hz	
		Norms	Study	Norms	Study	Norms	Study	Norms	Study	Norms	Study
Females right ear	3	30.4 (16.8)	22.5 (17.7)	34.9 (17.8)	22.3 (2.9)	41.4 (18.3)	28.3 (2.9)	47.1 (18.1)	37.5 (10.6)	54.3 (17.5)	37.0 (9.9)
Males right ear	4	31.8 (22.9)	13.7 (13.8)	38.2 (22.7)	16.7 (16.5)	52.3 (19.9)	42.5 (27.2)	63.5 (17.1)	51.7 (22.9)	70.5 (17.3)	60.0 (10.8)
Females left ear	3	21.9 (16.4)	27.5 (24.8)	34.4 (18.0)	25.0 (7.1)	41.3 (17.7)	37.5 (3.5)	48.3 (16.8)	45.0 (14.1)	54.6 (17.0)	42.0 (8.0)
Males left ear	4	27.8 (18.1)	16.7 (12.9)	34.8 (19.5)	21.2 (16.5)	50.4 (17.7)	36.7 (26.9)	63.4 (16.4)	58.5 (10.7)	71.3 (16.8)	65.0 (10.8)

Appendix F: Details of those who failed their hearing screening test

A total of 59 individuals were tested for their hearing. Nine (15%) failed, see Table F.1 for details of their age and sex.

Table F.1. Frequency of those who failed their hearing test by age and sex.

	65-74 yrs	75-85 yrs
males	3	2
females	2	2

Details of the reasons why people failed is as follows:

Failed lower frequencies only (500 Hz and 1 kHz):
- o one ear (n=1)
- o both ears (n=1).

Failed all/most frequencies (at least at and above 2 kHz):
- o one ear (n=2)
- o both ears (n=3).

Failed higher frequencies only (3 and 4 kHz):
- o one ear (n=2).

Appendix G: Sound measurement, calibration and signal delivery aspects

Speakers used for delivery: Altec Lansing Technology Powered Audio
System. VS.2120.

Sound meter type: Lutron Model SI-4001 (2) both recalibrated on 31/8/05, using
signals in the range of 80-130 dB, dBA and dBC, and frequencies of 244 Hz-1000 Hz .

Creation of sound files: For the sound delivery program it was necessary to have sound
files of each signal at levels from 35 dBA to 95 dBA in 5 dBA increments. This was
done in a sound attenuated TV studio at a Victoria University campus.

Once a signal was available at a particular volume it was played through the speakers
to be used in the study and the decibel level adjusted using acoustic software (Sound
Forge 6) so that it was measured to be received at a particular volume (eg 35 dBA) as
assessed by the sound meter. A tolerance range of plus or minus 1 dBA was allowed.
Table G.1 shows the sound meter settings.

Table G.1: Sound meter settings for creating the different sound files.

Meter settings	For recording of the following sound files
30-80 dBA	35-60 dBA inclusive
50-100 dBA	65-85 dBA inclusive
80-130 dBA	90-94 dBA inclusive

Thus for each of the four sounds, 13 sound files at different volumes were created. The
settings on the sound meter were "slow response" and "maximum hold" for all sound
level assessments. The volumes would fluctuate but the most dominant level was used.

Calibration in bedrooms: The procedure to be followed in the bedrooms of participants
was as follows: The mixed T-3 60 dBA sound file was played from the speakers, which
were located approximately one metre from the pillow, where the sound meter was

placed. The volume knob on the speakers were adjusted so that the sound level meter was showing as close as possible to 60 dBA (using all the settings as above).

The reliability of this calibration method in terms of what sound levels were at the pillow for all four signals was checked in the sleep laboratory for four volume levels and the results as shown in Table 2 were obtained. As all the values were received within 2 dBA of the original sound file level this was considered satisfactory. It is important to note however, that different bedrooms will have different sound absorption properties so these values will vary as a function of the materials and furnishings in each bedroom.

Table G.2: Actual sound levels measured at the pillow in the Victoria University sleep laboratory, using the mixed T-3 60 dBA sound file as the reference file for calibration.

	50 dBA sound file	60 dBA sound file	70 dBA sound file	80 dBA sound file
Mixed T-3	49.8	59.6	69.5	80.9
Male voice	51.6	59.9	70.4	79.2
High T-3	49.3	58.7	70.2	79.6
500 Hz	48.1	59.6	69.2	78.7

Appendix H: Text of male voice alarm

Danger! Danger! There is fire. Wake up!

You must get up and investigate.

There is fire. Get up now!

Duration of each loop was 9 seconds. Looped three times for each 30 second segment.

Appendix I: Consent Form for Research Participants and Information Sheet

We would like to invite you to be a part of a study that aims to develop the best smoke alarm signal possible for waking different groups within the population, including older people.

I, (insert name)

of (insert address)

certify that I am at least 18 years old and that I am voluntarily giving my consent to participate in the study entitled:

Reducing fire deaths in the aged: optimising the smoke alarm signal

being conducted at Victoria University of Technology by:

Professor Dorothy Bruck

I certify that the objectives of the study, together with any risks and safeguards associated with the procedures listed hereunder to be carried out in the research, have been fully explained to me by:

Belinda Gibson

and that I freely consent to participation involving the use on me of these procedures.

- o Be interviewed about my sleep quality, my medication intake, hearing capabilities and normal daily activities
- o Be asked questions about my memory, perhaps including answering some questions from a questionnaire
- o Participate in a free hearing test at a H.E.A.R. clinic
- o Attachment of surface electrodes to measure sleep
- o Awakening from sleep by different signals (beeps or voice)
- o Completion on several occasions of the Trail Making Task, 15 metre walk within my home and phone call (all timed).

I certify that I have had the opportunity to have any questions answered and that I understand that I can withdraw from this study at any time and that this withdrawal will not jeopardise me in any way.

I have been informed that the information I provide will be kept confidential.

Signed: .... }

Witness other than the researcher: }   Date: ....

....}

Any queries about your participation in this project may be directed to the researcher (Professor Dorothy Bruck ph. 9919 2336). If you have any queries or complaints about the way you have been treated, you may contact the Secretary, University Human Research Ethics Committee, Victoria University of Technology, PO Box 14428 MCMC, Melbourne, 8001 (telephone no: 03-9688 4710).

Information about the Research Project

Title: Reducing fire deaths in the aged: optimising the smoke alarm signal

At Victoria University our research team has, for several years now, been looking at the question of what smoke alarm signal is the best for waking up people. We have tested children and young adults and the results suggest that the current signal may not be as good as some alternative signals. We now want to test how well older people will wake up to different signals. This is especially important as we know that more older people die in fires than any other age group and most fatal fires occur during the time when people are asleep.

In this study we will be presenting some different signals to volunteers while they are asleep in their own home. Speakers will be set up in the bedroom. The sounds will be presented softly at first and then louder, and we are interested to know at what volume people will wake up to the different signals. The loudest volume is still within safe hearing limits. We want to always present the signals in the same type of sleep and because sleep changes across the night we will need to monitor the different stages of sleep of our volunteers. This is done by attaching ten small surface electrodes to the face and top of the head. A Sleep Technician is trained to do this and will present the signals from a hallway next to the bedroom.

The other issue that we are interested in is how groggy are people when they first wake up and does this change with age? With the first signal the volunteer will only need to press a button at their bedside. With the second signal, once the volunteer has been woken up they will sit up in bed and do a short paper and pencil task that they will have practised beforehand. This task involves connecting some dots in a certain order. They will then be asked to walk on a marked path for 15 metres around the inside of their home, ending at the phone. We then want them to make a short phone call to our answering machine, giving their name and address. These different tasks will all be timed by the Sleep Technician because we are interested to know how much the grogginess from sleep slows people down.

As the study involves disruption to sleep, volunteers need to be aware that they may be

sleepier than usual the next day and should be careful not to plan activities where sleepiness may be a problem. In particular the driving of a car should be avoided.

We need to present each volunteer with four signals during their sleep, and we will normally do two each night. Thus being involved will usually involve two nights of sleep study. However, if a person has trouble returning to sleep after the first awakening or some other problem arises, we may need three or even four nights. Because we realise that being part of our study involves some inconvenience we are paying each volunteer $200 for the awakenings to four signals. If for some reason not all four signals are completed (for example a volunteer chooses to withdraw) we will pay $30 for each night of involvement. If a social club has been helping us publicise the study we will be rewarding them separately.

For this project we need volunteers who meet our selection criteria. These are:
- Be aged between 65 and 85 years.
- Do not have a history of significant hearing problems in either ear.
- Usually do not have a lot of difficulty getting to sleep.
- Not be taking any medication that affects their sleep.
- Have no problem doing some simple paper and pencil tasks that involve counting and drawing some connecting lines
- Have no problem using the phone
- Be independently mobile (that is not in a wheelchair, but may use a walking frame or walking stick)

Because hearing levels are so important to this study all volunteers are asked to undertake a free hearing screening test. We will arrange this at a time that is convenient and at a H.E.A.R clinic as close as possible. We have taxi vouchers to cover the cost of getting to the clinic or, if a car is driven, we will cover the costs of the trip.

Thank you for your interest in our research. Contact details:
XXXXXXXXXXXXXXXXXXXXXXXX, Phone XXXXXXXXXX.

Appendix J: Demographics and Screening Form

(administered verbally by the Project Officer (recruitment)

Allocated ID number:

Name: Sex:

Date of Birth: Current Age:

Address:

Phone Number: (Best time to call?)

Do you have a history of any significant hearing problems in either ear?
(If yes, ...details)

Do you believe your hearing is *about average or above average* for someone of your age and sex? (If no, ...details)

How long does it usually take you to go to sleep when you first try to sleep at night?

If you wake up in the night do you usually have much difficulty getting back to sleep again? (If so, how long does it take you on average?)

Are you taking any medications regularly that might affect your sleep? (If so, what?)

Are you physically independently mobile? That is, you can walk satisfactorily either unaided, with a walking stick or with a walking frame?

Can you use a pen or pencil without any major difficulty?

Which nights of the week would suit you for the sleep tests? Any dates in next 4 weeks when you would not be available?

Cognitive functioning: If responses have given you any cause for concern administer the following two items:
Preface: I am going to ask you some things to see how your memory is. Is this OK?
Listen carefully, I am going to say three words. You say them back after I stop. Ready?
Here they are... ORANGE (pause), DOLLAR (pause), CHAIR (pause). Now repeat those words back to me. (Repeat once or twice if necessary.)
Spell EARTH forwards, then backwards. (Correct forward spelling if necessary)
If still concerned need to arrange a visit to administer the Mini Mental State to ensure cognitive functioning is acceptable.

Appendix K: Screening questionnaire re sleep deprivation and alcohol

Please complete this questionnaire prior to preparation for the sleep study.

ID Number _____

Please circle one: Night 1   Night 2   Night _____

1. Thinking about your sleep last night, compared to your usual sleep, was it: (please circle one of the options)

Much better than usual
A little better than usual
Same as usual
A little worse than usual
Much worse than usual

2. If you chose "much worse than usual", please comment on why your sleep was much worse. (Otherwise leave blank)

3. Have you consumed any alcohol since 4pm today? If so, please describe the type (beer, wine etc), the quantity and the time of day when it was consumed.

Type:
Quantity:
Time of Day:

In this research we are keen for your sleep to be as similar as possible on the different

nights of the study. Two factors that can especially affect your ability to wake up are

If you are quite sleepy from having had poor sleep on the previous night, or, if you have

consumed more than a glass or so of alcohol close to bedtime

If you think these may be of concern please discuss this with the Sleep Technician.

Thanks

(15)

Appendix L: Trail Making Task A
(Examples of version 1, of 4 different versions of both A and B – others were inverted or mirror image or both)

(17) (21)

(20) (19)

(16)

(18)

(4) (22)

(5)

(13) (6)

begin

(24)

(7) (1)

(14)

(2)

(8) (10) (3)

end

(9)

(11) (25) (23)

(12)
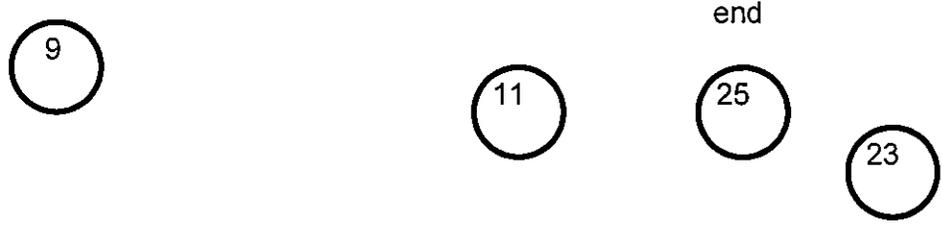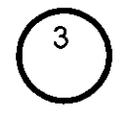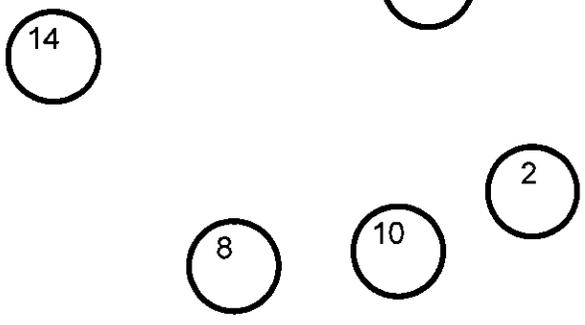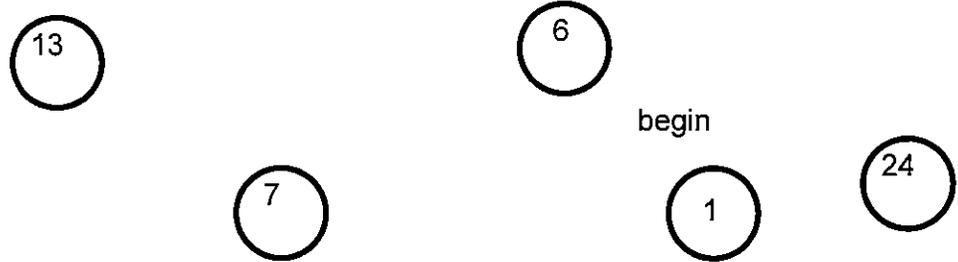
end

13

Appendix L (cont) : Trail Making Task B

10

8

9

I

D

B

4

3

begin

7

1

5

H

C

12

G

A

J

2

6

L

E

F

K

11

Appendix M: Graphs of the data as a function of the cumulative number of subjects.

Note: The purpose of these graphs was to help determine how many participants were needed to obtain relative stability in the data (see Results Section 5.3.2).
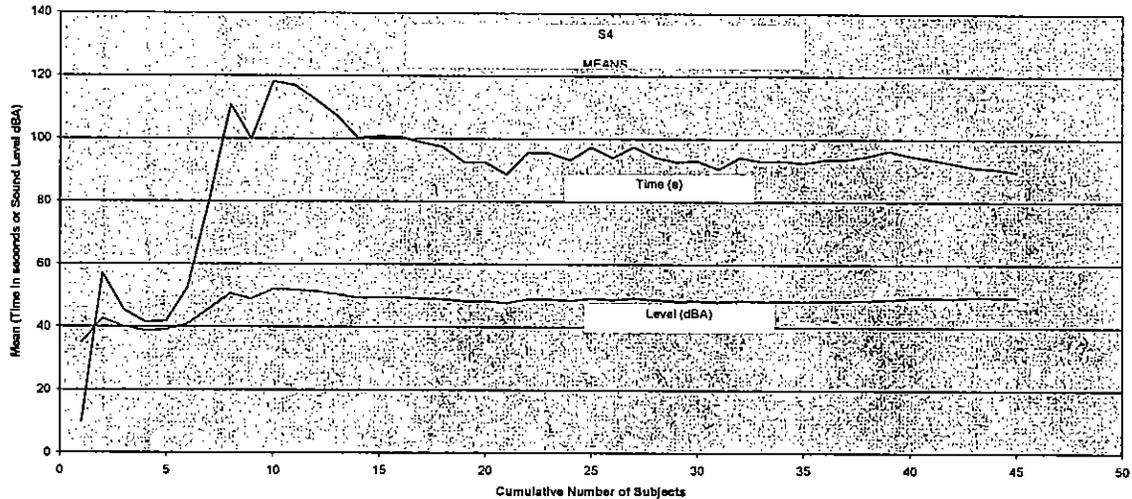
Figure M.1: Mean behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the mixed T-3 signal.
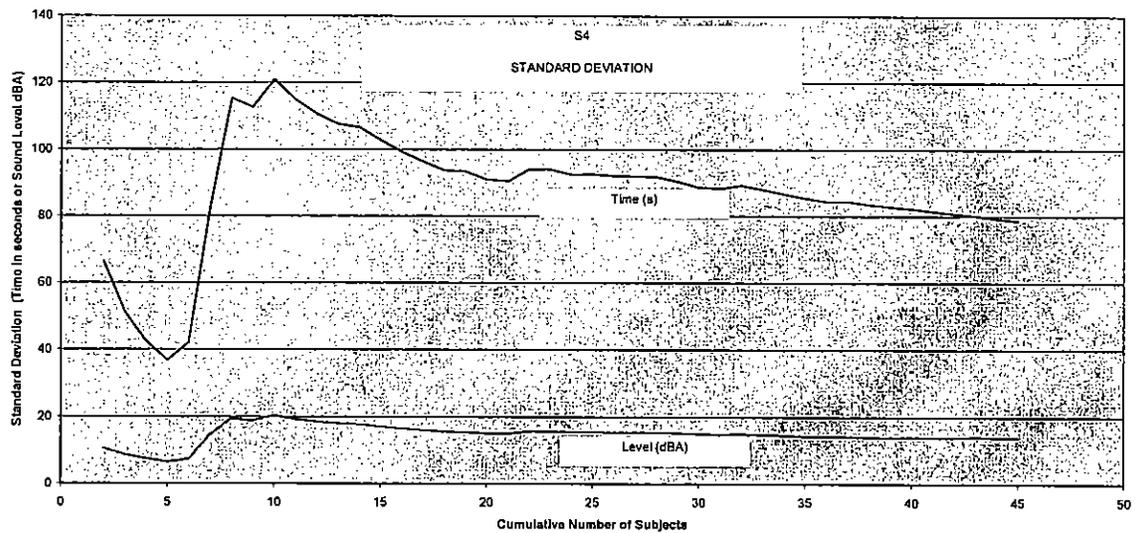
Figure M.2: Standard deviations of behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the mixed T-3 signal.
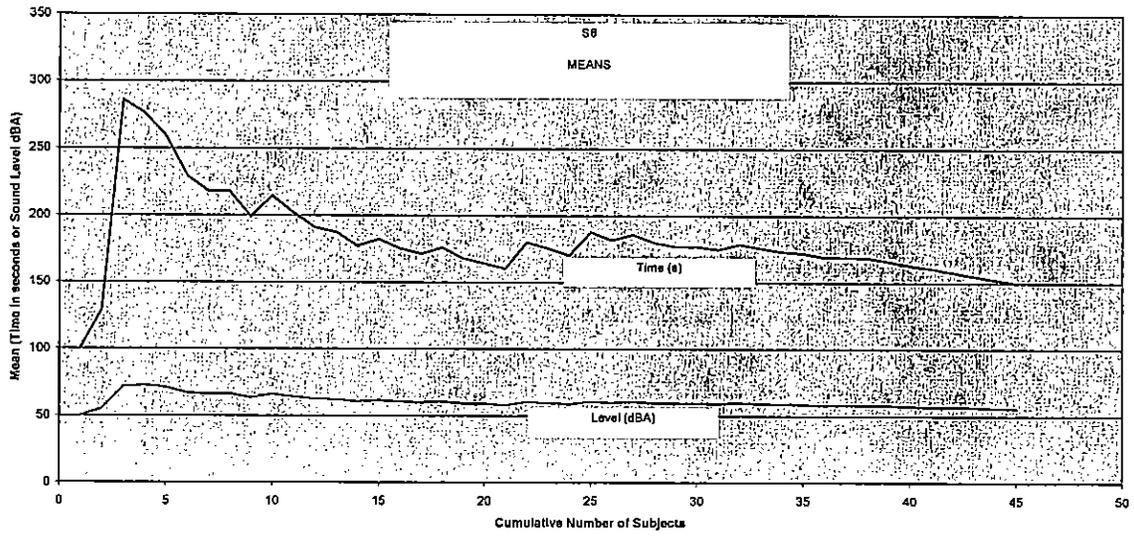
Figure M.3: Mean behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the male voice signal.
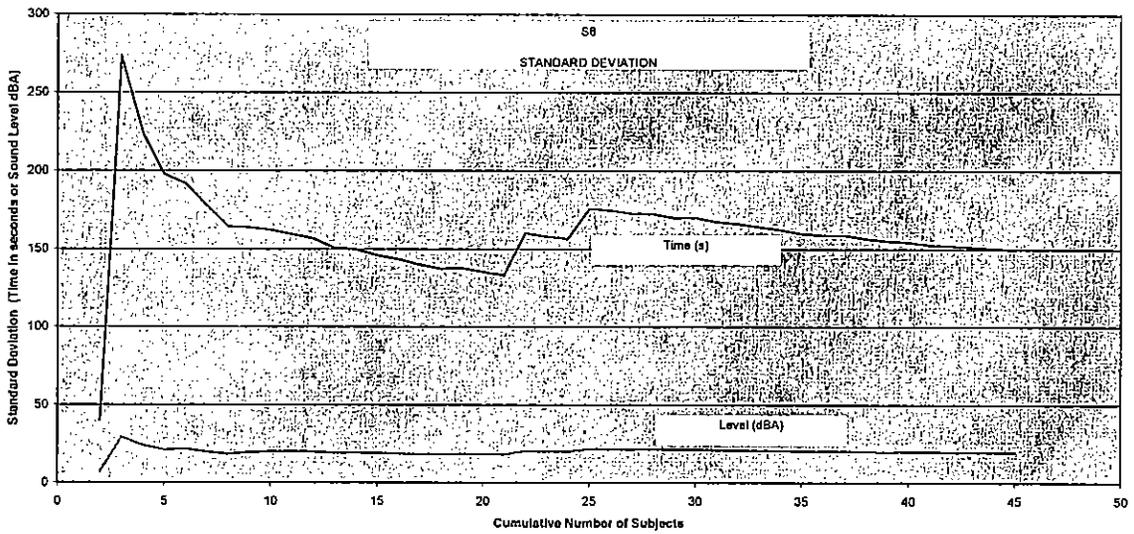

Figure M.4: Standard deviations of behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the male voice signal.
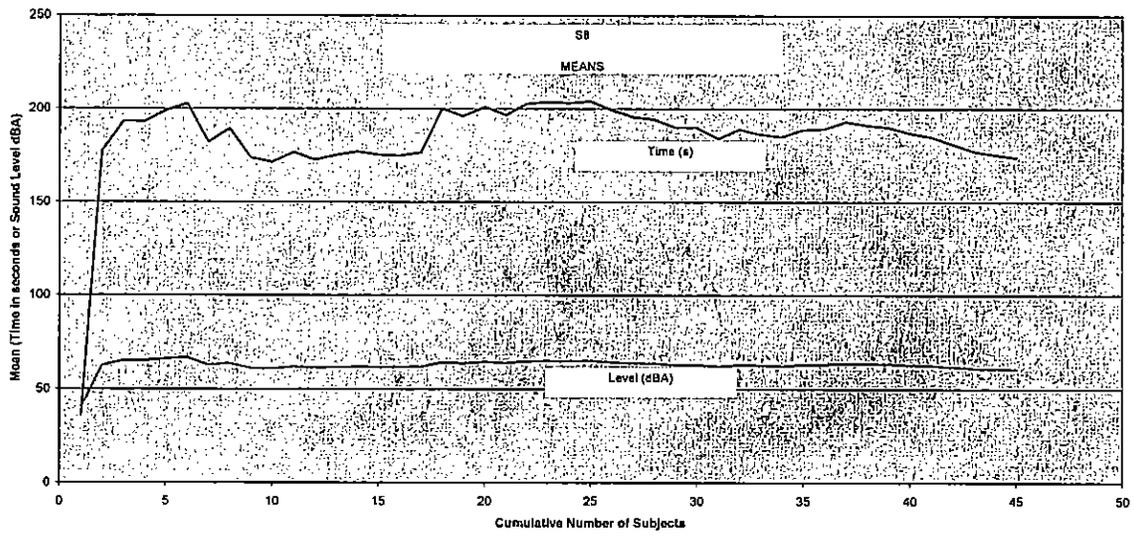
Figure M.5: Mean behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the high T-3 signal.
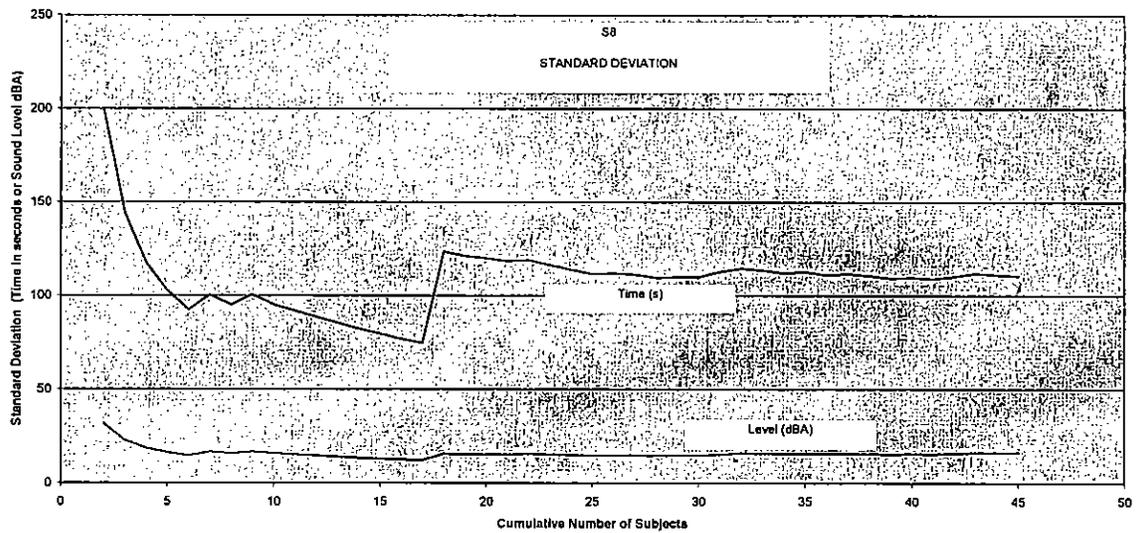

Figure M.6: Standard deviations of behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the high T-3 signal.
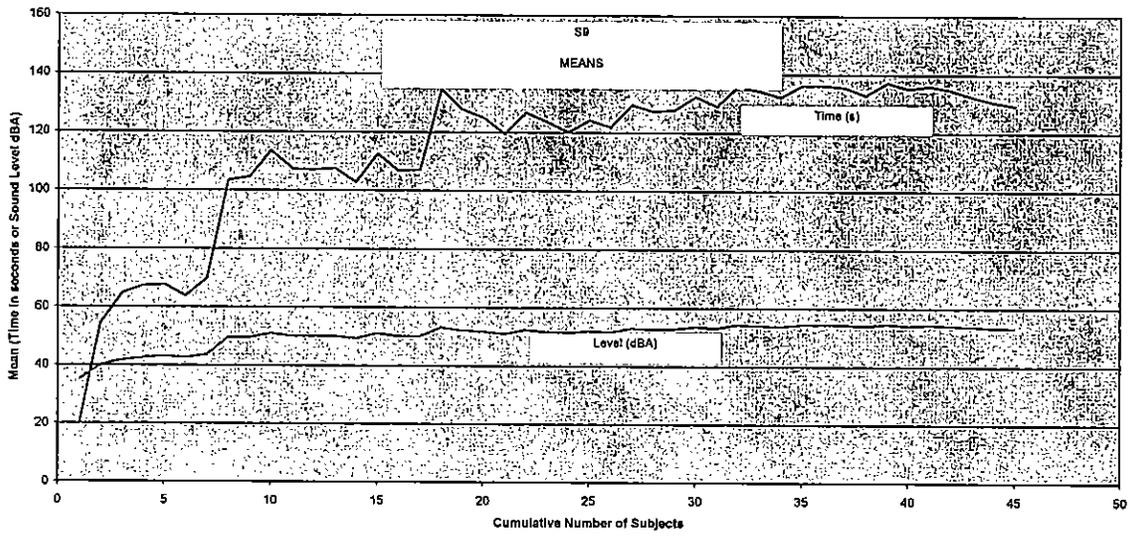
Figure M.7: Mean behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the 500 Hz T-3 signal.
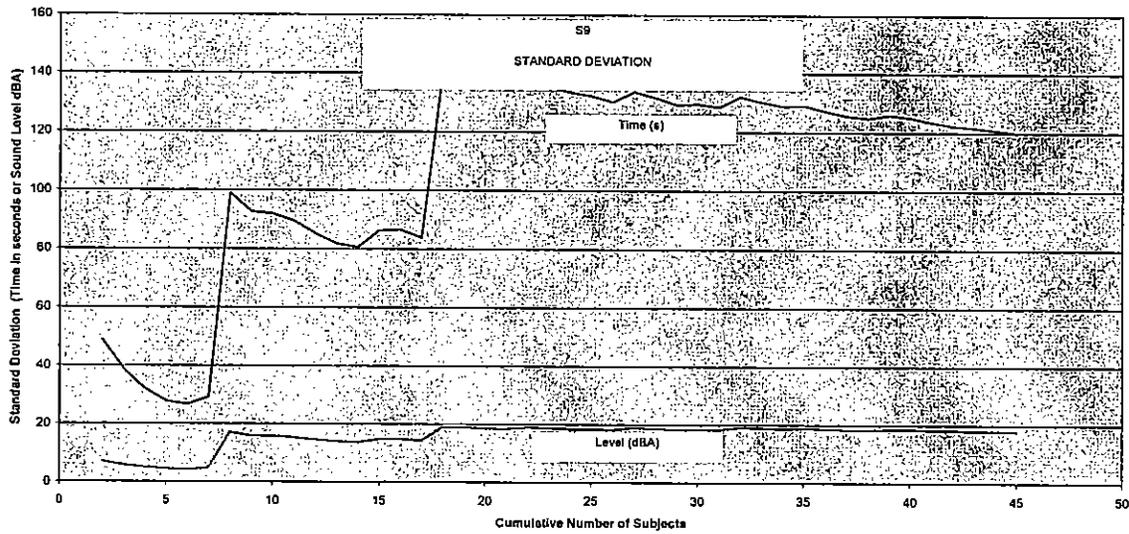

Figure M.8: Standard deviations of behavioural response time in seconds (upper graph) and AAT in decibels (lower graph) as a function of the cumulative number of subjects for the 500 Hz T-3 signal.