The battle of predicting vigilance: alpha power vs. skin temperature

S.M.A Buil, BSC (5630762)

Master: Applied cognitive psychology

1st Supervisor: Dr. Christoph Strauch

2nd supervisor: Prof. Dr. Leon Kenemans

Date: 15-07-2021

University Utrecht

Abstract

Predicting vigilance in the near future (30 seconds) could be beneficial in preventing mistakes in traffic situations, hospitals or border control. Previous research has shown that alpha power as measured with EEG and skin temperature might be used for predicting vigilance. The current study attempted to replicate those earlier findings and transfer them to other tasks. It also focussed on creating more of a temporal view surrounding the relationship between skin temperature and vigilance. Hits and misses on long sustained tasks were compared on their relative alpha power and distal-proximal gradient of skin temperature in the 30 second period pre-stimulus onset. The current study found no differences between hits and misses and was thus unable to replicate or transfer those earlier findings. There are multiple differences between the current study and earlier research discussed that might be the reason for the inability to replicate. As it stands however, the current results did not show a viable option for preventing mistakes with the use of alpha power or skin temperature in the near future.

Key words: *Vigilance, alpha power, skin temperature, long sustained tasks*

Introduction

Sleepiness is one of the main contributors to car accidents (National Transportation Safety Board, 1999), up to 20% of all car accidents occurred as a results of higher than normal sleepiness (Inoue & Komada, 2014). A direct consequence of higher than normal sleepiness is lower vigilance (Mathis & Hess, 2009). Vigilance can be described as a state of readiness to undertake action, react to stimuli and pay attention to one's environment. These characterizations imply that a lower level of vigilance can increase the chance of mistakes and thus the chance on causing a driving accident (Schmidt et al., 2009; Campagne, Pebayle, & Muzet, 2004). Being able to measure and predict someone's lower than normal vigilance might help in preventing errors in traffic situations, due to the effect that low vigilance has on the chances of creating a car accident. Predicting low vigilance might also be beneficial in other sectors such as hospitals (Geiger-Brown et al., 2012) or border control (Lee, Ang, Neo, Goh & Liew, 2019). In its predictivity for possible mistakes, and, by consequence accidents, vigilance can be seen as superior to sleepiness due to its direct link to performance (Reinke, Özbay, Deiperink & Tulleken, 2015). But how do we predict vigilance and thus upcoming mistakes?

Vigilance is normally measured with the use of long sustained tasks (Fisk & Schneider, 1981), but using these tasks in real-life situations is highly impractical. Therefore recent studies have focused on finding other measurements that might be useful in measuring vigilance and predicting upcoming mistakes. O'Connell and colleagues (2009) showed that alpha power as measured with an electroencephalogram (EEG) can be used to predict vigilance as indicated by mistakes on tasks. Lower alpha power was found to predict mistakes up to 20 seconds pre-target. In another study of Romeijn and van Someren (2011) they showed that skin temperature measured on the chest, finger and wrist might also be used due to its relationship with lower vigilance and thus mistakes. These two studies (O'Connell et al., 2009; Romeijn & van Someren, 2011) showed that bodily measurements could be used in predicting mistakes, but these are both based on just one experiment with the use of one task. The current study will therefore focus on replicating those findings to consolidate the state of knowledge. While replication of earlier studies is useful for validating earlier findings (Francis, 2012), it is not always enough to completely validate earlier findings. Hence, the current study also focuses on generalizing the findings of O'Connell and colleagues (2009), and Romeijn and van Someren (2011) by not only attempting to replicate findings on the exact same tasks (O'Connell and colleagues (2009): continuous temporal expectancy task (CTET), Romeijn and van Someren (2011): brief stimulus reaction time task (BSRT)), but also extending the findings to the relative other task (i.e., the BSRT for O'Connell and colleagues (2009) and the CTET for Romeijn and van Someren (2011)).

If these findings could be successfully replicated and transferred, then new avenues are opened in which alpha power and skin temperature could be used as indicators of vigilance. Before discussing the current experiment in more detail, the concept of vigilance and relevant theory will be discussed, than psychophysiological correlates of vigilance are introduced, namely electroencephalography (EEG) and skin temperature. Lastly the expected results and implications will be discussed.

Why vigilance matters

Sleepiness is very often mentioned as one of the main contributors to fatal car accidents around the world (Amini, Rezapur-Shahkolai, Khodaveisi, Gorijan & Soltanian, 2020; Horne & Reyner, 1999 ; Malek, Halvani & Fallah, 2011). Sleepiness can be described as a psychological need for sleep and it increases in a response to a lack of sleep (Cluydts, Valkc, Verstraeten, &

Theys, 2002). One of the consequences of higher sleepiness is lower vigilance. Vigilance is a concept that is closely linked to sleepiness, but they are not interchangeable or reciprocal to one and other. Vigilance is comprised of wakefulness, alertness and attention. Vigilance describes a state of readiness to undertake action, react to stimuli and pay attention to one's environment (Mathis & Hess, 2009). Vigilance is based on more factors then only sleepiness, factors like motivation, time on task or the consumption of substances (alcohol, caffeine or drugs) all play a role (Campagne, Pebayle, & Muzet, 2004). Levels of vigilance can also change rapidly during the day (Craig, Wilkinson, & Colquhoun, 1981). Lower levels of vigilance are characterized by less attention to the task at hand and lower alertness. Vigilance is closer linked to actual performance then sleepiness and therefore the current study will focus on the way that vigilance is linked to mistakes instead of sleepiness (Reinke, Özbay, Deiperink & Tulleken, 2015).

Vigilance can thus be seen as a factor affecting performance in a variety of tasks, one example for this set of tasks are the long sustained tasks (Fisk & Schneider, 1981). These tasks are created around the idea of relatively long and random interstimulus intervals. Longer so that a participant can dwell off and lose concentration and random so that a participant cannot anticipate on the upcoming target. The amount of mistakes on tasks like these can be seen as an indication of the levels of vigilance, as can reaction time (Torkamani-Azar, Kanik, Aydin,& Cetin, 2020). More mistakes and a longer reaction time correspond to lower levels of vigilance.

Implicitly measuring vigilance

The previously reviewed literature on vigilance is mostly built on behavioural data, i.e., mostly reaction times and error rates. Besides these behavioural data, psychophysiological data can also be useful in measuring vigilance and predicting upcoming mistake. In the current study two long sustained task will be used to gather a behavioural performance measure of vigilance;

the continuous temporal expectancy task (CTET) and the brief stimulus reaction time task (BSRT). The CTET is used by O'Connell and colleagues (2009) in their study to study the relationship between different aspects of the EEG and vigilance, while the BSRT was used by Romeijn and van Someren (2011) to study the relationship between skin temperature and vigilance. The relationship between the validated behavioral data that can be obtained with the use of these task can be put in relationship with psychophysiological data, in this case alpha power and skin temperature, in the hope to validate these new measurements of vigilance. The benefit of using psychophysiological data is the usage of these measurements in real-life situations. Measuring skin temperature during a driving tasks is a possibility, while conducting cognitive tasks during a driving task is highly impractical.

EEG alpha power and vigilance

Spectral analysis is one of the standard methods used for quantification of the EEG. The power spectral density (power spectrum) reflects the 'frequency content' of the signal or the distribution of signal power over frequency (Ahirwal & Londhe, 2012). In other words it represents the strength of different frequencies of brain waves. Alpha power is a section of the power spectrum for the frequencies that are commonly described as alpha waves (8 – 13 Hz) (Kalat, Cacioppo, & Freberg, 2017). Alpha power is normally highest during a calm and relaxed state and it increases when people go to sleep (Cantero, Atienza, & Salas, 2002). A decrease in alpha power normally coincides with higher visuospatial attention (Limbach, & Corballis, 2017), faster reactions (Yin, O'Halloran, Plon, Sandman, & Potkin, 2009) and lower sleepiness (Kaida et al., 2006).

These characterisations of alpha power overlap with some of the earlier mentioned definitions of vigilance. Vigilance, in the current study, is mostly considered to be a measure of performance on long sustained tasks. A predictor for vigilance is thus considered, when its able to distinguish between a good and a bad performance. Alpha power is proven to be higher for mistakes then for misses on a variety of tasks; cognitive (Valentino, Arruda & Gold, 1993), visual (Van Dijk, Schoffelen, Oostenveld & Jensen, 2008) and driving (Schmidt, 2009). This makes alpha power a prospect for predicting vigilance. However the predictive value for mistakes of alpha power is somewhat less researched, but it was the focus of the study of O'Connell and colleagues (2009). In this study they compared hits and misses on a long sustained task in their alpha power preceding the stimulus. Their results show a significant difference between hits and misses up to 20 seconds pre-stimulus, in which alpha power was higher for misses than for hits. A higher alpha could thus be used as a predictor of low vigilance. The study of O'Connell and colleagues (2009) only used the CTET and is of one few that show this strong of a predictive value, therefore extra validation could be beneficial.

Skin temperature and vigilance

Conducting an EEG during tasks such as driving or other complex tasks might be hard, because of the movements and the invasiveness of the EEG itself. Therefore a use of other indicators of activation might be more practical for measuring and predicting vigilance. Romeijn and van Someren (2011) proposed the use of skin temperature as a predictor for vigilance. The benefit of using skin temperature is that it is easier and cheaper to measure during real-life tasks with the use of sensors attached to the body or an infrared camera. A thermal infrared imaging-based fatigue detector could provide a valid, reliable, real-time, non-invasive measure of skin temperature variations (Diaz-Piedra, Gomez-Milan & Di Stasi, 2019; Costa et al., 2018). An infrared camera could thus be installed in cars or operating rooms to measure skin temperature and signal the operator when their vigilance as measured with skin temperature drops, making skin temperature much more applicable to real-life situations.

Core temperature follows a diurnal pattern in which it cools off at night to save energy and warms up during the day to prepare the body for action (Refinetti, 2010). To cool the body, it increases the blood flow to distal regions to facilitate the loss of heat to the surrounding (Romeijn et al., 2012b). This heat transfer changes the skin temperature at these more distal areas, creating a diurnal pattern opposite to that of core temperature (Lara, Molina, Madrid, & Correa, 2018). But in contrast to core temperature, skin temperature tends to fluctuate more during the day. These fluctuation may in fact represent affections (Rimm-Kaufman, & Kagan, 1996), stress (Herborn et al., 2015) or even cognitive processes such as inhibitory control (Lara et al., 2018) or exam scores (Khan, Villanueva, Vicioso, & Husmann, 2019). These fluctuation also correlate with alpha power and other neural markers of vigilance (Higuchi, Liu, Yuasa, Maeda & Motohashi, 2001; Nozawa & Tacano, 2009), showing that skin temperature is connected to vigilance as well.

Higher skin temperature in proximal regions, especially chest temperature, is connected to more mistakes and longer reaction times, (Romeijn & Van Someren, 2011). Using a gradient is more practical, because using the skin temperature on only one point such as the chest is less practical in real-life situations, because it can be influenced by the surroundings. An increase of the distal-proximal skin temperature gradient normally coincides with higher feelings of sleepiness and lower vigilance (Romeijn et al., 2012b). Skin temperature as measured with sensors on the skin is correlated to the amount of mistakes on long sustained tasks, and thus vigilance. The downside of the study of Romeijn and Van Someren (2011) is that they tested their hypothesis with the use of only one task, just like O'Connell and colleagues (2009) did for the predictive value of alpha power. Another downside of the study of Romeijn and Van Someren (2011) is that it does not offer a temporal indication of the relationship between skin temperature and vigilance. It was shown that skin temperature covaries with the probability of mistakes and reaction times (Romeijn & Van Someren, 2011), but it is yet unclear how this develops over time. Therefore the current study will not only focus on replicating and transferring these findings to the relative other task, but also on creating a better view on how this correlation develops over time.

Hypothesis

The current study will focus on investigating the predictiveness of alpha power and skin temperature for vigilance. If focusses on replicating and transferring earlier findings of O'Connell and colleagues (2019) and Romeijn and Van Someren (2011). The current study also tries to develop more of a temporal view around the relationship between skin temperature and vigilance. The goal of this study is to help in developing warning tools that work on predicting low vigilance.

1A. It is expected that alpha power will be predictive up to 20 seconds pre-stimulus as shown by O'Connell and colleagues (2009) on the BSRT. It is expected that alpha power will be higher for misses than for hits up to 20 seconds pre stimulus onset.

1B. The same pattern is predicted for the CTET in which alpha power is expected to be higher for misses than for hits

2A. Skin temperature is expected to be predictive for mistakes on the BSRT, but no temporal dimension is available in the current literature. The current study will in contrast to the original paper only focus on the distal-proximal gradient. This gradient is expected to be higher for mistakes than for misses.

2B. Skin temperature is also expected to be predictive for mistakes on the CTET, for which the distal-proximal gradient is expected to be higher for mistakes then for hits.

Method

Participants

To be eligible, participants had to meet certain requirements. First off, participants could not have any covid-like symptoms or recently been in contact with someone who tested positive on covid. Secondly the participant had to have a normal sleep rhythm. Thirdly, the participant had to have normal or corrected to normal eye-vision. Fourthly, participants had to be non-smokers. Lastly the participants had to fill in the informed consent form before they were able to participate in the experiment. To participate on the day of the experiment, the following requirements must be met: no stimulating substances up to 10 p.m. the night before the day of the experiment (like alcohol, drugs, or caffeine), a normal amount of sleep the night before the experiment, and they should not use hair product during the day of the experiment.

In total n = 34 participants took part in this experiment (15 male, 19 female; $M_{Age} = 25.9$, $SD_{Age} = 9.8$). Two participants were left-handed, the rest was right-handed. 13 participants participated during the afternoon, the others during the morning. The participants were compensated by receiving course credit or financial compensation. Because of a variety of reasons some participants were to be excluded before data analysis (see appendix A for a full overview of the exclusions).

Apparatus

All experiments were conducted within a lab environment at Utrecht University. Stimuli were presented on a windows monitor featuring a 60 Hz refresh rate. The headrest was placed at +/- 30 cm eye-distance to the screen. A standard keyboard was used to record key presses. The room was dimly lit with a temperature of 20.68 °C (SD=0.56).

Continuous EEG was recorded using the ActiveTwo Biosemi electrode system from 32 electrodes using the 10-20 system. Vertical eye movements were recorded using 2 external electrodes placed above and under the left eye. Horizontal eye movements were recorded using 2 external electrodes placed left from the left eye and right from the right eye. 2 external electrodes were also placed on the mastoid to be used as neutral reference. The data was analysed using BrainVision Analyzer 2.

Skin temperature was measured using two different techniques. One of which were iButtons (type DS1922L, Maxim/Dallas, USA) placed on different part of the body to measure the skin temperature. The iButtons sample with a .0625 °C resolution at 2-second interval. The method has been validated (van Marken Lichtenbelt et al., 2006). These iButtons were placed on the chest, finger, nose, forehead, behind the ear, on the ear, and one in the room for comparison.

Skin temperature was also measured using a FLIR infrared camera (type E53). The output of the camera was beyond the scope of the current study.

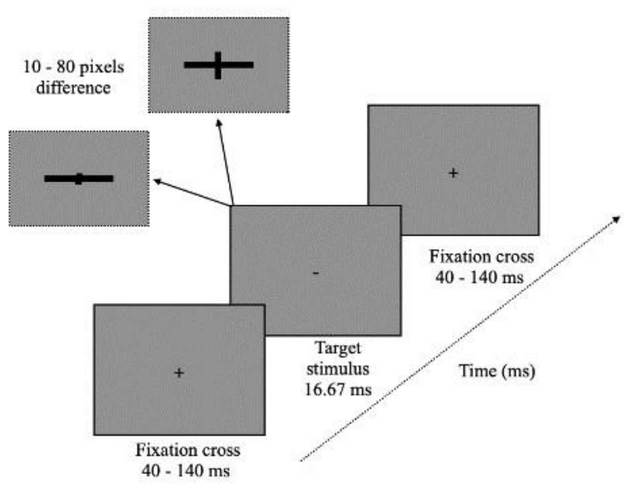
Tasks

brief stimulus reaction time task (BSRT)

Developed by Romeijn and van Someren (2011), the BSRT increases the likelihood of lapses as well as the range of reaction times compared to earlier sustained attention tasks; the here employed version was adapted slightly. Adaptations were made to cope with programming problems and to better EEG-processing by trying to increase the amount of misses. The Participants had to focus on a black cross (+) in the middle of a grey screen and respond as quickly as possible when the vertical line of the fixation cross became shorter (cue), in the original task the black cross turned in to a hyphen (-). The task had a duration of 19 minutes and was characterized by a short stimulus presentation time of the cue (16.67 ms, originally 25 ms) and a long interstimulus interval (4 – 14 s). A 2-up-1-down staircase design was added to change difficulty by changing the length of the vertical axis for the cue. The starting difference was 50 pixels, which could be decreased/increased to 10/80 pixels (see image 1 for a schematic overview of the BSRT). Reaction times were recorded as the time between presentation of the cue and key press. A lapse was recorded if participants failed to respond within 1000 ms. The task consisted of 144 targets (120 in the original); a trial block was presented upfront to familiarize with the task. For more information see Romeijn and Van Someren (2011).

Image 1

Schematic overview of the BSRT. A target is defined as a cross with a shorter vertical line

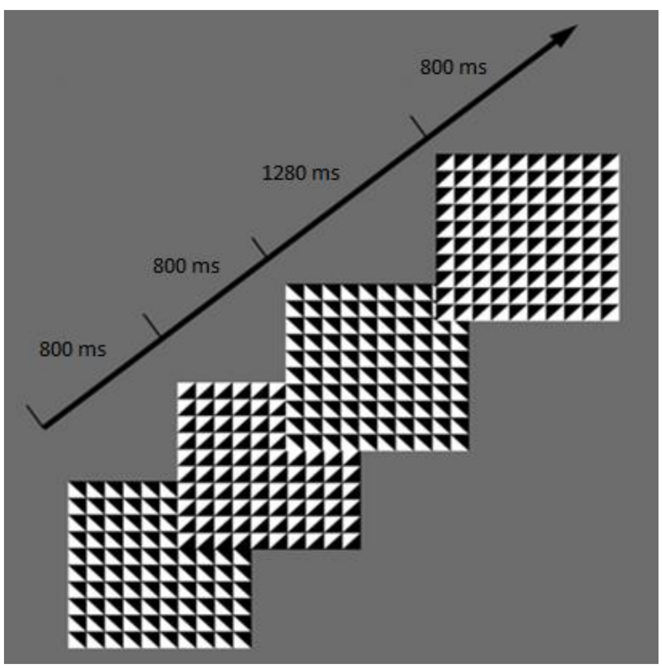


Continuous temporal expectancy task (CTET)

The CTET was used in O'Connell et al. in (2009), small adaptations were made however. Adaptation were made to cope with programming problems and in an attempts to increase the amount of misses for better EEG-processing. Participants were presented with a 8 cm^2 large square divided into a 10*10 grid of identical small squares. Each small square was diagonally split into black and white halves. The grid orientation shifted by 90° in a random direction (clockwise or counter clockwise) resulting in four different patterns. A white cross was presented in the middle of the grid to be used as a focus point for participants. Participants had to react by pressing the space bar when the grid was presented longer than non-target grids. Non-target grids were present for 800 ms, while target grids were present for 1280 ms (1120 ms in the original) (see image 2 for a schematic overview of the CTET).

Image 2

Schematic view of the CTET. Target is defined as the stimuli that is visible for 1280 ms.



Stimuli were pseudo-randomized in a way that there were 7 to 15 non-target grids present between the target, which accounts for 5.6 to 12 seconds. O'Connell et al. (2009) also measured steady-state visual-evoked potentials (SSVEP). To generate these SSVEP's, the stimulus flickered on and off at a constant rate of 25 Hz. All participants completed 10 blocks of the task and were given the choice to take a rest break in between each block. Each block consisted of 255 trials (225 in the original) with a total duration of about 3.5 minutes. The number of targets varied between 18 and 22 per block. Before conducting the real experiment participant had to perform two training blocks (one in the original). For more information on the CTET see O'Connell et al. (2009).

Questionnaire's

During the experiment the participants also had to fill in different questionnaires:

- Pittsburgh Sleep Quality Index (Buysse, Reynolds III, Monk, Berman, Kupfer, 1989)
- Morningness Eveningness Questionnaire (Horne & Östberg, 1976)
- Minimal demographic features (Sex, age, preferred hand, date/time)

For the current project, questionnaire data besides demographic information was not of interest

Procedure

At the start of the experiment the participant requirements earlier stated were checked. Then the experimenter started with applying the iButtons. The experimenter also placed electrodes for the EEG on the face and on the EEG-cap. After placing all the iButtons, electrodes, and EEG the experiment could be started. First a resting-state recording was made for the EEG. First 5 minutes with the eyes open and then 2.5 minutes with the eyes closed. After the resting-state recording the participants performed the BSRT and after that the participant performed the CTET. Upon completion, the experimenter disconnected and removed all the different measuring tools and debriefed the participant about the experiment. See image 3 and 4 for a view of the participant before the start of the experiment.

Image 3

Front view of the participant

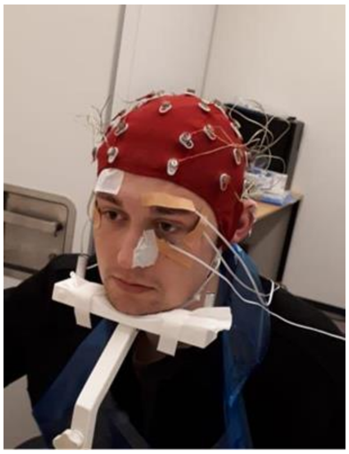
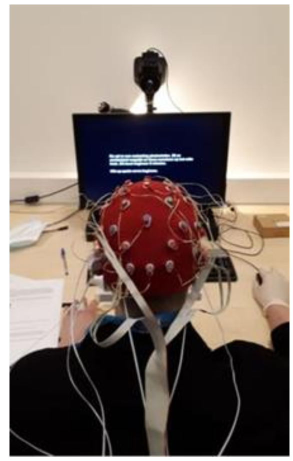


Image 4

Rear view of the participant



Data processing

Alpha power analysis

First the neural reference was changed to the external electrodes placed on the mastoids (instead of the central part of the skull). A sampling rate of 64 was deemed sufficient. The low cut-off filter was put on 0.5 Hz and the high cut-off filter was put on 31.9. O'Connell and colleagues (2009) used a higher high cut-off, but the current study only focusses on alpha power so a lower high-cut-off is sufficient. Then a 30 second segment was created ranging from 29.2 seconds pre-stimulus to 0.8 seconds post-stimulus for both hits and misses on both the CTET and the BSRT. Segments that included either data preceding the task or data from the breaks within the CTET were excluded. Segments containing an artifact (+/- 50 mV) ((+/- 90 mV in the original paper) were deleted. An ocular correction was performed to eliminate the effect of blinks and eye-movements on the electrical signal of the electrodes placed on the scalp, as was in the original study of O'Connell and colleagues (2009). A second artifact rejection was performed using the same criteria. Based on the original study of O ‘Connell (2009) concerning alpha power and vigilance the following electrodes were selected: Pz, P3, P4, PO3 ,and PO4.

The segments of 30 seconds were cut in to 15 segments of 2 seconds each. On each 2-second segment a Fast-Fourier transformation was performed to create a power spectrum (Bingham, Godfrey & Tukey, 1967). The same amount of trials for the hits and misses categories were created by selecting an equal amount of trials in the middle of the overrepresented group as of the underrepresented group. Then an average power spectrum was calculated for both hits and misses. O’Connell and colleagues (2009) defined alpha power as the activity between 8-14 Hz, but the current study focusses on the activity between 8-12 Hz. This difference in the definition of alpha power is chosen because 8-12 Hz overlaps more with what is common in the literature (Doppelmayr, Klimesch, Stadler, Pölhuber & Heine, 2002; Fink and Benedek, 2014; Schier, 2000). This resulted in 15 average powers for each of the 2-second segments at 5 electrode sites.

Skin temperature analyse

Skin temperature was measured using iButtons with a sampling 0.5 Hz. Three different temperatures were taken into account: finger, chest, and the distal-proximal gradient (difference between chest and finger temperature). For each temperature, an average for each 2-second segment preceding the stimulus was calculated. Resulting in a total of 15 average skin temperatures for each of the 2-seconds segments for the finger, the chest, and the distal proximal gradient, further analysis will only focus on the gradient. The analysis in the current study differs strongly from the one used by Romeijn and van Someren (2011). In the study of Romeijn and van Someren (2011) they compared deciles of skin temperature on the amount of misses and found higher temperatures corresponding with more mistakes. The current studies aims to explore more of the temporal dimension surrounding the relationship between skin temperature and vigilance by creating time segments for the period pre-stimulus, in a manner comparable to that of O'Connell and colleagues (2009) for alpha power.

Results

Behavioral Data

The average hit rate of the participants included in the data for the BSRT was 0.752 (SD = 0.121) and for the CTET the average hit rate was 0.558 (SD = 0.143).

Alpha power

BSRT

To compare the 15 time segments of both hits and misses on the BSRT in their alpha power preceding the stimulus for 5 electrodes, a one-way repeated measures analysis of variance (ANOVA) was used. The assumption of normality was not met, meaning the alpha power for close to none of the time segments for none of the electrodes was evenly distributed. This is not of big concern, because the ANOVA is fairly robust against violating the assumption of normality (Allen, Bennet & Heritage, 2014). The assumption of sphericity was not met as indicated by Mauchly's test of sphericity, therefore the Huynh-Feldt epsilon will be used. The assumption of homogeneity of variance was met.

The ANOVA yielded no difference between hits and misses when compared on their alpha power, $F(1,12) = 1.479$, $p = .247$. The results also showed no main effect of time on alpha power preceding the stimulus, $F(14,168) = 1.601$, $p = .213$ or an interaction effect present for time*performance $F(14,168) = 2.072$, $p = .155$, partial $\eta 2 = .147$. There was no effect present which indicated differences between hits and misses, different time segments or an interaction effect between time*performance for alpha power on the BSRT. See Figure 1A for the alpha power at electrode Pz pre-stimulus (see appendix B for other electrodes).

CTET

To compare the 15 time segments of both the hits and misses on the CTET in their alpha power preceding the stimulus for 5 electrodes, a one-way repeated measures analysis of variance (ANOVA) was used. The assumption of normality was not met, meaning the alpha power for close to none of the time segments for none of the electrodes was evenly distributed, but an ANOVA is fairly robust against the violation of this rule (Allen, Bennet & Heritage, 2014). The assumption of sphericity was not met as indicated by Mauchly's test of sphericity, therefore the Huynh-Feldt epsilon will be used. The assumption of homogeneity of variance was met.

The ANOVA yielded no difference between hits and misses when compared on their alpha power, $F(1,13) = 1.252$, $p = .283$. The results also showed no main effect of time on alpha power preceding the stimulus, $F(14,182) = .549$, $p = .670$ or an interaction effect present for time*performance $F(14,182) = .988$, $p = .405$. There was no effect present which indicated differences between hits and misses, different time segments or an interaction effect between

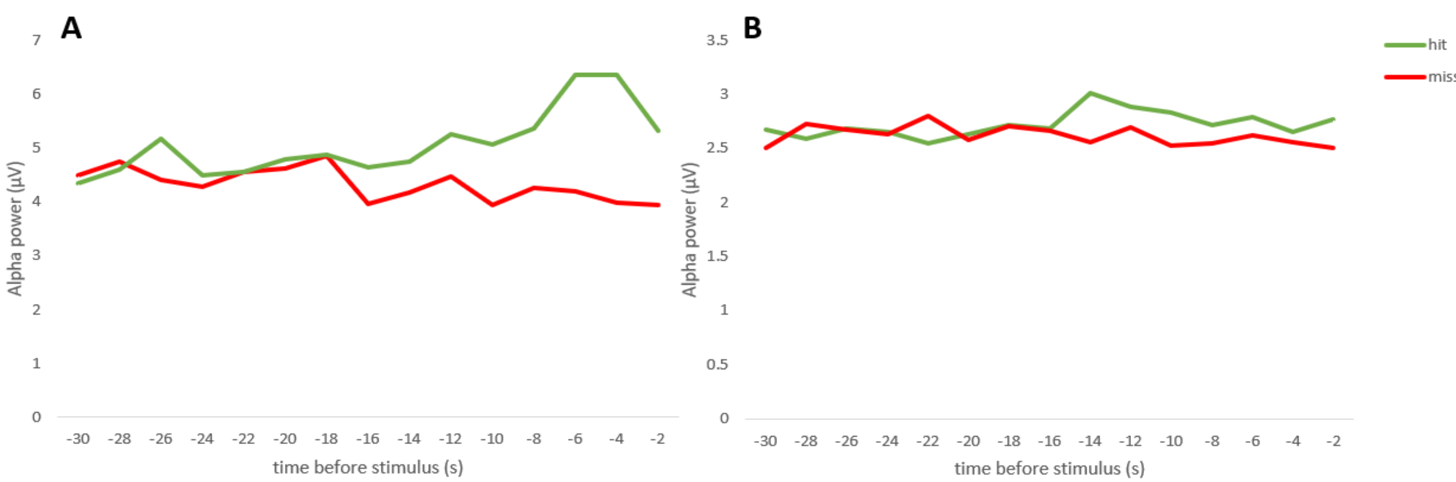


Figure 1. Alpha power as measured on electrode Pz calculated over the 30 second period preceding target onset and averaged over all participants. Green line represents hits and the red line represent misses. **A**. data concerning the BSRT. **B**. data concerning the CTET

time*performance for alpha power on the CTET. See Figure 1B for the alpha power at electrode Pz pre-stimulus (see appendix B for other electrodes).

Skin temperature

BSRT

to compare the 15 time segments of both the hits and misses on the BSRT for the distal-proximal gradient of skin temperature preceding the stimulus, a one-way repeated measures analysis of variance (ANOVA) was used. The assumption of normality and homogeneity were met. The assumption of sphericity was not met as indicated by Mauchly's test of sphericity, therefore the Huynh-Feldt epsilon will be used.

No difference was found between hits and misses when compared on their distal-proximal gradient, $F(1,15) = .006$, $p = .941$. The results also showed no main effect of time on the distal-proximal gradient preceding the stimulus, $F(14,210) = 2.588$, $p = .102$ partial $\eta 2 = .15$. and there was also no interaction effect present for time*performance, $F(14,210) = .326$, $p = .818$. There was no effect present which indicated differences between hits and misses, different time segments or an interaction effect between time*performance. See figure 2A for the average distal-proximal gradient in the 30 seconds preceding the target.

CTET

to compare the 15 time segments of both the hits and misses on the CTET for the distal-proximal gradient of skin temperature preceding the stimulus, a one-way repeated measures analysis of variance (ANOVA) was used. The assumption of normality and homogeneity were met. The assumption of sphericity was not met as indicated by Mauchly's test of sphericity, therefore the Huynh-Feldt epsilon will be used.

No difference was found between hits and misses when compared on their distal-proximal gradient, $F(1,15) = 1.214$, $p = .288$. The results showed a main effect of time on the distal-proximal gradient preceding the stimulus, $F(14,210) = 6.458$, $p = <.05$ partial $\eta 2 = .301$. and there was no interaction effect present for time*performance, $F(14,210) = .876$, $p = .436$. There was no effect present which indicated differences between hits and misses or an interaction effect between time*performance. There was an effect of time in which the gradient was lower for time segments closer to the target onset. See figure 2B for the average distal-proximal gradient in the 30 seconds preceding the target.

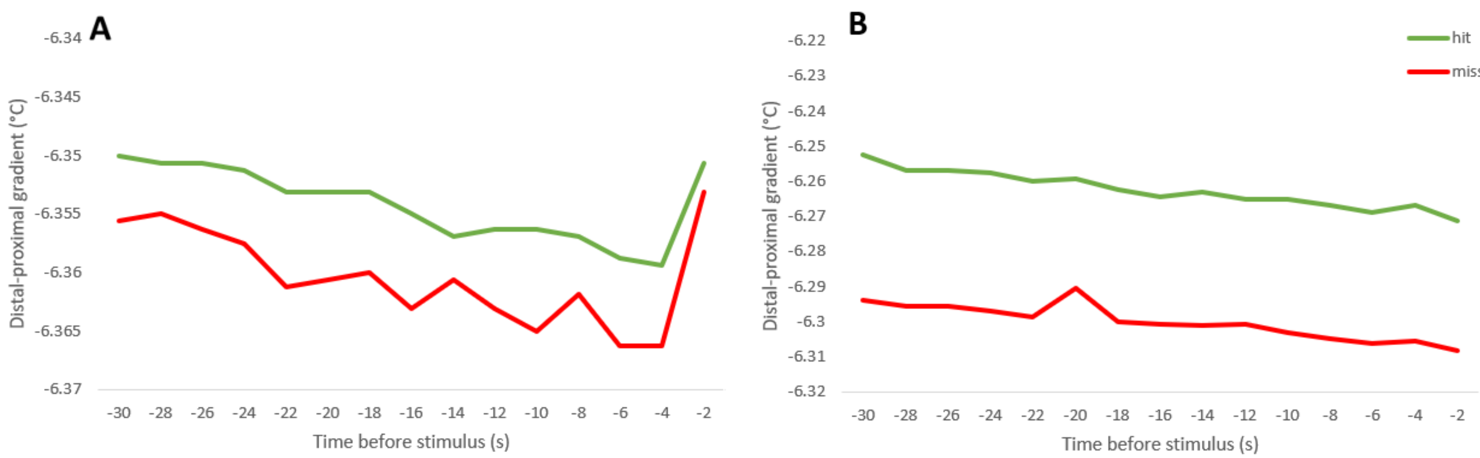


Figure 2. The distal-proximal gradient of skin temperature calculated over the 30 second period preceding target onset and averaged over all participants. Green line represents hits and the red line represent misses. **A**. data concerning the BSRT. **B**. data concerning the CTET

Discussion

The current study investigated the predictiveness of alpha power and skin temperature for vigilance as measured by misses on long sustained tasks. The study tried to replicate and transfer earlier findings of O'Connell and colleagues (2009) on alpha power and Romeijn and van Someren (2011) on skin temperature. The current findings suggest that there is no foreshadowing found for misses in the alpha band, in contrast to the findings of O'Connell and colleagues

(2009). Alpha power is commonly associated with calm and relaxed states and it increases during the period before sleep onset (Cantero, Atienza, & Salas, 2002). These characterization make alpha power the perfect candidate for predicting vigilance. Vigilance can be seen as a state of preparedness for action and of alertness (Mathis & Hess, 2009), thus the opposite state of the state commonly associated with alpha power. It was hypothesized that a decrease in vigilance, as indicated by misses on the tasks performed, would be associated with an increase in alpha power. The BSRT and CTET were successfully used in creating misses due to lack of vigilance and these misses were compared to the hits in their relative alpha power, however no differences were found between hits and misses. The data even suggest a pattern in the opposite direction, in which alpha power increases for hits and not for misses.

The current findings are also not in line with earlier findings of Romeijn and van Someren (2011) on the relationship between skin temperature and vigilance. Skin temperature can fluctuate because a variety of reasons; e.g. the diurnal pattern of core temperature (Lara, Molina, Madrid, & Correa, 2018), or environmental changes (Romanovsky, 2014). Skin temperature fluctuations are associated with a variety of different neural processes, such as affections (Rimm-Kaufman, & Kagan, 1996), stress (Herborn et al., 2015), neural markers of vigilance (Higuchi, Liu, Yuasa, Maeda & Motohashi, 2001; Nozawa & Tacano, 2009), or even cognitive processes such as inhibitory control (Lara et al., 2018). The study of Romeijn and van Someren (2011) also showed a connection between fluctuations of skin temperature and vigilance as measured on a long sustained task. The current study successfully used the same task and a second comparable task and fluctuations of skin temperature were successfully measured and compared between hits and misses, but no differences were found. The current study failed to show any connection between skin temperature and performance on either of the long sustained tasks.

This raises the question concerning to why the current study did not find any results that are in line with previous research, not for alpha power nor skin temperature. There are multiple differences between the current study and the original studies that might have led to the inability to replicate earlier findings. First differences and problems concerning the results on alpha power are discussed, then the same for skin temperature and lastly general limitations are discussed.

Alpha power

One of the reasons for the inability to reproduce earlier findings might be the frequency of the steady-state visual evoked potentials (SSVEP). The CTET uses a flickering of the stimuli to elicit a SSVEP (O'Connell et al., 2009). In the original study this led to activity at 25 Hz, while in the current study this activity was at a much lower frequency (+/- 12 Hz). This frequency overlaps with the frequency of the alpha band width (8-12 Hz). An SSVEP at this frequency (12 Hz) leads to relatively strong and fast peaks, which could interfere with finding other neural patterns (Bakardjian, Tanaka & Cichicki, 2010). The interference of the SSVEP in the current study might have made finding an effect for alpha power impossible. Adopting a new band width for alpha power might solve this problem for the CTET. The issue however is that for the BSRT no differences were found either, meaning the impact of the SSVEP alone is not enough to explain the current results.

Secondly, for the analysis of EEG it is important to have the same amount of hits and misses before averaging, because of the equalizing of signal-to-noise ratio (SNR) (Luck, 2014). It is also important to have at least 10 trials before averaging because of the SNR. The results of some participants for tasks used (BSRT & CTET) did not always meet these requirements and therefore some participants had to disqualified. Especially the BSRT often resulted in less than 10 mistakes after artifact rejection, for which participants had to disqualified. A differentiation of the task could be used in further research to increase the amount of misses, so that less participants have to be disqualified. The BSRT as used in the current study is fairly short and only exists of 1 block, an extra block could be added to increase the amount the mistakes. For the CTET the difference in time between the target grid and non-target grids could be reduced to increase the amount of misses.

Skin temperature

First of all, the current study only found really small changes in skin temperature for the distal-proximal gradient within participants (about 0.1 °C). In the original study of Romeijn and van Someren (2011) they found changes up to 1.5 °C for chest temperature, but no precise change for the gradient is mentioned. In another study by Molina, Sanabria, Jung & Correa (2019) they found a bigger difference (about 1.5 °C) between the highest and lowest distal-proximal gradient. The small changes found in this study are not completely in line with existing literature and make it hard to find significant results. These small changes in skin temperature also make it harder to use skin temperature in a practical sense in a real-life situations. Skin temperature can change quickly due to environmental changes (Romanovsky, 2014), so picking up this small of change is almost impossible in real-life situations.

Secondly, skin temperature is heavily influenced by the circadian effect of time of day on skin temperature (Romeijn et al. 2012a). Especially proximal skin temperature is strongly influenced by the sleep-wake cycle (Molina, Sanabria, Jung & Correa, 2019). The time of day on which the experiment takes place can thus influence not only chest temperature, but also the distal-proximal gradient. The distal-proximal gradient in the current study was lower at the target onset then 30 seconds pre-stimulus on the CTET, again showing the effect that time of day can have on skin temperature. Participants did participate on different hours and can have different circadian rhythms (Kandeger, Selvi & Tanyer, 2018). Future research should add extra measurements which take in account the time of day and the circadian rhythm of participants, because this can influence the distal-proximal gradient and thus the results (Romeijn et al. 2012a).

Thirdly, the current study differs strongly in methodology concerning the statistics in comparison to the original paper of Romeijn and van Someren (2011). The current study compares hits and misses on their skin temperature gradient in the 30 second window pre-target, while the original paper (Romeijn & van Someren, 2011) compares different deciles of skin temperatures on their amount of hits and misses. This change was made to see if there could be an application for real-life situations in which people could be warned on the basis of their skin temperature. Discovering more about the temporal view was necessary to develop a warning system that would be able to warn people in a timely manner. However, this differentiation could lead to the inability to find similar results, although the current study also found no main effect for performance (hit vs. miss) either. A more precise replication of the original study might be beneficial in forming a conclusive view on the relationship between skin temperature and vigilance.

General limitations

The current study falls within a larger research-project into predicting vigilance using EEG and skin temperature. One of the downsides of this organization is the multitude of researchers. Although a protocol was clearly formed, differences can exist between the different experimenters which may lead to different results. The exact location of the iButtons and/or electrodes can for example differentiate between researchers. Another downside of the multitude of researchers is the initial difficulty of performing EEG research. Many participants had to be disqualified because of unusable EEG data, sometimes only 13 participants remained of the 34 in total. This is a downside for the current study as more participants is beneficial for scientific research and increase the chance of finding firm and conclusive findings (Hackshaw, 2008).

Blinking during targets during the BSRT could lead to misses not because of low vigilance but because the participants just did not see the target. With the lay-out of the current study we were unable to delete misses on the basis of eye-blinks. A check was performed in which the amount of eye-blinks between hits and misses was comparable for all the participants (10% marge). Future research could still adapt to deleting the trials in which eye-blinks occur altogether to overcome possible problems.

Conclusion

The current study was unable to replicate and/or transfer earlier findings concerning alpha power and skin temperature on their relationship to vigilance. Possible reasons for this and future adaptations are discussed above. For the time being a gap exist between the findings of the current study and earlier studies, which further research might close. The small changes in skin temperature found in this study make measuring skin temperature hard to imagine as viable option for application as of today. The same accounts for the invasiveness of the EEG. The current study hoped to find a relationship that could help in preventing mistakes in traffic, or other setting as hospitals or border control. As it stands however, the current results do not show a viable option for preventing mistakes with the use of alpha power or skin temperature in the near future.

Literature list

Ahirwal, M. K., & Londhe, N. (2012). Power spectrum analysis of EEG signals for estimating visual attention. International Journal of computer applications, 42(15), 22-25.

Amini, R., Rezapur-Shahkolai, F., Khodaveisi, M., Gorjian, S., & Soltanian, A. R. (2020). The relationship between sleep quality and road traffic crashes of urban drivers in Hamadan, Iran. Journal of injury and violence research, 12(1), 47.

Bakardjian, H., Tanaka, T., & Cichocki, A. (2010). Optimization of SSVEP brain responses with application to eight-command brain–computer interface. Neuroscience letters, 469(1), 34-38.

Bingham, C., Godfrey, M., & Tukey, J. W. (1967). Modern techniques of power spectrum estimation. IEEE Transactions on audio and electroacoustics, 15(2), 56-66.

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry research, 28(2), 193-213.

Campagne, A., Pebayle, T., & Muzet, A. (2004). Correlation between driving errors and vigilance level: influence of the driver's age. Physiology & behavior, 80(4), 515-524.

Cantero, J. L., Atienza, M., & Salas, R. M. (2002). Human alpha oscillations in wakefulness, drowsiness period, and REM sleep: different electroencephalographic phenomena within the alpha band. Neurophysiologie Clinique/Clinical Neurophysiology, 32(1), 54-71.

Costa, C. M. A., Moreira, D. G., Sillero-Quintana, M., Brito, C. J., de Azambuja Pussieldi, G., de Andrade Fernandes, A., ... & Marins, J. C. B. (2018). Daily rhythm of skin temperature of women evaluated by infrared thermal imaging. Journal of thermal biology, 72, 1-9.

Craig, A., Wilkinson, R. T., & Colquhoun, W. P. (1981). Diurnal variation in vigilance efficiency. Ergonomics, 24(8), 641-651.Daniel, R. S. (1967). Alpha and theta EEG in vigilance. Perceptual and Motor Skills, 25(3), 697-703.

Diaz-Piedra, C., Gomez-Milan, E., & Di Stasi, L. L. (2019). Nasal skin temperature reveals changes in arousal levels due to time on task: An experimental thermal infrared imaging study. Applied ergonomics, 81, 102870.

Doppelmayr, M., Klimesch, W., Stadler, W., Pöllhuber, D., & Heine, C. (2002). EEG alpha power and intelligence. Intelligence, 30(3), 289-302.

Van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. Journal of Neuroscience, 28(8), 1816-1823.

Fink, A., & Benedek, M. (2014). EEG alpha power and creative ideation. Neuroscience & Biobehavioral Reviews, 44, 111-123.

Fisk, A. D., & Schneider, W. (1981). Control and automatic processing during tasks requiring sustained attention: A new approach to vigilance. Human factors, 23(6), 737-750.

Francis, G. (2012). The psychology of replication and replication in psychology. Perspectives on Psychological Science, 7(6), 585-594.

Geiger-Brown, J., Rogers, V. E., Trinkoff, A. M., Kane, R. L., Bausell, R. B., & Scharf, S. M. (2012). Sleep, sleepiness, fatigue, and performance of 12-hour-shift nurses. Chronobiology international, 29(2), 211-219.

Hackshaw, A. (2008). Small studies: strengths and limitations.

Herborn, K. A., Graves, J. L., Jerem, P., Evans, N. P., Nager, R., McCafferty, D. J., & McKeegan, D. E. (2015). Skin temperature reveals the intensity of acute stress. Physiology & behavior, 152, 225-230.

Higuchi, S., Liu, Y., Yuasa, T., Maeda, A., & Motohashi, Y. (2001). Diurnal variations in alpha power density and subjective sleepiness while performing repeated vigilance tasks. Clinical neurophysiology, 112(6), 997-1000.

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. International journal of chronobiology.

Horne, J., & Reyner, L. (1999). Vehicle accidents related to sleep: a review. Occupational and environmental medicine, 56(5), 289-294.

Inoue, Y., & Komada, Y. (2014). Sleep loss, sleep disorders and driving accidents. Sleep and Biological Rhythms, 12(2), 96-105.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical neurophysiology, 117(7), 1574-1581.

Kalat, J.W., Cacioppo, J.T., & Freberg, L.A. (2017). Brain and Behaviour (Utrecht university customized edition). Cengage learning.

Kandeger, A., Selvi, Y., & Tanyer, D. K. (2019). The effects of individual circadian rhythm differences on insomnia, impulsivity, and food addiction. Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity, 24(1), 47-55.

Khan, T. H., Villanueva, I., Vicioso, P., & Husman, J. (2019, October). Exploring relationships between electrodermal activity, skin temperature, and performance during. In 2019 IEEE Frontiers in Education Conference (FIE) (pp. 1-5). IEEE.

Lee, V., Ang, J., Neo, A., Goh, L. Y., & Liew, N. (2019). Operational Vigilance in Border Security: the Singapore Experience. Journal of Police and Criminal Psychology, 34(3), 330-339.

Limbach, K., & Corballis, P. M. (2017). Alpha-power modulation reflects the balancing of task requirements in a selective attention task. Psychophysiology, 54(2), 224-234.

Luck, S. J. (2014). An introduction to the event-related potential technique. MIT press.

Malek, M., Halvani, G. H., & Fallah, H. (2011). A study of the relationship between the Pittsburgh sleep quality index and road accidents among truck drivers. Occupational Medicine quarterly journal, 3(1), 14-20.

van Marken Lichtenbelt, W. D., Daanen, H. A., Wouters, L., Fronczek, R., Raymann, R. J., Severens, N. M., & Van Someren, E. J. (2006). Evaluation of wireless determination of skin temperature using iButtons. Physiology & behavior, 88(4-5), 489-497.

Mathis, J., & Hess, C. W. (2009). Sleepiness and vigilance tests. Swiss Med Wkly, 139(15-16), 214-219.

Molina, E., Sanabria, D., Jung, T. P., & Correa, A. (2019). Electroencephalographic and peripheral temperature dynamics during a prolonged psychomotor vigilance task. Accident Analysis & Prevention, 126, 198-208.

National Transportation Safety Board. (1999). Evaluation of U.S. Department of Transportation Efforts in the 1990s to Address Operator Fatigue. National Transportation Safety Board, Safety Report NTSB/SR-99/01, Washington, D.C.

Nozawa, A., & Tacano, M. (2009). Correlation analysis on alpha attenuation and nasal skin temperature. Journal of Statistical Mechanics: Theory and Experiment, 2009(01), P01007.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. Journal of Neuroscience, 29(26), 8604-8611.

Refinetti, R. (2010). The circadian rhythm of body temperature. Front Biosci, 15(1), 564-594.

Reinke, L., Özbay, Y., Dieperink, W., & Tulleken, J. E. (2015). The effect of chronotype on sleepiness, fatigue, and psychomotor vigilance of ICU nurses during the night shift. Intensive care medicine, 41(4), 657-666.

Rimm-Kaufman, S. E., & Kagan, J. (1996). The psychological significance of changes in skin temperature. Motivation and Emotion, 20(1), 63-78.

Romanovsky, A. A. (2014). Skin temperature: its role in thermoregulation. Acta physiologica, 210(3), 498-507.

Romeijn, N., & Van Someren, E. J. (2011). Correlated fluctuations of daytime skin temperature and vigilance. Journal of biological rhythms, 26(1), 68-77.

Romeijn, N., Raymann, R. J., Møst, E., Te Lindert, B., Van Der Meijden, W. P., Fronczek, R., ... & Van Someren, E. J. (2012a). Sleep, vigilance, and thermosensitivity. Pflügers Archiv-European Journal of Physiology, 463(1), 169-176.

Romeijn, N., Verweij, I. M., Koeleman, A., Mooij, A., Steimke, R., Virkkala, J., ... & Van Someren, E. J. (2012b). Cold hands, warm feet: sleep deprivation disrupts thermoregulation and its association with vigilance. Sleep, 35(12), 1673-1683.

Schier, M. A. (2000). Changes in EEG alpha power during simulated driving: a demonstration. International Journal of Psychophysiology, 37(2), 155-162.

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. Accident Analysis & Prevention, 41(5), 1087-1093.

Torkamani-Azar, M., Kanik, S. D., Aydin, S., & Cetin, M. (2020). Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state EEG in a long sustained attention task. IEEE journal of biomedical and health informatics, 24(9), 2550-2558.

Valentino, D. A., Arruda, J. E., & Gold, S. M. (1993). Comparison of QEEG and response accuracy in good vs poorer performers during a vigilance task. International Journal of Psychophysiology, 15(2), 123-133.

Appendix A

Table 1

Participant exclusions for each psychophysiological measurement for both tasks

	Alpha power on the BSRT	Alpha power on the CTET	Skin temperature on the BSRT	Skin temperature on the CTET
Outlier (hit-rate) (2*SD)	1	0	1	0
Not enough misses (10+)	5	0	5	0
Technological errors	3	2	3	2
Missing data	1	1	9	5
Did not finish task	0	3	0	3
False hit/false alarm rate (H/FA <3)	0	8	0	8
Not enough misses (10+) after artefact rejection	10	3	-	-
Outlier (alpha power) (2*SD)	1	3	-	-
Total without excluded	**13**	**14**	**16**	**16**

Appendix B

Alpha power pre-stimulus for electrodes P3, PO3, PO4 and P4

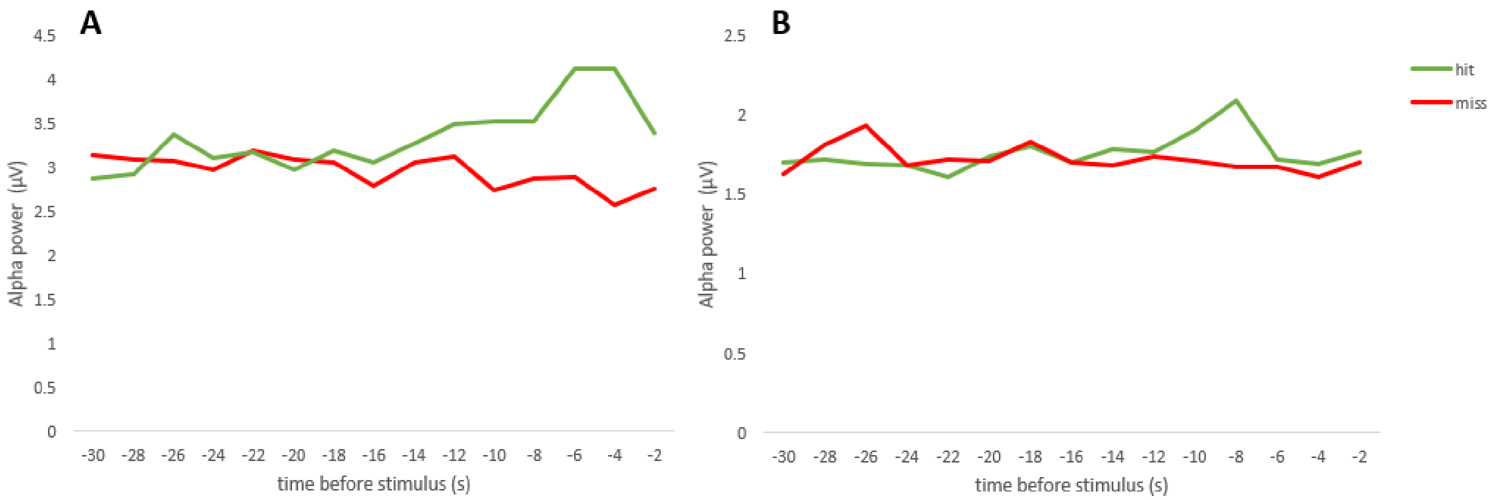


Figure 3. Alpha power as measured on electrode P3 calculated over the 30 second period preceding target onset and averaged over all participants. Green line represents hits and the red line represent misses. **A**. data concerning the BSRT. **B**. data concerning the CTET

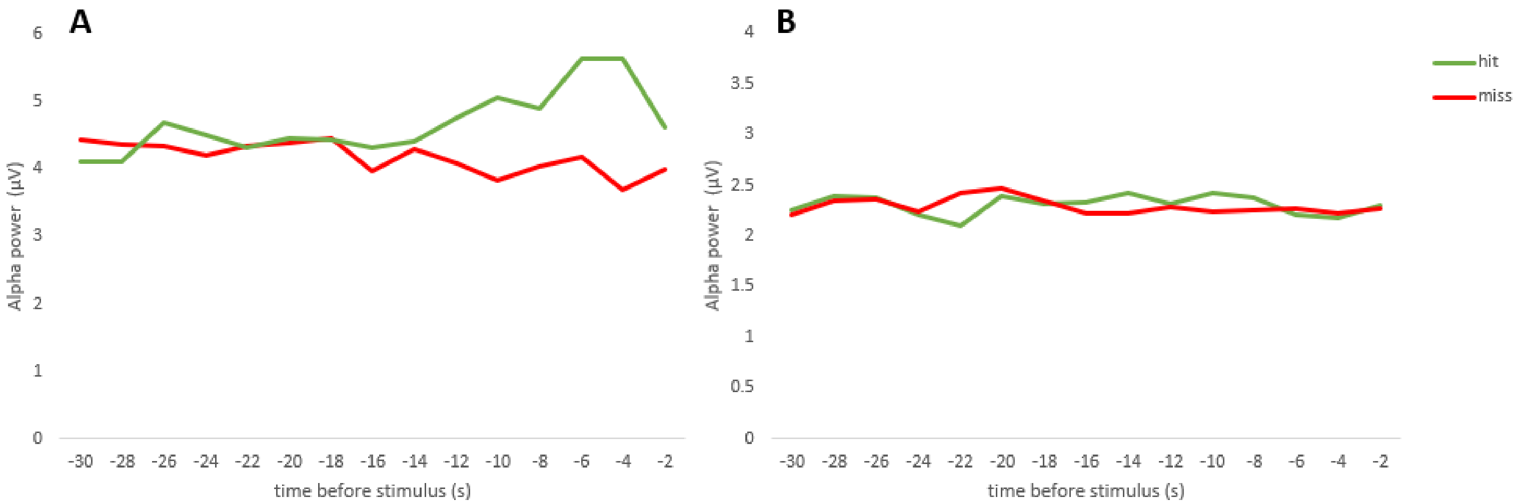


Figure 4. Alpha power as measured on electrode PO3 calculated over the 30 second period preceding target onset and averaged over all participants. Green line represents hits and the red line represent misses. **A**. data concerning the BSRT. **B**. data concerning the CTET

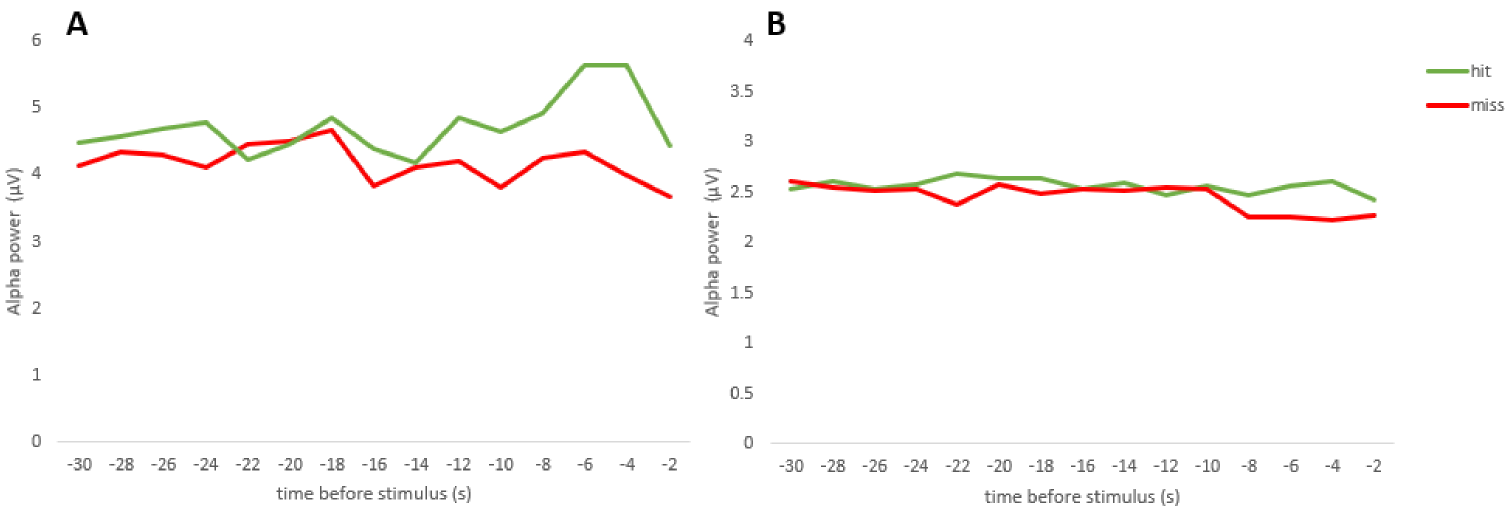


Figure 5. Alpha power as measured on electrode PO4 calculated over the 30 second period preceding target onset and averaged over all participants. Green line represents hits and the red line represent misses. **A**. data concerning the BSRT. **B**. data concerning the CTET

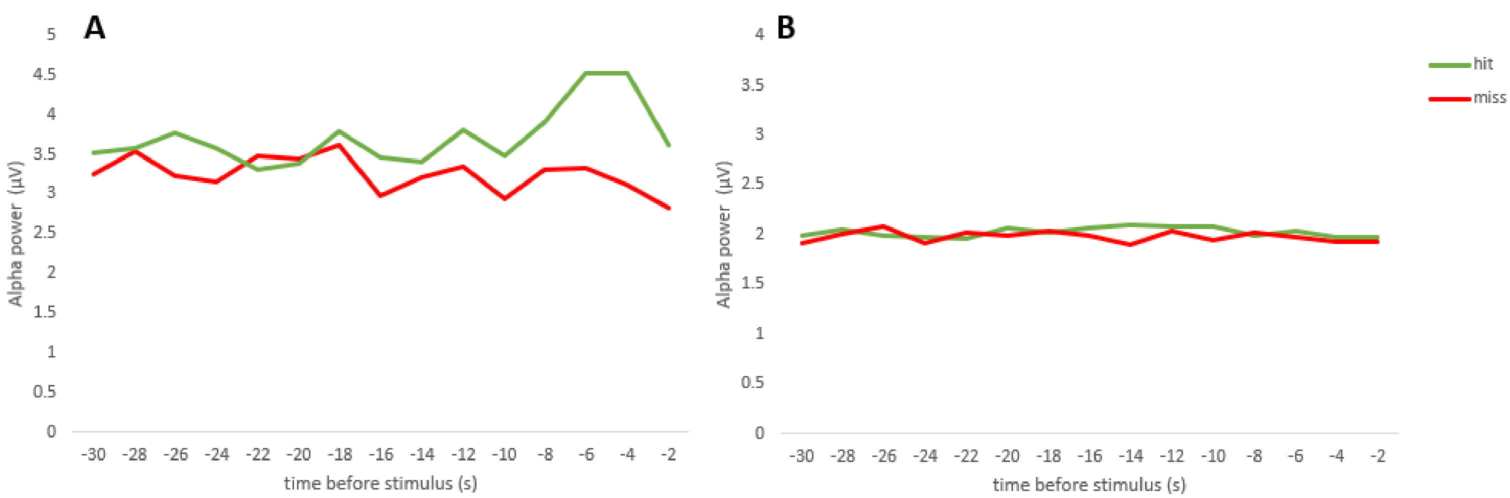


Figure 6. Alpha power as measured on electrode P4 calculated over the 30 second period preceding target onset and averaged over all participants. Green line represents hits and the red line represent misses. **A**. data concerning the BSRT. **B**. data concerning the CTET