Daytime fatigue as a result of reducing nocturnal sleep in order to work longer hours is a common problem in modern societies. Recent research has shown that a nap of 10 minutes in length can increase people's awareness and subsequent performance. However, the problem arises of how to undertake a 10 minute nap without over-sleeping. Behavioural measurements of sleep onset may provide the answer, as they are an inexpensive and convenient way to measure sleep onset. This study examined which out of a passive and active behavioural device was the best measure of sleep onset as defined by polysomnographic sleep onset. The present study used a repeated measures design which involved six participants who were measured using both the active and passive devices, each for three days. On each occasion the participants were measured, they undertook a total of nine sleep onset sessions. Six of the sleep onset sessions involved participants falling asleep whilst using the device, while three sessions involved them falling asleep without the device (control condition). The results revealed no significant difference between the active and passive device both for their discrepancy magnitude from polysomnographic sleep onset and for their discrepancy standard deviations. In addition there were no learning effects for the devices, and the devices did not prolong sleep onset compared to the control condition. The results obtained suggested that both devices were effective measures of sleep onset. However, further research will need to be undertaken to address the limitations of this study and provide a more detailed analysis of the two devices before the full extent of their effectiveness can be determined.
1 Introduction

1.1 Overview

It is well documented that in modern industrialised societies people are often required to reduce their nighttime sleep to permit longer periods in the workplace as a means of increasing their productivity. However, such sleep restriction often decreases people’s daytime level of alertness and work performance. Recent research has shown that a brief nap containing 10 minutes of sleep may be a means of reducing fatigue and increasing performance. A problem occurs, however, in determining how to obtain just 10 minutes of sleep. Simply setting an alarm for 10 minutes would not suffice as people cannot accurately gauge how long it takes them to fall asleep. Thus they could not determine exactly how much sleep they have had. To overcome this difficulty, a device that can measure sleep onset and wake the person up after 10 minutes sleep, is required. Behavioural measures of sleep onset have offered the most potential for providing such a device.

Behavioural measures of the time taken to fall asleep, or sleep onset latency (SOL), have been the focus of considerable research over the past 30 years. Behavioural measures may provide a reliable, inexpensive and subject operable alternative to polysomnography (PSG) sleep measures, which despite being the ‘gold standard’ measure is considered to be an expensive, time consuming and highly labour intensive measure of sleep. There have been two major types of behavioural devices developed, namely, active measures (requires subject to make a response to an external cue) and passive measures (assessing physiological changes). Research has shown each type of behavioural measure to be a moderately accurate measure of SOL. The present study compared an active and passive device and evaluated which provides the best behavioural measure of SOL as determined by PSG SOL. The
results obtained will be employed to determine which device would be the most efficacious when used as a napping measure device.

1.2 The sleep onset process

Sleep onset is considered to be a gradual process during which behavioural and physiological changes take place. For many years researchers have used the physiological changes that occur during the transition from wakefulness to sleep to define a specific point of sleep onset. Rechtschaffen and Kales (1977) developed the most commonly used sleep manual for the determination of the different sleep stages and sleep onset. Sleep onset is considered to have occurred when a person reaches stage 1 sleep. Rechtschaffen and Kales (1977) physiologically defined stage 1 sleep as either a rapid drop out of alpha or when it declines to less than 50 percent of alpha found in wakefulness. Stage 1 sleep is characterised by low voltage, mixed frequency waves, theta (2.0-7.0 Hz) and beta (20-30 Hz). As people go deeper into sleep they reach what is called stage 2 sleep. Stage 2 sleep is physiologically characterised and defined by the presence of sleep spindles (lasting approximately 0.5 seconds with a frequency of 12-14 Hz) and K complexes (a clear large amplitude negative component followed by a positive deflection).

1.3 Background of sleep measurement

Objective sleep measurement originated with the development of the electroencephalogram (EEG) following the observations by Loomis, Harvey and Hobart (1936) that sleep and wakefulness contained different electrical brain activity. Subsequently, further objective measures of sleep were discovered such as rapid eye movement (REM), respiration, and leg movements. Since this time Polysomnography
(PSG) has been well established as the ‘gold standard’ in terms of objective sleep measurement and analysis (Reite, Buysse, Reynolds & Mendelson, 1995). While PSG is considered the best objective measure of sleep, it has limitations. One significant limitation of PSG concerns the need for technical expertise to operate a large amount of complex and expensive equipment. In response to this limitation, researchers have attempted to design instruments capable of recording sleep with the accuracy of the PSG while being both less complex and expensive.

To date, extensive studies have shown that, behavioural measurements of sleep have offered the most promise in terms of the development of such an instrument. Behavioural measures have fewer logistical and financial difficulties compared to PSG measures (Hauri, 1999).

There have been two main types of behavioural measures utilised in sleep research. These are commonly termed ‘active’ and ‘passive’ behavioural measures. The active measures usually involve subjects pressing a microswitch in response to an external cue that may be auditory, tactile or visual. A response to the cue indicates that the person is awake, whereas non-responding indicates that the subject is asleep. In contrast, the passive measures assess physiological changes that occur at sleep onset. These include muscle relaxation or change in respiration with specific changes in these physiological measures indicating sleep (Hauri, 1999).

1.4 Passive behavioural measures

A widely used passive measure of sleep, designed to overcome some of the limitations of PSG sleep recording, is the actigraph. Actigraphs are placed on the wrist to measure the movement or activity of the individual concerned. Modern actigraphs work by recording movement several times per second and storing this
activity data. This data can then be analysed to indicate sleep/wake parameters (e.g., percentage of time asleep, percentage of time awake, number of awakenings, total time awake, total time asleep and SOL) (Ancoli-Israel, Cole, Alessi, Chambers, Moorcroft & Pollak, 2003). Actigraphy offers some significant advantages over PSG, e.g., it is much more convenient to use, it does not require participants to come into a laboratory, is more cost effective, requires little intervention from sleep technicians and can record information for days, weeks and even months (Ancoli-Israel et al., 2003).

Many studies have compared actigraphy to PSG and reported that the two measures are quite highly correlated when measuring sleep and wakefulness. A study undertaken by Cole, Kripke, Gruen, Mullaney and Gillin (1992) showed that the actigraphy monitors were able to correctly differentiate sleep from wakefulness approximately 88% of the time. Jean-Louis, Kripke, Mason, Elliot and Youngstedt (2001) found that correlations between PSG and actigraphy for sleep duration ranged between 0.79 and 0.94. Results of a study by Jean-Louis, von Gizycki, Ferdinand, Fookson, Spielman, Nunes, Fillilove and Taub (1996) indicated a correlation of 0.97 for actigraphy and PSG in relation to total sleep time. Similarly, Sadeh, Sharkey and Carskadon (1994) reported overall agreement rates for PSG and actigraphy to range between 91-93%.

As described above, much of the research employing actigraphy has indicated it to be a suitable instrument for measuring total sleep time and consequently it is now a widely used objective measure of sleep/wake patterns for both research and clinical purposes. Nevertheless, in spite of these advantages, a major limitation of actigraphy is the fact that it is not an accurate measure of SOL.
Previous research has shown that as inactivity is unavoidably required before sleep onset occurs, actigraphic sleep onset criteria is consistently established far before the onset of PSG stage one sleep criteria (Cole & Kripke, 1989; Cole et al., 1992; Hauri & Wisbey, 1992; Mullaney Kripke & Messin, 1980; Webster, Kripke, Messin, Mullaney & Wyborney, 1982). For example a study by Hauri (1999) which included actigraphy in its latency measures reported a correlation of .60 between actigraphy and PSG for stage 1 sleep. The means reported in the study showed that actigraphy latency stage one sleep was 5.5 minutes (SD=6.2) in contrast to the PSG latency which was 17.3 minutes (SD=13.0). These findings clearly reveal the limitations of actigraphs in giving accurate measures of sleep onset due to the fact that a considerable period of quiescence inevitably precedes PSG sleep onset. On average this period is about 12 minutes, but it is quite variable between subjects. Therefore, the onset of quiescence is not an accurate prediction of sleep onset (Hauri, 1999).

Although actigraphy is not an accurate measure of SOL, there have been other passive behavioural measures of sleep that have been more accurate in the measurement of sleep onset. These measures have focused on the decreasing muscle tonus associated with sleep onset.

One of the earliest studies investigating behavioural sleep measures was conducted by Blake, Gerard and Kleitman (1939). Their study employed eight normal subjects to investigate the factors influencing brain potentials. One of the measures Blake et al. (1939) used was muscle tonus. This was measured by using a light spool which was held between two of the subjects fingers as they were falling asleep. Results showed that after the subjects entered stage 1 sleep the muscle tonus relaxed and the subjects dropped the spool within 0.5 to 25 seconds of their alpha rhythm.
diminishing. This showed that certain behavioural changes coincided with stage 1
sleep.

Many of the more recent passive measures of sleep have also focused on
decreasing muscle tonus during stage one sleep. With advances in technology the
measurement devices have become both more sophisticated and more accurate.

Given that self-reports have been criticised as an unreliable measure of sleep
onset, Franklin (1981) designed a device he believed would be able to measure sleep
onset more accurately than self reports. The device consisted of a spring loaded hand
held switch which was attached to an electric clock. When subjects went to bed they
depressed the switch which started the clock. When they entered stage 1 sleep, muscle
relaxation would occur and they would release the switch stopping the clock. Franklin
(1981) took measures of SOL by self report as well as the switch activated clock. The
results showed that the SOL’s gained from the patient’s self reports were significantly
longer than those obtained from the switch activated clock. Franklin (1981) concluded
that although there was no ‘ideal’ measure of sleep onset latency, the ease by which
the switch activated clock could be used together with its cost effectiveness, allowed
it to be a useful tool for both research and clinical applications.

Veins, De Koninck, Van den Bergen, Audet and Christ (1988) designed an
improved SOL monitor to address the main limitation of Franklin’s device, which was
that the clock stopped recording with any brief deactivation of the switch. Hence, any
accidental or short release could result in inaccurate SOL recordings. Veins et al.
(1988) device consisted of a display module and a time-based counter. Similar to
Franklin’s device, the subject depressed the button on a hand held device, with their
thumb. However, this caused the display to go blank thus resulting in the subject not
being able to read the time. In addition, if the subject released the button within a
predetermined interval (in their case, 5 mins) this was ignored by the device. If the release surpassed this period, the clock was reactivated and displayed the SOL in a coded format. Therefore, the SOL was measured from the time the button was depressed to the time it was released, if this time is greater than five minutes. The results of the study showed good correlation for the device and stage 2 sleep onset. Veins et al. (1988) interpreted their results as demonstrating their sleep onset latency device to be a reliable measure of SOL as defined by stage 2 sleep onset. However, their device was not a very good measure of SOL as defined by stage 1 sleep.

A more recent study by Hauri (1999) evaluated the use of a sleep switch device in both good sleepers and insomniacs. Its performance was compared against PSG, wrist actigraphy and subjective self report measures of sleep. The device used in Hauri’s study was a small hand held instrument held between the thumb and finger. Similar to earlier studies, the switch was designed to have enough resistance to release as the subject fell asleep. One side of the device has two metal buttons where the finger rested. When the finger touched these buttons they utilised the subjects galvanic skin response to close a circuit. On the other side of the device is the location of the switch, which is depressed by the thumb. Depression of this switch completes the closing of the circuit and starts the ‘awake’ recording. The clock continues to record awake time until either the button is released or the finger falls out of contact with the electrodes. The results of the Hauri’s study showed that the sleep switch device had a correlation of 0.98 with PSG with regards to measures of the onset of solid sleep, which was defined as 10 or more minutes of sleep without any epoch being scored as wakefulness. However, the devices correlation with PSG stage 1 sleep onset was considerably weaker ($r = 0.60$). Furthermore, the PSG latency to stage 1 sleep was 17.3 minutes ($SD=13.0$), compared to the sleep switch device latency of
32.4 minutes (SD= 30.7). Thus, although this device had a high correlation with PSG latency to solid sleep, it clearly was not as accurate a measure of stage 1 SOL. i.e., there was a 15 minute difference between the sleep switch and PSG measures of stage 1 SOL. As can be observed from the above studies passive behavioural measures have not been a very good accurate measure of SOL with regards to stage 1 and sometimes 2 sleep. However, active behavioural measures of SOL have shown more promise in this regard.

1.5 Active behavioural measures

The second major type of behavioural measure used in sleep research are described as active behavioural measures. Mair (1994) reported that the majority of active behavioural measurements systems used either a micro computer or a portable reaction time unit. These units produce a quiet 5 to 8 second duration tone, using either headphones or a speaker near the bed. The tone is generated at random intervals as subjects approach sleep onset. The subjects are told to press the micro switch in their hand upon hearing the tone. A certain minimum interval is left between tones to permit the subject to de-arouse and allow the possibility of sleep before the presentation of the next tone. Behavioural sleep onset was deemed to have occurred if the subject failed to respond to the tone between two and seven consecutive presentations. Active behavioural measures involving auditory stimuli rely on the assumption that during sleep onset, people’s auditory thresholds rise dramatically, that is, as they fall into sleep their ability to hear sounds in the surrounding environment decreases. Rechtschaffen, Hauri & Zeitlin (1966) claimed they rise approximately 15dB above waking levels.
Another group of researchers have made a considerable contribution to the study of active behavioural measures of sleep. Ogilvie and Wilkinson (1983) investigated behavioural and physiological indicators of sleep, including, EEG and reaction times. Their device produced quiet tones of varying intertone intervals between 3 to 15 seconds. Subjects were instructed to press the button when they heard the tone. Their results showed a consistent lengthening of reaction times, or delay, as the subject approached sleep onset. These findings were replicated by Ogilvie and Wilkinson (1984) in a study which recorded reaction time and EEG recordings to further examine the process of sleep onset. Subjects were instructed to push a switch as soon as they heard a tone (8 seconds in duration). The subjects were further informed that while it did not matter if they fell asleep, they were to push the switch whenever they heard a tone. Subjects were woken if they fell asleep and were told to continue pressing the switch when they heard the tone. The results of this study revealed strong relationships between tone response and sleep stage. The also results showed that subjects failed to respond to the tone 0.7% of the time when awake, 27.8% in stage 1, and 76.0% in stage 2. Thus as subjects only failed to respond 27.8% of the time in stage 1 sleep showed that their measurement lacked accuracy.

A later study by Ogilvie and Wilkinson (1988) examined the inconsistencies between behavioural and EEG measures of sleep. They used a similar methodology to that described for the above studies, with subjects responding to intermittent tones occurring between 1 and 32 seconds. Their results showed that the mean probability of a response was 0.25 in stage 1 and 0.016 in stage 2, and supported the claims that response to an auditory stimulus decreased as subjects descended through the sleep stages.
The relationship between EEG and behavioural sleep onset was investigated by Mair (1994). Mair suggested that as previous studies (e.g., Ogilvie and Wilkinson, 1984) using active behavioural measurement of sleep had generally used long tones (varying between 5 and 8 seconds) it was possible that alpha return (showing a waking state) occurred when the tone was presented. That is, the tone was waking the subjects. Thus Mair examined whether a short duration tone of 0.2 seconds would have a better correspondence with PSG measures than the commonly used 8 second tone. The methodology of Mair (1994) study was similar to previous studies using auditory behavioural measures of sleep. Specifically, subjects were required to respond to tones using a hand held switch device. The results revealed that the long tone (8 seconds) condition resulted in significantly greater amount of time between EEG and behavioural sleep onset, lower correlations with PSG measures and significantly more alpha returns when the tone was presented. In contrast the short tone condition had a median elapsed time difference between PSG sleep onset of 1.91 minutes (SD=2.97) compared to 2.47 minutes (SD=6.73) for the long tone condition. These results demonstrated that shorter tones provide a more accurate behavioural sleep onset measure which was closer to PSG sleep onset, being less variable, while demonstrating that information processing still occurs a short time into PSG sleep onset.

1.6 Review of behavioural measures of sleep onset

The research reviewed above showed that behavioural measures of sleep onset have mixed success in measuring SOL. The time line below (see Figure 1) shows the most successful active and passive devices which met the SOL criteria in relation to PSG.
As can be seen from the timeline, actigraphy is the first behavioural measure to meet sleep onset criteria. This occurs quite a while before PSG sleep onset and is the reason why actigraphy is not a good measure of SOL. The active and passive devices are more accurate measures of SOL however their SOL criteria are usually met after PSG SOL is met.

1.7 *Why is it important to produce accurate behavioural measures of sleep onset?*

A number of recent studies have shown that brief afternoon naps have recuperative effects on people's alertness and cognitive functioning. Tietzel and Lack (2001) looked at the short term benefits of both brief and long naps in people who were sleep restricted (e.g. 5 hours of nocturnal sleep). Their results showed that the detrimental effects of acute sleep restriction were most effectively and most rapidly overcome (in the first hour at least) by a brief 10 minute afternoon nap.

A subsequent study by Tietzel and Lack (2002) focussed on the restorative value of ultra-brief (e.g. 30 and 90 seconds) and brief naps (10 minutes) on alertness and cognitive functioning after nocturnal sleep deprivation. They measured subjective alertness, objective alertness, fatigue, vigour and cognitive performance both before the nap and at several stages post nap. Consistent with their earlier findings, the 10 minute nap provided significantly greater cognitive performance and improved
alertness than the ultra-brief nap. Interestingly, they reported that the ultra brief napping conditions (30 and 90 seconds) provided no significant benefits in comparison to the no nap condition.

A recent study by Brooks (2004) supported the above findings. In this study 5, 10, 20 and 30 minutes naps were compared. Their results once again showed that the 10 minute nap provided instant, reliable and continued improvements (some lasting up to 155 minutes) in all variables measured. In comparison the 20 minute nap showed improvements from 35 minutes after napping to around 125 minutes after napping. However, directly after the 30 minute nap the subjects showed reduced subjective alertness, performance and increased fatigue. This gradually improved until 155 minutes post nap. Therefore, naps of 30 minute were seen as having carry over sleep inertia effects that had subsequent detrimental effects on performance.

The above findings were also supported by Gillberg, Kecklund, Axelsson and Akerstedt (1996) and Horne and Reyner (1996) who have shown that brief naps of 19.8 minutes and 10.8 minutes, respectively, have improved performance in sleep restricted subjects. Thus, brief naps have been shown in the above studies to be effective in enhancing alertness and performance following sleep restriction which appears to be commonplace in modern-day western societies.

1.8 Rational for the present study

There are numerous areas in which napping might be used to address the cognitive and behavioural deficits associated with sleep restriction and thus provide beneficial effects on performance. A means of overcoming these deficits would appear to have significant relevance in modern industrial societies where production and efficiency are seen as key indicators of success. In order to achieve success,
domestic or social activities as well as working long hours. In such situations, it is probable that a brief afternoon nap could be used as an alertness management strategy to reduce fatigue and improve performance. Another example of the benefits of napping on performance concerns fatigued drivers of motor vehicles. Rather than continue driving and risk serious injury to themselves, others or considerably delay their schedule for a long sleep, fatigued drivers might have a brief nap and thereby gain the benefits described above.

However, while napping could be beneficial in improving performance in a variety of settings, the problem arises concerning how an individual can undertake a brief nap outside of environments such as sleep laboratories, and ensure a nap of a given amount (e.g., 10 minutes) of sleep.

In a non-research environment, a brief 10 minute sleep may be difficult to obtain simply because people cannot set their alarm for ten minutes sleep time given that they do not know how long it will take them to fall asleep. In addition, there remains the risk that some individuals who decide to have a 10 minute nap may not wake after ten minutes without an alarm clock. The detrimental effects on performance of napping for 30 minutes compared to ten minutes have previously been highlighted (Brooks, 2004). A possible solution to this problem might be a behavioural sleep onset device that would both measure the onset of sleep, and set an alarm for ten minutes once the person had fallen asleep. Such a device would need to be small, portable, subject operable and affordable.

As discussed above, there has been a large volume of research utilising behavioural measures of sleep onset. Not all of these measures described however, would be applicable in the proposed study. Given that it is not a good measure of
sleep onset, actigraphy would not be suitable for such a napping device. The most preferred devices in terms of accuracy of sleep onset, affordability, convenience and self operability are Hauri's (1999) passive hand held device, and the active behavioural measurement device developed by Mair (1994).

Therefore, the aim of this study is to compare two behavioural devices (one active and the other passive) to see which is best measure of SOL in regards to PSG. Previous research has not compared both an active and passive behavioural device in the same study. Consequently, analysing the results of previous studies to determine which type of behavioural device (active or passive) is the best measure of SOL, can be difficult due to the influence of many other external and internal variables (e.g. differing labs, individual differences). However, by observing two different behavioural devices in one sample in the one study, these variables can be controlled, resulting in a study design that permits analysis of less confounded data. Furthermore, having participants undertake both conditions will allow observations to be made as to whether the devices recorded behavioural sleep onset the same for each participant (i.e., does the active condition occur before the passive condition for everybody) or whether people differ in their behavioural responses (i.e., for some people active behavioural onset occurs before passive behavioural onset and vice versa).
1.9 Experimental Hypotheses

Based on the above rational this study will test the following hypotheses

1) The mean discrepancy magnitudes for the active behavioural condition will be significantly less than the mean discrepancy magnitudes for the passive behavioural condition.

2) The mean discrepancy standard deviations for the active condition will be significantly less than the mean discrepancy standard deviations for the passive behavioural condition.

In addition to these hypotheses additional analysis will be undertaken

- **Testing for learning effects** - in order to explore whether participants' behavioural responses (both active and passive) change over the sessions.

- **Testing for differences in SOL between the active, passive and control condition** - to determine whether the active and passive behavioural devices prolong SOL.

- **Testing for shortening of SOL** - to determine whether a familiarisation with the environment and using the devices results in participants falling asleep sooner in the later sessions than they do in the earlier sessions.
2 Method

2.1 Participants

Six paid participants (3 male, 3 female $M_{age} = 21.00, SD = 1.09$) were recruited using advertising notices placed around Flinders University and other locations (see Appendix G). Each participant received a monetary payment of $200 for their voluntary participation. A small number of participants were employed because this study focused mainly on exploring within-subjects differences as opposed to between-subject differences. Thus each participant was measured on repeated occasions. Subjects were screened and assessed as good sleepers (i.e., estimated SOL < 20 minutes, Total Sleep Time 7-8 hours) in good health by using the Sleep History Questionnaire (see Appendix A) and 7-Day Sleep Wake Diary (see Appendix B). Participants were also screened for extreme phase types (i.e., people who are extreme morning or evening types) using the Time of Day Preference Questionnaire (Horne & Ostberg, 1976, see Appendix C).

Participants were also screened, using a cap test, to ensure they had an adequate level of alpha wave activity (8-12 Hz). Since alpha wave activity is a critical part of the sleep onset criterion a certain level of alpha activity ($10 \mu \text{V}^2, \pm 2$) is required in order for SOL to be established confidently. There is no conclusive evidence in the literature of alpha wave activity has any relationship to participants behavioural responses or SOL’s, consequently excluding participants with low alpha activity will bias the results.

This study was approved by the Flinders University Social and Behavioural Research Ethics Committee. Prior to undertaking the study participants were provided
with a Letter of Introduction (see Appendix D), a Study Information Sheet (see Appendix E) and were asked to sign a Consent Form (see Appendix F).

2.2 Design

The present study used a repeated measures design with three conditions (active device, passive device, and PSG SOL without any device). The participants' sleep was recorded on six different afternoon sessions. For three of the sessions the active measure was used to record SOL along with PSG recording, and for the other three sessions the passive measure was used to measure SOL along with PSG measures. During each session the participant’s SOL was measured without a behavioural device on the first, middle and final trial, to observe whether using either of the devices results in a lengthening of sleep onsets. Thus, each afternoon session consisted of 9 trials, 3 trials measuring PSG SOL without any device and 6 trials, using either the active or passive device. To balance any order effects, subjects undertook the two conditions alternatively, with 3 of the subjects beginning with the passive condition (e.g., ABABAB) and 3 of the subjects beginning with the active condition (BABABA) (see Table 1).
Table I.

**Laboratory Testing Schedule**

<table>
<thead>
<tr>
<th>Estimated time for each testing stage</th>
<th>Active Condition Sessions</th>
<th>Passive Condition Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:50 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITHOUT</strong> device</td>
<td><strong>WITHOUT</strong> device</td>
</tr>
<tr>
<td>1:15 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITH</strong> active device</td>
<td><strong>WITH</strong> Passive device</td>
</tr>
<tr>
<td>1:40 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITH</strong> active device</td>
<td><strong>WITH</strong> Passive device</td>
</tr>
<tr>
<td>2:05 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITH</strong> active device</td>
<td><strong>WITH</strong> Passive device</td>
</tr>
<tr>
<td>2:30 pm</td>
<td><strong>WITHOUT</strong> device</td>
<td><strong>WITHOUT</strong> device</td>
</tr>
<tr>
<td>2:55 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITH</strong> active device</td>
<td><strong>WITH</strong> Passive device</td>
</tr>
<tr>
<td>3:20 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITH</strong> active device</td>
<td><strong>WITH</strong> Passive device</td>
</tr>
<tr>
<td>3:45 pm</td>
<td>PSG SOL</td>
<td>PSG SOL</td>
</tr>
<tr>
<td></td>
<td><strong>WITH</strong> active device</td>
<td><strong>WITH</strong> Passive device</td>
</tr>
<tr>
<td>4:10 pm</td>
<td><strong>WITHOUT</strong> device</td>
<td><strong>WITHOUT</strong> device</td>
</tr>
</tbody>
</table>

Note: each participant undertakes each condition three times, each interchangeably (e.g. A,B,A,B,A,B).

### 2.3 Materials

#### 2.3.1 Sleep questionnaires

Three questionnaires were used in order to screen for good sleepers.

#### 2.3.1.1 Sleep History Questionnaire.

This questionnaire has been extensively used by the Flinders University Sleep Laboratory to assess sleep patterns and general health related to sleep (see Appendix A). It contains a total of 20 questions which assess whether participants are good sleepers and other information about their general sleep patterns.
2.3.1.2 Time Of Day Preference Questionnaire.

This is a 19-item questionnaire (adapted from the ‘Morningness-Eveningness’ questionnaire designed by Horne & Ostberg, 1976) which is used to assess an individual’s time of day activity preference, subjective alertness and effectiveness functioning. In this study it was used to screen for extreme phase types, that is people who awake very early or very late in the morning (see Appendix C).

2.3.1.3 7-Day Sleep/Wake Diary

This is a widely used subjective tool which assesses participants’ typical sleep patterns to ensure that they meet the study’s set criteria (see Appendix B). In this study it was used to screen for good sleepers. The diary is competed over the period of a week. Each day participants are required to fill out certain information such as when they fell asleep, when they woke up, when they ate/drank caffeine or alcohol, their time taken to fall asleep etc.

2.3.2 The Sleep Laboratory

The bedrooms used in this study had no windows, time cues (e.g., clocks) and were sound attenuated to avoid any external interference to sleep onset. The room temperature was kept constant 22 degrees Celsius.

2.3.3 Electrode application

EEG and EOG data were recorded to measure PSG SOL. EEG data was obtained through an electrode cap (Electro-cap International, Eaton, OH, USA) which used standard bipolar recording from Cz to Oz, with a reference electrode positioned in the front of the cap which sat in the middle of the subjects forehead. EOG data was
obtained through an electrode which was positioned on the right cheek of participant's 1cm below and 1cm lateral the right outer eye canthus, whilst the opposing electrode was located on the nasion. EOG electrodes were nicolet 10mm gold-platted electrode cups.

Preceding the attachment of the EEG and EOG electrodes the appropriate skin areas were cleaned using Lemonprep (Mavidon Medical Products, Lake Worth, FL, USA), which was applied with a piece of cotton wool that was soaked in 70% ethanol solution. Electro-gel (Electro-gel; Electro-Cap International, Eaton, OH, USA) was then placed into the cups of the EOG electrodes using a blunted special precision guide needle. The EOG electrodes were fixed to the cleansed facial sites using Alupore (Hypafix; Smith & Nephew, France, a low allergy microporous adhesive tape. Once the electrode cap was fitted a blunted special precision guide needle was used to fill the electrodes on the cap with conductive electrode gel. Because impedances were required to be below 5000 ohms, for clear reading, the blunted needle was moved gently moved around to abrade the skin underneath the electrode.

Impedances were measured at the beginning (and if required during) of each session using a digital multimeter (Q-1419 Dick Smith Electronics, Sydney, NSW).

2.3.4 LabVIEW5 computer program

A modified version of Laboratory Virtual Instrument Engineering Workbench 5 (LabVIEW5; National Instruments Corporation, Texas, U.S.A) computer programme was run using a Power Macintosh 7300/200 computer which received signals via an opto-isolated amplifier (Flinders University of South Australia). During each trial LabVIEW5 displayed and recorded a power spectral analysis graph of EEG data for alpha (8 - 12 Hz), theta (4 - 7 Hz) and delta (0.5 - 4Hz) brain waves. This
The power spectral analysis graph was used to determine sleep onset and was positioned in the middle of the computer monitor. The Y-axis of the graph showed the relevant power of each waveform for each epoch and the X-axis of the graph showed the number of 30-second epochs there had been since the beginning of the trial. Raw EOG data was graphically displayed on the monitor above the power spectral analysis graph. This showed 10-second phases of time (0-10 seconds, 10-20 seconds, 20-30 seconds) of EOG activity during each epoch followed by the full retrospective 30-second period in a condensed format.

2.3.5 Behavioural measurements of sleep.

2.3.5.1 Active behavioural measure.

A purpose built Programmable Stimulus Tone Generator (PSY 381, Flinders University of South Australia) was designed and constructed for this research. It was configured to deliver 1000Hz tones of 200 millisecond duration at a random intervals of 16 to 41 seconds. The tone generator was controlled internally by a Peripheral Interface Controller (PIC). This also generated a digital pulse when a tone had occurred that was sensed by the input/output card and viewed on the computer.

The ambient noise level of each bedroom was measured using a Bruei And Kerr Type 2231 Precision Sound Level Meter (PSLM). The tone level was then set at 6dB above this level using the PSLM. The tone was set at 6dB as it was thought the tones used in previous studies e.g., Mair (1994) i.e., 15dB above ambient noise level, were too loud and as a result, may have been arousing the participants. In fact 15dB actually believed to be of alarm level (Sanders & McCormack, 1993). Bruei and Kerr (internet reference) leaders in sound measurement stated that a tone of 3dB is barely noticeable by most people (should be easily heard by the participants in this study due
to their young age) and that a tone of 6dB would result in a doubling of that sound pressure. Thus in this study, a 6dB tone was utilised as it considered to be less arousing to participants but still loud enough to be easily responded to when awake.

A hand-held device was constructed using a Linear Force Distance Transducer (RS Electronics Code 317-780) and a DC Line Level monitor (PSY384, Flinders University of South Australia) was used to generate a digital pulse when the plunger was depressed 3mm. This pulse was fed to the input/output card so that the LabView5 software could record the time activated in conjunction with tones generated by the speaker. Reaction times were then computed and if the response time was within 10 seconds from the onset of the tone then a green indicator was displayed on the monitor to the right of the power spectral analysis graph. If a participant did not respond to the tone within 10 seconds the indicator would turn red. The monitor also displayed the number of responses for each epoch in a graph directly below the power spectral analysis graph. The onset and response times were recorded to file for further analysis.

2.3.5.2 Passive behavioural measure.

The same hand-held device was used to measure the passive behavioural response when undertaking the passive condition. The monitor displayed the percentage the device was released for each epoch and the number of releases each epoch in a graph directly below the power spectral analysis graph. The force required to depress this hand-held device was greater (nominally 350gms plus/minus 50gms) than that used in Hauri (1999) (nominally 65gms plus/minus 10gms) This was the case because it was suspected the time subjects in Hauri’s (1999) study were able to hold down the switch because the required force was too low. Therefore the device in
this study is designed to provide enough resistance to avoid being held down too easily but not too much that it became a strain or arousing for the participant to depress the switch.

2.4 Procedure

2.4.1 Prior to the experiment

Prior to the experiment subjects completed the Sleep History Questionnaire (see Appendix A and Time of Day Preference Scale (see Appendix C) to assess whether they are suitable candidates for the study. If they met the requirement of the study they then underwent an EEG cap test to establish whether they have enough alpha to be a clear EEG sleep onset indication. During the cap tests a cap was applied to the participants and their alpha waves were observed. Suitable subjects were then scheduled times for their 6 afternoon sessions.

2.4.2 Preparation for the experiment

The following procedures are consistent with Brooks (2004).

On the night prior to each laboratory session participants restricted their sleep from 2am to 7am. They confirmed these times by ringing the sleep lab and leaving a message on the voicemail. The messages were then accessed the next day and the time in which they were received recorded. On the day of their laboratory sessions, participants were required to come into the sleep lab at 12.15 pm where they were fitted with an electrode cap and EOG electrodes. Following this, the subjects were directed to the bedroom and were allowed 10 – 15 minutes to settle and get comfortable in the bed. At 1.50pm the participants undertook the first condition.
2.4.3 PSG sleep onset determination

Rechtschaffen and Kales (1977) developed the most widely used sleep onset criteria. They define sleep onset as a decrease in alpha waves to half that present during relaxed wakefulness. Subsequently, in this study stage 1 sleep onset latency was said to occur when the power spectral analysis on the Macintosh computer showed a 50 percent drop in baseline alpha power, with this reduction remaining for three successive epochs.

2.4.4 During the experiment

2.4.4.1 Control trials

In the control trials participants were simply told that the light was to be turned off and that they were to lie down, relax and try to fall asleep.

2.4.4.2 Active trials

In the active trials participants were told to hold the hand-held device in the palm of their preferred hand with their thumb resting on the button. They were then told that the light was going to be turned off and that they were to lie down, relax, try to fall asleep and to push the button down if they heard the tone.

2.4.4.3 Passive trials

In the passive trials participants were told to hold the hand-held device in the palm of their preferred hand and to depress the button. They were then told that the light was going to be turned off and that they were too lie down, relax, try to fall asleep and to continue to depress the button. They were informed that if they felt they
had released the button for any reason that they were simply to depress it again and maintain the button in the depressed position.

2.4.4.4 Protocol during the trials

Participants were provided with a 25-minute opportunity to fall asleep for each trial. In the control condition when subjects met Stage 1 sleep onset criteria (three consecutive epochs of stage 1 or another stage of sleep) the participants were immediately woken up. In the active and passive conditions participants were woken when both Stage 1 sleep and behavioural sleep onset criteria were reached. If it occurred that behavioural sleep onset criteria had not been reached five minutes after PSG SOL had occurred the participants were woken up. This five-minute protocol was specifically designed to give the participants enough time to reach behavioural sleep onset after PSG SOL had occurred while not allowing them too much sleep that could affect their sleep latency on the next trial (Brooks, 2004). After subjects were woken they had on average between 5 and 10 minutes of quiet awake time until the next trial (depending on how long it took them to fall asleep). Each trial lasted for a maximum of 25 minutes allowing 9 trials over a period of approximately 3 hours.

2.4.4.5 Number of times behavioural sleep onset criteria had not been met before five-minute cut off

On only four occasions did participants have to be woken up five-minutes after PSG SOL had occurred and behavioural onset latency criteria had not been met. As these scores could not be left out (this would have excluded the worst scores) the behavioural onset latency was said to occur 5 minutes after PSG SOL.
2.4.5 Measurement of behavioural sleep onset for the active and passive devices

In the active condition, as in Mair (1994) the tones had an intertone interval varying randomly between 16 to 41 seconds. Subjects were considered to have reached behavioural sleep onset with a failure to respond to two consecutive tones. The sleep onset time was defined as the time at which the first of these two consecutive failures to respond occurred. In the passive condition behavioural sleep onset was defined as the subject failing to depress the button for an entire 30-second epoch.
3 Results

3.1 Missing data

There were three sessions during the experiment where the subjects did not fall asleep thus resulting in missing data. Due to there only being three sessions where this happened the results were simply analysed without this data.

3.2 Hypotheses testing

3.2.1 Evaluation of how each participant performed during the experiment

Before presenting the results of hypothesis one which examines whether there was a significant difference in the mean discrepancy score between the active and passive conditions for the whole group, the data for each individual will be presented in order to provide a clearer and more detailed picture of the results.

3.2.1.1 The calculation of discrepancy scores

It was decided, for this purpose of this study, that analysis of the behavioural onset latencies (BOL's) and the polysomnographic sleep onset latencies (PSG SOL) would be undertaken in terms of discrepancy scores. For example, if a participant recorded a PSG SOL of 12 minutes and an BOL of 15 minutes, when using the active device, the discrepancy score would be 3. Similarly, if a participant recorded a PSG SOL of 12 minutes and a BOL of 9 minutes the discrepancy score would be –3. This method of scoring provides a clear indication of where the participants average active and passive BOL's are in relation to PSG SOL. Thus, a discrepancy score of 0 is the optimum result as it shows that there was no discrepancy between the BOL and the PSG SOL.
3.2.1.2 Reading the scatterplot for each subject

The scatterplots below show the participant’s PSG SOL for each trial (excluding control trials) and the corresponding BOL for that trial. Points above the midline mean that BOL had occurred after PSG SOL, and points below the midline mean that BOL occurred before PSG SOL. Therefore, a majority of the points for either the active or passive condition occurring above the midline would indicate that the participant had fallen asleep before he/she released the button or ceased to respond to the tone. Similarly, a majority of points for either the active or passive condition occurring below the line would denote that the participant had released the button or stopped responding to the tone before he/she have fallen asleep.
3.2.1.3 Participant One

![Scatterplot indicating participant one's PSG SOL and corresponding BOL for all of their active and passive trials.](image)

**Figure 2.** Scatterplot indicating participant one's PSG SOL and corresponding BOL for all of their active and passive trials.

**Note:**

- Mean BOL ($M=8.05$) / PSG SOL ($M=10.33$) for active condition
- Mean BOL ($M=7.11$) / PSG SOL ($M=9.64$) for passive condition

---

**Table 2.**

<table>
<thead>
<tr>
<th></th>
<th>Active Device</th>
<th>Passive Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Discrepancy</td>
<td>.94</td>
<td>.69</td>
</tr>
<tr>
<td>Standard Deviation (discrepancy)</td>
<td>1.16</td>
<td>1.55</td>
</tr>
<tr>
<td>Correlations between BOL and PSG SOL</td>
<td>.95</td>
<td>.97</td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td>.91</td>
<td>.94</td>
</tr>
</tbody>
</table>
Figure 2 shows that for participant one the active and passive BOL usually occurs very soon after PSG SOL. This was confirmed by the mean discrepancies in Table 2 which showed that on average for the active device, participant one stops responding to the tones .94 minutes after PSG SOL, and .69 minutes with the passive device.

An independent-samples t-test was conducted to compare the mean discrepancy magnitudes for the active and passive condition for participant one. When undertaking this analysis however, the discrepancy scores had all negative numbers (e.g. -1,-3) were converted to positive numbers. This was carried out in order to compare the magnitude of the discrepancies rather than the direction of the discrepancy from the PSG SOL (i.e., before and after). For example, if a participant had an average active BOL of 2.00 minutes and an average passive BOL of -2.00 minutes, a t-test would show these to be significantly different measures of BOL. However, in relation to the PSG SOL of the participant they are equally accurate (both out by 2 minutes). Thus, when determining whether the two devices (active or passive) are significantly different measures of BOL the negative signs will be removed.

The results of the t-test showed that there was no significant difference between the active discrepancy magnitudes \( (M=1.11, SD=.99) \) and the passive discrepancy magnitudes \( (M=1.08, SD=1.28) \), \( t(34)=0.07, p=.94 \). Therefore, for participant one there was no significant difference in discrepancy magnitude between the active and passive devices as a measure of PSG SOL.
3.2.1.4 Participant Two

Figure 3. Scatterplot indicating participant two’s PSG SOL and corresponding BOL for all of their active and passive trials.

Note:  

\[+\] = Mean BOL (M=7.58) / PSG SOL (M=5.78) for active condition 

\[\times\] = Mean BOL (M= 6.36)/ PSG SOL (M=5.47) for passive condition 

--- = midline: points below - BOL occurred before PSG SOL

: points above – BOL occurred before PSG SOL

Table 3.

Mean Discrepancy and Standard Deviation (in minutes), Correlations and Regression

Coefficients for Participant Two.

<table>
<thead>
<tr>
<th></th>
<th>Active Device</th>
<th>Passive Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Discrepancy</td>
<td>1.81</td>
<td>0.89</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.13</td>
<td>0.85</td>
</tr>
<tr>
<td>Correlations between</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td>0.84</td>
<td>0.93</td>
</tr>
</tbody>
</table>
Figure 3 shows that for participant two the active and passive BOL usually occurred soon after PSG SOL. This was confirmed by the mean discrepancies in Table 3 which revealed that on average for the active device, participant two ceases responding to the tones 1.81 minutes after PSG SOL, and .89 minutes with the passive device.

An independent-samples t-test was performed to compare the mean discrepancy magnitudes for the active and passive condition for participant two. As described above, when undertaking this analysis the discrepancy scores had all negative numbers converted to positive numbers. The results of the t-test revealed that there was a significant difference between the active discrepancy magnitudes ($M=1.81$, $SD=1.13$) and the passive discrepancy magnitudes ($M=.89$, $SD=.85$), $t(34)=2.76$, $p<.05$. Therefore, for subject two the average passive BOL occurred closer in time to PSG SOL than active BOL.
3.2.1.5 Participant Three

![Scatterplot indicating participant three’s PSG SOL and corresponding BOL for all of their active and passive trials.](image)

**Figure 4.** Scatterplot indicating participant three’s PSG SOL and corresponding BOL for all of their active and passive trials.

Note: 

+ = Mean BOL (M=6.53) / PSG SOL (M=7.08) for active condition

× = Mean BOL (M=5.81) / PSG SOL (M=7.97) for passive condition

--- = midline: points below - BOL occurred before PSG SOL

: points above - BOL occurred before PSG SOL

**Table 4.**

*Mean Discrepancy and Standard Deviation (in minutes), Correlations and Regression Coefficients for Participant Three.*

<table>
<thead>
<tr>
<th></th>
<th>Active Device</th>
<th>Passive Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Discrepancy</td>
<td>-.56</td>
<td>-.216</td>
</tr>
<tr>
<td>Standard Deviation (discrepancy)</td>
<td>2.20</td>
<td>2.51</td>
</tr>
<tr>
<td>Correlations between BOL and PSG SOL</td>
<td>.59</td>
<td>.67</td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td>.35</td>
<td>.45</td>
</tr>
</tbody>
</table>
Figure 4 illustrates that for participant three the active and passive BOL usually occur before PSG SOL. This was confirmed by the mean discrepancies in Table 4 which showed that on average for the active device, participant three stops responding to the tones .56 minutes before PSG SOL, and 2.16 minutes before PSG SOL with the passive device.

An independent-samples t-test was conducted to compare the mean discrepancy magnitudes for the active and passive condition for participant three. The results of the t-test showed that there was no significant difference between the active discrepancy magnitudes ($M = 1.61, SD = 1.59$) and the passive discrepancy magnitudes ($M = 2.27, SD = 2.40$), $t(34) = -0.98, p = .33$. Therefore, for participant three there was no significant difference between the active and passive devices as a measure of PSG SOL.
3.2.1.6 Participant Four

![Scatterplot indicating participant four's PSG SOL and corresponding BOL for all of their active and passive trials.](image)

**Figure 5.** Scatterplot indicating participant four's PSG SOL and corresponding BOL for all of their active and passive trials.

**Note:**
- + = Mean BOL ($M=10.85$) / PSG SOL ($M=10.12$) for active condition
- $\times$ = Mean BOL ($M=4.94$)/ PSG SOL ($M=5.88$) for passive condition
- Midline: points below – BOL occurred before PSG SOL
- Points above – BOL occurred before PSG SOL

**Table 5.**

**Mean discrepancy and Standard Deviation (in minutes), Correlations and Regression Coefficients for Participant Four.**

<table>
<thead>
<tr>
<th></th>
<th>Active Device</th>
<th>Passive Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Discrepancy</td>
<td>.74</td>
<td>-.94</td>
</tr>
<tr>
<td>Standard Deviation (discrepancy)</td>
<td>2.07</td>
<td>2.45</td>
</tr>
<tr>
<td>Correlations between BOL and PSG SOL</td>
<td>.95</td>
<td>.64</td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td>.90</td>
<td>.41</td>
</tr>
</tbody>
</table>
Figure 5 shows that for participant four the active BOL usually occurred soon after PSG SOL whereas the passive BOL usually occurred just prior to PSG SOL. This was confirmed by the mean discrepancies in Table 5 which show that on average for the active device, participant four stopped responding to the tones .74 minutes after PSG SOL, and released the passive device on average around .94 minutes before PSG SOL.

An independent-samples t-test was performed to compare the mean discrepancy magnitudes for the active and passive condition for participant four. The results of the t-test showed that there was no significant difference between the active discrepancy magnitudes ($M=1.56$, $SD=1.52$) and the passive discrepancy magnitudes ($M=2.12$, $SD=1.47$), $t(32)=-1.08$, $p=.28$. Thus, for subject four there was no significant difference between the active and passive devices as a measure of PSG SOL.
3.2.1.7 Participant Five

![Graph showing BOL (mins) vs PSG SOL (mins) for active and passive conditions.]

Figure 6. Scatterplot indicating participant five’s PSG SOL and corresponding BOL for all of their active and passive trials.

Note: + = Mean BOL (M=9.27) / PSG SOL (M=6.97) for active condition
     \* = Mean BOL (M=10.97) / PSG SOL (M=8.64) for passive condition
     ----- = midline: points below - BOL occurred before PSG SOL
     : points above - BOL occurred before PSG SOL

Table 6.

**Mean Discrepancy and Standard Deviation (in minutes), Correlations and Regression Coefficients for Participant Five.**

<table>
<thead>
<tr>
<th></th>
<th>Active Device</th>
<th>Passive Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Discrepancy</td>
<td>2.29</td>
<td>2.33</td>
</tr>
<tr>
<td>Standard Deviation (discrepancy)</td>
<td>1.77</td>
<td>1.37</td>
</tr>
<tr>
<td>Correlations between BOL and PSG SOL</td>
<td>.95</td>
<td>.98</td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td>.89</td>
<td>.96</td>
</tr>
</tbody>
</table>
Figure 6 reveals that for participant five the active and passive BOL usually occurs after PSG SOL. This was confirmed by the mean discrepancies in Table 6 which show that on average for the active device participant five ceased responding to the tones 2.29 minutes after PSG SOL, and released the passive device on average around 2.33 minutes.

An independent-samples t-test was conducted to compare the mean discrepancy magnitudes for the active and passive condition for participant five. The results of the t-test showed that there was no significant difference between the active discrepancy magnitudes ($M=2.41, SD=1.59$) and the passive discrepancy magnitudes ($M=2.33, SD=1.37$), $t(33)=0.15, p=.87$. Hence for subject five there was no significant difference between the active and passive devices as a measure of PSG SOL.
3.2.1.8 Participants Six

Figure 7. Scatterplot indicating participant six's PSG SOL and corresponding BOL for all of their active and passive trials.

Note: $\uparrow$ = Mean BOL ($M=15.25$) / PSG SOL ($M=14.11$) for active condition

$\times$ = Mean BOL ($M=8.75$) / PSG SOL ($M=12.08$) for passive condition

= midline: points below - BOL occurred before PSG SOL

: points above - BOL occurred before PSG SOL

Table 7.

Mean discrepancy and Standard Deviation (in minutes), Correlations and Regression Coefficients for Participant Six.

<table>
<thead>
<tr>
<th></th>
<th>Active Device</th>
<th>Passive Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Discrepancy</td>
<td>1.34</td>
<td>-3.33</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.05</td>
<td>4.79</td>
</tr>
<tr>
<td>(discrepancy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlations</td>
<td>.98</td>
<td>.59</td>
</tr>
<tr>
<td>between BOL and PSG SOL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression Coefficients</td>
<td>.97</td>
<td>.35</td>
</tr>
</tbody>
</table>
Figure 7 illustrates that for participant six the active BOL usually occurred after PSG SOL but that passive BOL occurs before PSG SOL. This is confirmed by the mean discrepancies in Table 7 which show that on average for the active device participant six stopped responding to the tones 1.34 minutes after PSG SOL, and releases the passive device on average around 3.33 minutes prior to PSG SOL.

An independent-samples t-test was conducted to compare the mean discrepancy magnitudes for the active and passive condition for participant six. The results of the t-test showed that while there was no significant difference between the active discrepancy magnitudes \(M = 1.14, SD = 1.05\) and the passive discrepancy magnitudes \(M = 3.39, SD = 4.75\), \(t(34) = -1.96, p = .058\), it was approaching significance. While the results revealed there were no significant difference between the active and passive devices as a measure of PSG SOL for participant six, they suggested, for this participant, that the active device was a better measure of PSG SOL.

3.2.1.9 Summary of individual cases

Overall, the above results revealed that only participant’s two (better with passive device) and possibly participant six’s (better with active device) showed a significant difference between the two devices, whilst the other four participants showed no significant difference. Although the scatterplot’s and discrepancy means showed that the participants reacted differently using each device, the individual data suggests that any differences observed between the two devices was not statistically significant.
3.2.2 Hypothesis One

A paired-samples t-test was performed to test whether the mean discrepancy magnitude scores for the active behavioural condition were significantly less than the mean discrepancy magnitudes for the passive behavioural condition. Again, using the rationale described earlier, the negatives signs were removed from the discrepancy score. The results showed that there was no statistically significant difference between the active behavioural condition \((M = 1.61, SD = .48)\) and the passive behavioural condition \((M = 2.01, SD = .92)\), \(t(5) = -0.94, p = .39\).

Table 8.

Mean Discrepancy Magnitudes and Standard Deviations for the Active and Passive Conditions.

<table>
<thead>
<tr>
<th></th>
<th>Mean discrepancy magnitude</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>1.61</td>
<td>.48</td>
</tr>
<tr>
<td>Passive</td>
<td>2.01</td>
<td>.92</td>
</tr>
</tbody>
</table>

3.2.3 Hypothesis Two

As the variability of the device (i.e., how consistent it is) is an extremely important factor to consider when evaluating which of the two devices (active, passive) is a better measure of PSG SOL, a paired-samples t-test was conducted to examine whether there was a difference between the standard deviations (of the discrepancy scores) of the active behavioural condition and the passive behavioural condition. The negatives of the discrepancy scores were not removed for this analysis because doing so would detract from the BOL variance recorded in the devices. The results indicated that there was no statistically significant difference between the
active behavioural conditions standard deviation scores ($M=1.53$, $SD=.51$) and the passive behavioural conditions standard deviation scores ($M=2.32$, $SD=1.36$), $t(5)= -1.32$, $p=.24$.

Table 9.

**Standard Deviation Scores and Standard Deviation for the Active and Passive Condition.**

<table>
<thead>
<tr>
<th></th>
<th>Standard deviation scores</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>1.53</td>
<td>.51</td>
</tr>
<tr>
<td>Passive</td>
<td>2.32</td>
<td>1.36</td>
</tr>
</tbody>
</table>

3.2.4 **Testing for learning effects**

In order to determine whether there was a learning effect for the two devices over the three sessions a two-way repeated measures ANOVA was performed using participants average discrepancy scores (with negatives). The results revealed that there was no statistically significant effect for device, $F(1, 5)= 5.06$, $p=.07$.

Additionally there was no main effect for time, $F(2,10)=1.44$, $p=.28$, and there was no interaction between time and device, $F(2,10)=0.45$, $p=.65$. Thus, no learning effect were found over the three session for the two device (see Figure 8)
3.2.5 Testing for differences in PSG SOL

In order to investigate whether there were significant differences between the SOL for the active, passive and control conditions a one-way repeated measures ANOVA was conducted using the participants raw PSG SOL scores. There was no statistically significant difference in PSG SOL between the active (M=8.53, SD=3.09), passive (M=8.28, SD=2.17), and control conditions (M=8.62, SD=2.17), $F(2, 10)=0.10, p=.90$. Therefore participants PSG SOL was the same in the active, passive and control condition.

3.2.6 Testing for shortening of PSG SOL

A two-way repeated measures ANOVA was used to test whether PSG SOL was shorter in the later sessions than in the earlier sessions for the active, passive and control conditions. The results showed that there was no significant main effect for time, $F(2,10)=2.73, p=.11$. Additionally, there was no significant main effect for device, $F(2,10)=0.04, p=.95$, and there was no significant interaction effect $F(4,20)=2.07, p=.12$. These results revealed that PSG SOL remained the same in the three conditions from the initial sessions to the later sessions (see Figure 9).
Figure 9. PSG SOL scores for the active, passive, and control conditions across the sessions.
4. Discussion

4.1 Overview

This study investigated which out of an active and passive behavioural device would be the most accurate at determining sleep onset in regards to PSG SOL. Two main parameters were analysed to determine which was the best measure of SOL, 1) whether the discrepancy magnitudes for the devices as a measure of SOL differed significantly between the two devices and 2) whether the mean discrepancy standard deviations differed between the two devices. The results revealed no significant differences between the active and passive devices both in regards to their discrepancy magnitudes from PSG SOL and their mean discrepancy standard deviations.

The following section discusses the findings of this study and provides a detailed interpretation of the results. It also outlines the limitations of the study, and provides some suggestions on what might be done to improve the study and finally suggests some possible directions for future research.

4.2 Interpretation of Hypothesis One

The results of hypothesis one showed that there was no significant difference between the mean discrepancy magnitudes for the active behavioural condition and the passive behavioural condition. This finding failed to support the hypothesis that the mean discrepancy magnitudes for the active behavioural condition would be significantly lower than those for the passive behavioural condition. The results revealed that, on average, the BOL for the active device occurred within 1.61 minutes of PSG SOL occurring, while the BOL for the passive device occurred within 2.01
minutes of PSG SOL occurring. These results demonstrate that both devices were accurate measures of SOL, particularly in regards to the passive device, than had been reported in previous research. For example, Hauri’s (1999) passive device had a latency to PSG SOL of 17.3 minutes which was much longer than the device developed for this study. This may have been a function of the changes in design concerning the passive behavioural device used in this study, which resulted in the passive device requiring a greater force to depress the button than was the case in Hauri’s (1999) study. Similarly, the active device developed for the present study appeared to be a better measure of BOL in comparison to the device used by Mair’s (1994) study which reported a median elapsed time difference to PSG SOL of 1.91 minutes. Again, an explanation for this finding could be the improvement made for the active device in this study, specifically, the use of quieter tones. Overall while the two devices did not differ significantly from each other as measures of BOL, they were both very good measures of BOL in their own respect and in comparison to devices used in previous research.

### 4.3 Interpretation of Hypothesis Two

The results concerning hypothesis two showed that there was no significant difference between the discrepancy standard deviations for the active behavioural condition and the passive behavioural condition. Again, this failed to support the hypothesis that had predicted the discrepancy standard deviation for the active behavioural condition would be significantly lower than those for the passive behavioural condition. In particular, the active device had a discrepancy variation in BOL of 1.53 minutes and the passive device had a discrepancy variation in BOL of 2.32 minutes. Both these discrepancy variations were quite low, that is, both device’s
BOL only varied within 2.32 minutes of BOL. Thus both devices not only had BOL in close proximity to PSG SOL (as shown in hypothesis one) but also had little variability in their BOL thereby making them a consistent measure of BOL. As discussed previously, the variability of behavioural devices is probably the most important factor to consider when determining their effectiveness. Given that both devices displayed low discrepancy variability they were effective measures of BOL.

4.4 Learning effects

There were no learning effects observed for the two devices across the three days each were used. This indicated that participants responded the same with both devices even after practise using them. This was a noteworthy finding in the present study, in view of the fact an absence of learning effect is crucial to a good behavioural sleep measure. If learning effects had occurred e.g., if participants learnt to hold down the button/ respond to the tones continually longer into sleep, it would be difficult to determine whether the devices would remain accurate devices after prolonged use. However, because no learning effects occurred, it is possible to interpret this result as supporting the assertion that these two devices will remain accurate measures of behavioural sleep onset latency with continued use.

4.5 PSG SOL for the active, passive and control condition

There was no difference found between the active, passive and control condition with regards to PSG SOL. This result revealed that using both the active and passive device did not prolong the time taken to fall asleep. This, is an important finding because if the devices significantly prolonged PSG SOL they would be rendered quite useless as a napping device regardless of their accuracy or consistency.
To illustrate, if a person wanted to take a power nap, however, it took them 30+ minutes longer to fall asleep due to using the device, then clearly the device is not at all practical.

4.6 Testing for a shortening of PSG SOL

There were no differences observed between the PSG SOL in the initial sessions compared with the later sessions. These results suggested that despite familiarisation with the sleep laboratory, the times participants fell asleep in the later sessions were similar to those observed in the initial session. Therefore, it could be concluded that participants were relatively comfortable coming into the laboratory and participating the experiment.

4.7 Summary of effectiveness of devices

The results of this study showed that both the active and passive behavioural devices were effective measures of BOL. Firstly, the results of hypotheses one and two revealed that both the active and passive device were good measures sleep onset. The BOL for both devices occurred in close proximity to PSG SOL, particularly when compared to previous research. Furthermore, both devices had low discrepancy standard deviations which indicated they were relatively consistent measures of BOL.

In addition to this, the results also demonstrated that the devices neither delayed PSG SOL, nor had any learning effects. Thus it can be seen that not only are the devices accurate and consistent measures of BOL, the results also revealed them to be practical (do not prolong SOL) and perform the same with continued use, thereby making them very effective measure of BOL.
4.8 Limitations and future research

Two major limitations of this study were small sample size, only 6 people were able to be tested and the fact that they were only tested six times each (3 times with each device). While this is clearly a limitation of this study, the logistics of recruitment and screening and testing for suitable participants required over 250 hours of time in the laboratory. Thus while a greater sample size may have contributed greater confidence in interpreting the results of the study, it was not possible, given the nature of this particular project, to increase the sample size or frequency of testing. Clearly further research, which overcomes this limitation, is required further to investigate the findings reported in this study.

Firstly, future research would need to explore more closely whether or not prolonged use of the devices results in a learning effect. Although the data in this study suggested that no learning effect was associated with the devices, that is, participant’s responses with the devices was the same in the early sessions as it was in the later sessions, each devices was only tested for three days each. Thus, future studies might include designs where participants are tested over greater number of days. If the devices were to be developed into ‘power nap devices’ they would be used by people tens of times or possibly hundreds of times a year. Hence, future studies would need to measure participants over many days (each day may need fewer trials in it) in order to simulate prolonged use of the devices and subsequently determine whether learning effects occurred after extensive use. If learning effects did occur, i.e., people were able to hold down the button or respond to the tone continually longer after PSG SOL, then the devices would appear to have limited efficacy as a power nap measure.
Secondly, future studies would need to employ more participants to permit results to be generalised to a larger population. Studying more participants would not only provide a more accurate indication of the devices discrepancy magnitude and discrepancy standard deviations, it would also allow researchers to determine if different types of people respond to the devices in a different way. The results found in the present study demonstrated that there were no differences between the devices both for individuals and over the group, with the exception of one participant. However by measuring more participants, future research would be able to tease out any differences between the devices for different people, and if so, investigate possible explanations for these differences.

Thirdly, future research could extend on this study by examining how well these devices work as measure of a power nap. In such studies, participants could be sleep deprived the night before coming into a laboratory and then instructed to take a power nap using one of the devices. This would not only provide further information concerning the accuracy of the devices as a measure of BOL, but would provide data that could be analysed to determine if they were effective in facilitating a 10 minute nap. Participants in future studies might undertake a 10 minute nap as determined by the device and then undergo a series of fatigue and awareness test, as used in Brooks (2004), to assess if the nap they had using the device was beneficial and if so, the degree of benefit derived.

Finally, future research might also provide further information concerning how subject friendly the devices are. If future studies were to show that both devices were accurate measures of BOL, the further research would be required to investigate which device people preferred using and which device would be more practical to use. For example, did subjects prefer using the passive device as it was less intrusive, and
would the passive device be more practical as the whole device could be contained in one unit.

Overall, the current study does have limitations in that it cannot produce all of the information needed to determine categorically how effective these devices are as a measure of BOL. However, it does provide a good foundation for future research to build on which could determine if the findings of this study occur also occur in a larger population.

4.9 Conclusion

The current study has shown that there were no differences between the active and passive devices in terms of their discrepancy magnitudes from PSG SOL and in terms of their discrepancy standard deviations. The subsequent scores of both devices for these measures suggested that they would be effective measures of BOL. Additionally, the effectiveness of the devices was given further support as there were no learning effects found for the devices, and using the device did not prolong PSG SOL in comparisons to the control condition. Thus, overall, this study produced results that supported the conclusions that both devices were effective measures of BOL. The limitations of this study were described and discussed, and although the data suggested that both of these devices were effective measures of BOL, such findings must be interpreted with caution, and further research is required to provide significant evidence for such conclusions. Such research could extend on the findings of this study in order to establish the validity of these devices as measure of BOL, in particular whether the devices are effective measures of BOL and whether learning effects occur if the devices are used over a long period of time.
References


