University of Groningen

The key to mental fatigue

de Jong, Marlon

DOI:
10.33612/diss.192205878

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
de Jong, M. (2021). *The key to mental fatigue: monitoring and counteracting performance decline during prolonged office work*. University of Groningen. https://doi.org/10.33612/diss.192205878

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 29-10-2022

Design
publiek.

Paranimfen
Berry van den Berg
Rob Nijenkamp

This thesis, part of SPRINT@Work, is part-financed by the European Regional Development Fund, the province and municipality of Groningen, and the province of Drenthe [grant number T-3036, 2013].

Publication of this thesis was financially supported by the University of Groningen and the Research School of Behavioural and Cognitive Neurosciences

rijksuniversiteit
groningen

The key to mental fatigue

Monitoring and counteracting performance decline during prolonged office work

Proefschrift

ter verkrijging van de graad van doctor aan
de Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 25 november 2021 om 11.00 uur

door

Marlon de Jong

geboren op 22 februari 1991
te Deventer

Promotor
Prof. dr. M.M. Lorist

Co-promotor
Dr. J. Jolij

Beoordelingscommissie
Prof. dr. D. de Waard
Prof. dr. D. van der Linden
Prof. dr. J.B.F. van Erp

Chapters

Chapter 1

General introduction

Theoretical background

Mental fatigue

Mental fatigue is a state that emerges during or after prolonged task performance and is generally characterised by feelings of aversion against continuing the current task and a decline in task performance. Around the time of the Industrial Revolution, mental fatigue became increasingly important as a topic for scientific study (e.g., Thorndike, 1900). Due to the excessive use of machines, researchers started explaining human biology and behaviour similarly to the working of machines, thereby establishing a link between human behaviour and mental energy. The machine energy metaphor provided understanding of the meaning of fatigue, that is, (mental) fatigue was explained as a battery that ran low and needed to reload before (mental) work could be performed efficiently again. This metaphor achieved a stronghold in the English language, and still, if we talk about mental fatigue, we often refer to the feeling of being 'out of energy' or 'drained'.

The machine energy metaphor remained an important influence on scientific theories about mental fatigue throughout the 20th century. Grandjean (1979) for example, compared the manifestation of fatigue with a container that is filled with experiences during the day. According to Grandjean, this container could only be emptied by a recovery period, which would reverse or reduce the effects of (mental) fatigue. Moreover, in early versions of his model, Hockey (1993, 1997) set a limited budget for the expenditure of effort, which he defined as the energetic response to the demands of the environment. If this effort budget was smaller than the effort-costs of the task, aversion against the task would emerge, task performance would decline, or someone would disengage from the task. Later, several researchers, including Hockey himself, found evidence that participants were able to perform well on a task if they were sufficiently motivated (Hockey & Earle, 2006; Hopstaken et al., 2016). For instance, Boksem et al. (2006) observed that participants were almost able to perform as efficiently as they did at the start of the experiment if they received a monetary reward based on their performance, contradicting previous energy-depletion theories stating there was a limited energy budget. In contrast to these energy-depletion theories, Boksem and Tops (2008) argued that mental fatigue is the result of a cost-benefit evaluation of effort: if the costs of executing a task come to exceed the benefits of finishing a task, mental fatigue emerges. That is, people consciously or unconsciously decrease the amount of effort they invest in the task, with deteriorations in performance as a consequence.

Ever since researchers began to study the phenomenon, the implications of mental fatigue on daily life activities (Thorndike, 1900) and especially its negative effects on productivity (Ricci et al., 2007) and safety in the working environment (Baker et al., 1994; McCormick et al., 2012) became more evident. In general, decline in productivity due to mental fatigue is characterised by a decrease in speed and accuracy. People not only make more mistakes, they are also less able to correct for these mistakes, resulting in a decline in the overall accuracy of performance (Boksem et al., 2006; Lorist et al., 2005). Besides decline in performance, prolonged task performance results in changes in subjective state. These changes typically reflect an increase in experienced feelings of fatigue, aversion against continuing the tasks that one has to perform, and eventually the urge to stop performing this task (Hockey & Earle, 2006; van der Linden, 2010). Interestingly, decline in performance and changes in subjective state do not always go hand in hand. Previous research showed that people are very bad at estimating when their performance declines. Usually, when people begin to notice the effects of mental fatigue, performance has already deteriorated substantially (Hockey, 2013; Zhang et al., 2011).

Although in general mental fatigue has been related to a decline in performance, findings of scientific research are mainly based on average effects of time-on-task, while there can be large differences within and between participants in the effects of prolonged task performance. Davis (1946) for example, identified three groups in a participant pool of experienced airline pilots: a normal group with participants who did not show any deteriorations in task performance with time-on-task, a withdrawal group with participants that reduced the amount of effort they allocated to the task, resulting in decline in task performance, and an over-activity group consisting of participants who increased the effort that they allocated to the task, resulting in increased arousal and improved performance.

How performance decrements emerge during or after prolonged task performance depends on a wide range of personal and environmental factors. One of the factors that might influence the manifestation of mental fatigue is age. Over the past century, the average age of the working population has significantly increased. A major challenge caused by this ageing workforce is to keep workers fit for work. It is therefore specifically important to gain a better understanding of the effects of mental fatigue in older adults in working environments, and thereby investigate how working conditions and tasks could be adapted in order to meet the needs of the ageing worker. Although relatively little is known about the interaction between age and mental fatigue on performance, in general, younger and older adults tend to perform more slowly and more error-aversive than younger adults (Starns & Ratcliff, 2010). Moreover, in a meta-analysis, Verhaeghen and Salthouse (1997) showed that there is a high correlation between age and measures of reaction time and perceptual speed. These differences in behavior suggest that effects of mental fatigue might be expressed differently in younger and older adults.

Besides age, several factors have been found to influence the way in which mental fatigue is manifested in the working environment. For instance, improved motivation can counteract the effects of mental fatigue (Boksem et al., 2006; Hopstaken et al., 2016). Additionally, the effects of mental fatigue can vary over tasks. Tasks that require cognitive control and extensive planning are prone to the effects of mental fatigue (Lorist et al., 2000), while tasks that are performed automatically are relatively unaffected (van der Linden et al., 2003). Moreover, the course of mental fatigue also depends on the setting in which a task is performed. For instance, during prolonged task performance, people are less able to focus attention solely on important information and are more easily distracted by irrelevant information (Boksem et al., 2005), especially when perceptual load of the task is low (Csathó et al., 2012), resulting in larger deteriorations in task performance in turbulent settings, such as open-plan offices (Smith-Jackson & Klein, 2009).

Electro-encephalography

To investigate which cognitive processes are affected specifically, researchers can investigate the effects of mental fatigue on brain activation using electro-encephalography (EEG). EEG is a high-temporal resolution technique that measures the communication between neurons via electrodes on the scalp. Neurons communicate by transmitting electrical signals through their axons (action potential), which can then be received by other neurons through their dendrites (post-synaptic potentials). EEG mainly measures the electrical signal that is carried through the dendrites, however, only if a large number of neurons (thousands or millions) is active at the same time, the electrodes placed on the scalp can measure the electrical activity.

The processing of specific stimuli or events has been related to specific patterns of electrical activation. This unique pattern can be derived from the EEG signal by averaging brain activity recorded during multiple trials of the same condition. The resulting pattern is called an event-related potential (ERP). Comparing ERPs across different conditions can provide information about differences in the timing and organization of cognitive processes during task performance (Luck, 2005). The P300, for example, is a positive wave observed between 300 and 500 ms after stimulus presentation that has been associated with attention and memory operations (Polich, 2007, 2012). Previous research has shown that the parietal P3 amplitude decreased as a result of mental fatigue elicited by prolonged task performance. In a controlled lab setting, this smaller P3 amplitude has found to be correlated with a decline in task performance (Hopstaken et al., 2015; Lorist & Jolij, 2012; Wascher & Getzmann, 2014)

Another way to investigate brain dynamics during task performance is by extracting information about neural oscillations from the EEG signal. The EEG signal can be described by multiple oscillations, varying in frequency. Modulations in specific power frequency bands can be linked to specific cognitive processes. For instance, in the present thesis, we were especially interested in power in the alpha frequency band (7-14 Hz), because modulations in alpha power have been related to cortical activation and excitability (Scheeringa et al., 2016). Specifically, alpha power has been used as a marker of neural arousal, where an increase in alpha power reflects an increase in mental fatigue (Kenemans & Lorist, 1995). Moreover, reduced levels of alpha activity have been associated with increased levels of attention and improved levels of preparatory attention (Foxe & Snyder, 2011). Van den Berg and colleagues (2014), for example, showed that if participants could gain a reward based on their level of performance, alpha power before stimulus presentation decreased more compared to stimuli that did not lead to a potential reward, indicating increased preparatory attention.

Sensor technology and ethics

Mental fatigue has substantial negative influences on performance in daily life. Additionally, deteriorations in performance due to mental fatigue vary between and within subjects and depend on environmental context. Therefore, effectively detecting and counteracting mental fatigue is a major challenge, specifically in the working environment, where the effects of mental fatigue are particularly potent. Although brain activation can provide information about the course of mental fatigue on an individual level, and portable EEG devices could in theory monitor mental fatigue outside the lab, it is an intrusive method which strongly disrupts daily life activities. Since the industrial revolution, researchers and other innovators have developed new sensor technologies to monitor performance that are less intrusive than EEG, such as eye tracking and other performance-based methods, and could serve the purpose of detecting mental fatigue during daily life activities.

One device that seems to meet the terms of these conditions in an office environment, an environment in which employees are easily subject to the effects of mental fatigue, is the computer. The computer has become an increasingly important part of our working life. From 2005 until 2015, the percentage of employees that are intensively working with a computer, network, mainframe or smartphone has increased from 21 % to 37 % (Eurofound, 2015), and it is likely these trends remain similar, or even increase the coming years. Computers might form an opportunity to monitor behaviour dynamics that are susceptible to the effects of mental fatigue in the working environment. For instance, previous research developed a program that was able to monitor behaviour related to computer use, such as typewriting or the trajectory of the cursor (Pimenta et al., 2013, 2014), during computer activities. These measures might also be susceptible to changes in speed and accuracy on daily life tasks, and given that a large part of the working population is already using the computer, it is a simple and non-intrusive solution to monitor behaviour.

Although in principle sensor technologies are meant to support the user, they can have unforeseen consequences that are unintentionally harmful to the user or to society (Schukat et al., 2016). For instance, some health insurance companies already ask their customers to share their personal activity data, monitored via a pedometer or step counter on their phones. By doing so these individuals could earn back part of their insurance fee. Although these marketing strategies are being framed in a way that they are beneficial to the user, health insurance companies could eventually use these large amounts of data to decide on the level of insurance fees or even on whether they should take someone as a customer. While pedometers or step counting devices were firstly developed in order to help individuals improve their own health, commercial institutes already took the first steps to use these simple devices for their own commercial benefits, potentially harming the user and others in the process. So, in order to endorse responsible research and innovation, and prevent or deal with the eventual negative consequences of new technologies, guidelines have been developed that empower researchers and other innovators to incorporate their responsibilities in the entire innovative process (Grunwald, 2014; Stilgoe et al., 2013). These guidelines specifically focus on anticipation of future scenarios regarding new findings, reflection on researchers' own role, inclusion of the different stakeholders involved, and responsiveness to societal challenges.

Overview of this thesis

Previous literature showed that mental fatigue negatively affects productivity during regular work. These effects vary between people and over days due to internal and external factors. People themselves, however, are not very good at detecting these changes in performance and technologies that are now being used to investigate the effects of mental fatigue in the lab are too intrusive for use in an office environment. Therefore, the main objective of the present thesis was to monitor (**Chapter 2 and 3**) and counteract (**Chapter 4**) mental fatigue during prolonged office work, taking advantage of new developments offered by sensor technology. The development and implementation of new technologies comes with responsibilities, however. The second part of this thesis will therefore describe how we reflected on this process (**Chapter 5**).

Part I: Monitoring and counteracting mental fatigue

In the second chapter of this thesis, we investigated how mental fatigue affects neural processing during typewriting and in addition if these changes in neural activation were correlated with changes in typewriting performance for young and older adults. We were specifically interested in the relation between mental fatigue, typewriting performance, and the P3 component to examine whether typewriting indices were related to deteriorations in attentional and memory processes because of mental fatigue. Subsequently, the third chapter investigates whether the markers that were found to be susceptible to the effects in mental fatigue in a lab setting, could also describe behaviour dynamics in the working environment, and whether mental fatigue would manifest differently in younger and older adults. Additionally, we investigated how these markers were modulated over time during regular office work, investigating the effects of prolonged task performance on different time-scales (i.e., time-on-task, time-of-day, and day-of-week).

One way to enhance or maintain performance during prolonged task performance is by drinking caffeine-containing beverages, such as coffee and tea. Caffeine improves performance through an enhancing effect on the attentional system. First, caffeine has a broad sustained enhancing effect on neural arousal (Kenemans & Lorist, 1995). These effects are consistent with the stimulating effects of caffeine on behaviour that relate to an overall increase in performance. Second, caffeine also biases the processing of relevant information over irrelevant information, resulting in more effective behaviour (i.e., faster and more accurate responses) (Kenemans & Lorist, 1995; Lorist et al., 1994a; Ruijter, De Ruiter, et al., 2000). As mentioned before, deteriorations in task performance due to the effects of mental fatigue have found to be larger in rural environments, such as open-plan offices. Therefore, in the third chapter of this thesis, we focused on the relationship between caffeine consumption and the bias towards processing of behavioural relevant information, specifically investigating the effect of caffeine on preparatory attention for stimuli that could potentially lead to monetary rewards.

Part II: Responsible Research and Innovation

The first part of this thesis focuses on the technical usability of monitoring typing indices in order to detect mental fatigue and in order to improve cognitive performance during prolonged office work. However, given that meddling with employees in the working environment is accompanied by serious ethical concerns, even more so if large amounts of personal data is involved, it is important to think about the impact of a new technology on its potential users or even on society as a whole. Therefore, the second part of this thesis focuses on reflecting on and dealing with the implications of new technologies in the working environment.

In **Chapter 5**, we explored the interplay between ethical issues in the design and implementation of technologies for workplace health promotion on the basis of two case studies of the project SPRINT@Work (See Appendix I for more information about SPRINT@Work). We specifically focused on two ethical issues, privacy and autonomy, that play an important role in the worker-employer relation, especially given the (financial) dependency of workers on their employers (Autoriteit Persoonsgegevens, 2020). A context-specific approach of ethics was applied during the reiterating process of development and small-scale implementation of health-related technologies in the workplace. During several intervision sessions, the researchers of SPRINT@Work and an ethicist investigated whether the legal framework of privacy sufficiently protects the worker from harm when using of the new technologies in the context of workplace health promotion, and whether the technologies could sufficiently enable the autonomy of workers when using these technologies in the context of workplace health promotion. This approach makes it possible to already take into account the possible impact of new technologies in their design. By using this context-sensitive approach, we aimed to overcome the divide between the development and implementation phase of new technologies, prevent out-of-context generalization during the implementation of new technologies and prevent vagueness of responsibilities of the stakeholders involved in the design and implementation process.

Chapter 2

Age modulates the effects of mental fatigue on typewriting

This chapter is based on:

De Jong, M.
Jolij, J.
Pimenta, A
Lorist, M.M. (2018)

Age Modulates the Effects of Metal Fatigue on Typewriting.
Frontiers in Psychology, 9(July), 1–15.
https://doi.org/10.3389/fpsyg.2018.01113

Abstract

In the present study, we examined whether age influences the effects of mental fatigue on task performance, and if we could validate the use of measures based on typing behaviour as an index of the effects of mental fatigue on different aspects of cognition. Young (N=24, 18-30y) and middle-aged (N=24, 50-67y) participants performed a typewriting task and a mouse targeting task for 120 minutes. At the beginning and at the end of the experiment the level of subjective fatigue was assessed. During task performance measures based on typing behaviour and EEG were recorded. Results showed that subjective fatigue increased over the experiment in both the young and the middle-aged group. Typing speed decreased with time-on-task in both age groups, reflected in larger general interkey intervals and in an increase in typing time. In addition, typing accuracy decreased with time-on-task in the young group, however, not in the middle-aged group, reflected by an increase in typing errors. Moreover, the young group used the backspace key more often with time-on-task due to delayed error-correction, reflected in larger backspace sequences, resulting in larger interkey intervals and increased typing time. This effect was absent in the middle-aged group. In the young group, the P3 brain potential amplitude decreased over the experiment, which was related to an increase in typing time, longer general interkey intervals, and an increase in typing errors, suggesting that decreased task engagement was related to less efficient typewriting, at least in the young group. These results indicate that measures based on typing behaviour could give information about the process of mental fatigue, and in addition suggest that age influences the effect of mental fatigue on typewriting. More specifically, younger adults more often adopt a strategy that emphasizes speed, while middle-aged adults act more error-aversive than younger adults.

Keywords: *mental fatigue, aging, typewriting, performance, ecologically sensitive, EEG, P3*

Introduction

Mental fatigue refers to a complex, and multifaceted state that may occur during or after prolonged task performance. Although the effects vary within and between people, mental fatigue is often manifested in aversive feelings against the current task and deteriorations in task performance. More specifically, with time-on-task people often perform more slowly, make more mistakes, and are less able to correct for these mistakes (Boksem et al., 2006; Lorist et al., 2005)

Since the early 1900s, the effects of mental fatigue and its implications for daily life, and working life in particular, have been addressed by numerous researchers (e.g., Ackerman, 2011; Bartley & Chute, 1949; Hockey, 2013; Muscio, 1921; Thorndike, 1900). In order to explain the effects of mental fatigue and the cognitive processes involved, several theoretical models have been developed. Bartlett (1941), for example, argued that the effects of mental fatigue are related to a decline in cognitive control, resulting in a decreased ability to coordinate and accurately time complex activities, which is accompanied by a growing irritability and awareness of physical discomfort. Kanfer & Ackerman (1989) put forward that both a decrease in cognitive resources and a decline in the motivation to allocate the residual resources to the current task play a role in explaining the effects of mental fatigue. More recently, mental fatigue has been argued to develop as a result of a cost-benefit evaluation of effort and might serve as a warning system for a decrease in performance efficiency; if the costs of performing a cognitive task come to exceed the benefits of finishing the task, subjective experience of mental fatigue (e.g., aversion against task performance, low vigilance) emerges and performance deteriorates (i.e., increase reaction times and decrease accuracy; Boksem & Tops, 2008). This subjective warning system is not optimal, however. Previous research has shown that people are not able to accurately decide whether they are still capable of effectively performing a task. Comparing performance and subjective measurements, Zhang and colleagues (2011), for example, showed that performance already declines before people actually become aware of being fatigued. This has real-world consequences: mental fatigue not only affects productivity (Ricci et al., 2007), but it is also one of the most frequent causes of accidents in a working environment (Baker et al., 1994; McCormick et al., 2012).

The effects of mental fatigue can be influenced by internal (e.g., motivation; Boksem et al., 2006) and external (e.g., caffeine; Lorist et al., 1994) factors. Age is an important example of one such factor that might influence the manner in which mental fatigue is expressed. Given that over the past century the average age of the working population has increased, it is important to gain a better understanding of the effects of mental fatigue in older adults in working environments. Although relatively little is known about the interaction between age and mental fatigue on performance, in general, older adults are slower to respond (Verhaeghen & Salthouse, 1997), and act more error-aversive than younger adults, while younger adults are faster and make more mistakes than older adults (Starns & Ratcliff, 2010). The main question in the present study was whether and how the effects of mental fatigue are modulated by age. Because of the differences in overall processing speed and error aversion, we hypothesized that the effects of mental fatigue are expressed differently in younger and older adults, as well.

Research on mental fatigue in ecologically valid settings has clearly established the implications of mental fatigue in the working environment (e.g., Baker et al., 1994; Hockey, 2013; McCormick et al., 2012; Ricci et al., 2007). One way to monitor performance during a typical working day for office workers, who are prone to mental fatigue, is by using computer input devices (e.g., keyboard and mouse). For example, Pimenta and colleagues (2014) used keyboard-based performance measures of speed and accuracy to monitor changes in typewritten performance during working days.

Based on effects of time-of-day, Pimenta et al. argued that these keyboard-based performance measures could be used to monitor mental fatigue. Additionally, Kalfaoğlu & Stafford (2014) examined the relation between typing errors and speed. They investigated how typists behaved before, during, and after correct typewriting compared to incorrect typewriting, and found that typists often slow down during and after mistakes, while speed is constant during correct typewriting. Moreover, these mistakes can be predicted by increased variability in typing speed before making mistakes. These results indicate that typing characteristics are susceptible to variability in performance, and may thus be used as a proxy for studying and monitoring mental fatigue.

Although monitoring typewriting behaviour might increase our knowledge about the consequences of mental fatigue in the working environment, it should be noted that behavioural measures only describe the consequences of mental fatigue on behaviour. That is, behaviour is the final result of many different cognitive processes that can be influenced by internal (e.g., mental fatigue) and external factors (e.g., task demands). Measures of behaviour can provide some understanding about these cognitive operations, however, they do not give information about these processes directly and therefore their relation with these underlying cognitive processes remain elusive. Measuring brain activity using electro-encephalography (EEG) can offer direct insight in the cognitive processes underlying task performance. Event-related potentials (ERPs) can be extracted from the EEG signal and describe the timing and organisation of cognitive processes before, during, or after external (e.g., stimulus presentation) or internal (e.g., making a mistake) events. The P1 (i.e., a positive wave between 100 – 160ms after stimulus presentation) and N1 (i.e., a negative wave between 160 – 210ms after stimulus presentation) ERP components, for example, are associated with early processing of visual information (Luck, 2005, p36-p37), The N400, a negativity wave in the 250 –500ms interval after stimulus presentation, was found to be inversely correlated with the semantic fit of a presented word in the context of a sentence (Hoeks et al., 2004; Kutas & Hillyard, 1984). The P3, a positive wave between 300 – 400ms after stimulus presentation, has been associated with attention and memory operations (Polich, 2007). Although the exact mechanisms underlying the P3 are unclear and topic of debate, it is known that mental fatigue and decreases in performance, are associated with decreased P3 amplitudes (Hopstaken et al., 2015; Lorist & Jolij, 2012; Wascher & Getzmann, 2014). The second aim of this study was to elucidate the relation between behavioural indices of typing behaviour and underlying cognitive processes. To validate if indices of typing behaviour could monitor the effects of mental fatigue on specific cognitive operations, we recorded brain activity in addition to behavioural measures.

In sum, in the present study we examined whether age influences the effects of mental fatigue on task performance. During the experiment, participants performed a typewriting task (i.e., an adapted version of the task of Hoeks et al. (2004)) and a mouse targeting task for 2 hours. Time-on-task was used to induce mental fatigue. In line with previous ageing research (Hertzog et al., 1993; Starns & Ratcliff, 2010; Verhaeghen & Salthouse, 1997), we hypothesized that speed would decline more with time-on-task in the middle-aged group than in the young group, and accuracy would decline more in the young group than in the middle-aged group. Based on different measures derived from typing behaviour we examined different aspects of human performance. Besides measures of typing speed (typing time, general interkey interval, and stimulus-key interval) and typing accuracy (incorrect words, corrected words, backspace use, and backspace sequence length), we investigated the letter-letter interval, an index for speed during correct typewriting, the letter-backspace interval, which is an index for the initiation of error-correction, and the backspace-backspace interval, which is an index of speed during error-correction. With regard to the relation between effects of mental fatigue on different measures of typing behaviour and underlying cognitive processes, we expected that top-down cognitive control (e.g., attentional and working memory processes) would be most prone to the effects of mental fatigue during typewriting. Therefore we hypothesized a decrease of the P3 amplitude over the experiment and an increase in P3 latency with time-on-task. Furthermore, we expected that changes in the P3 component would be related to fatigue-related changes in typewritten performance.

Materials and methods

Participants

Twenty-four healthy young adults (8 males), ranging in age from 18 to 30 (*M*=22.4, *SD*=3.4), and twenty-four healthy middle-aged adults (11 males), ranging in age from 50 to 67 (*M*=57.8, *SD*=6.0) gave their written informed consent to participate in this experiment that was approved by the local Ethics Committee. All participants were Dutch, right-handed, had normal or corrected-to-normal visual acuity. Both the young and the middle-aged participants worked at least 1 hour every day on a computer (Young: *M*=4.0, *SD*=2.2; Middle-aged: *M*=4.6, *SD*= 2.9). Sixteen young and 13 middle-aged participants were capable of typing according to the 10-finger system without looking at the keyboard. The participants were not dyslectic, did not use prescription medication, and did not work night shifts. The young adults were students who received either course credits or 20 euro in exchange for their participation; the middle-aged adults were recruited via partner companies of the SPRINT@Work project, local media outlets, social media, and personal contact, and received a lunch or 20 euro in exchange for their participation. Data from two participants (one participant in each age group) was excluded from the analysis due to excessive noise in the EEG data (>30% of the epochs).

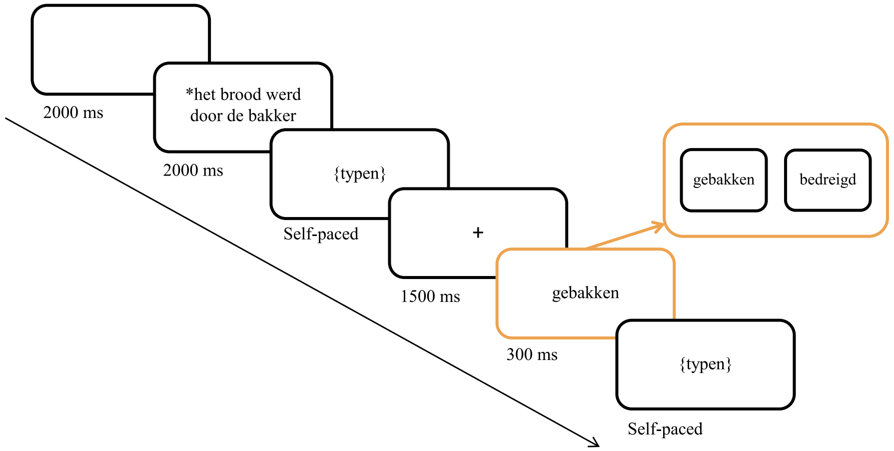
Apparatus and materials

The experiment was conducted in a sound-attenuated room. The experimental setup consisted of an office chair placed behind an adjustable desk, a windows computer with a QWERTY keyboard, screen support, and an Iiyama prolite G2773HS monitor, model PL2773H. The desk-top lighting was fixed at 800 lux, set according to the Dutch norms for lighting in an office environment. The experimental task was written in PsychToolbox version 3.0 (Brainard, 1996; Kleiner et al., 2007; Pelli, 1997), a toolbox for Matlab R14b (The Mathworks Inc., 2014). Performance data were acquired using PsychToolbox and a keylogging utility (CAMCof, version 7/10/14; Pimenta et al., 2014). EEG activity was recorded using 20 Sn electrodes on an electrode cap (ElectroCap International). EEG analysis was performed using Matlab R14b in combination with EEGlab (Delorme & Makeig, 2004). Statistical analysis was conducted in R (R Core Team, 2015; RStudio Team, 2016) using the packages lme4 (Bates, Mächler, et al., 2015) and lmerTest (Kuznetsova et al., 2017).

Typewriting task

The experiment consisted of two tasks; a typewriting task and a mouse targeting task, alternating on a trial by trial basis. A trial of the typewriting task had a mean duration of 16.5s (SD=4.2) and the mean duration of a trial in the mouse targeting task was 6.6s (SD=0.7). In the present paper we discuss the typewriting task, which is an adapted version of the task used by Hoeks and colleagues (2004).

Figure 1: *Schematic overview of a trial of the typewriting task.*



**het brood werd door de bakkers gebakken/bedreigd. 'the bread by the bakers baked/threatened.' [the bread was being baked/threatened by the baker]*

Each trial of the typewriting task started with a blank screen presented for 2000ms, followed by the first part of a sentence (i.e., all, but the last, word; Figure 1). After 2000ms the text was replaced with a blank screen, and the participants had to type the text they had read during sentence presentation. After finishing typing, the participants pressed Enter and a fixation cross appeared (1500ms). Then, the last word of the sentence appeared in the middle of the screen, which was replaced with a blank screen after 300ms. The participant had to type the last word, again followed by pressing the Enter key to continue to the next trial. The sentences were presented in the Courier New font, with a size of 30 points.

Two types of sentences were presented, differing with regard to the semantic fit of the last word. The last word could either semantically fit with the first part of the sentence (congruent sentences; 75% of the sentences; e.g., het brood werd door de bakker gebakken. 'the bread by the bakers baked.' [the bread was baked by the baker]) or no semantic fit (incongruent sentences; 25% of the sentences; e.g., het brood werd door de bakker bedreigd.'the bread by the baker threatened.'[the bread was threatened by the baker]). Before the start of the experiment, sentence types were randomized with restriction, so that different incongruent sentences were separated by at least one congruent sentence. The duration of the experimental task was set to 120 min. As a consequence, the number of trials performed by each participant was dependent on how fast participants completed the trials (Young: *M*= 357, *SD*=31.9; Middle-aged: *M*=281, *SD*=34.9).

Subjective measures

The Activation-Deactivation Adjective Check List (AD ACL; Thayer, 1989) was used to measure momentary activation states experienced by the participants. The participants indicated to what extent their mood could be described by 20 adjectives on a 4-point scale, ranging from 1 (definitely) to 4 (definitely not). The questionnaire categorizes the momentary state of activation into 4 clusters: general activation, deactivation/sleep, high activation and general deactivation.

Performance measures

The keylogging software registered a timestamp at the start and at the end of each keystroke with a code that referred to the identity of the pressed key (i.e., letter or backspace). Based on these markers, performance measures related to typing speed and typing accuracy were calculated for the sentence and the last word of each trial (Table 1). Typing speed is reflected in typing time, interkey interval, and stimulus-key interval. Typing time reflects the total time it takes to type the sentence or the last word of one trial. Interkey interval reflects the time between two subsequent keypresses, and is divided into four categories: general interkey interval (average time between two subsequent keypresses irrespective of the type of key), the letter-letter interval (average time between two subsequent letters), the letter-backspace interval (average time between a letter and a backspace), and the backspace-backspace interval (average time between two subsequent backspaces). The stimulus-key interval reflects the time between the offset of the last word of the sentence and the first keypress made to type this word. In addition, typing accuracy is reflected in the following measures: incorrect words, corrected words, backspace use, and backspace sequence length. The incorrect words represent the percentage of incorrectly typed words, the corrected words represent the percentage of words that are corrected, backspace use reflects the percentage of keystrokes that are backspaces, and backspace sequence length reflects the average key length of a sequence of consecutive backspaces

Table 1: *Keyboard-based performance measures related to typing speed and typing accuracy.*

Typing speed	Typing accuracy
• Typing time • Interkey interval: ○ General interkey interval ○ Letter-letter interval ○ Letter-backspace interval ○ Backspace-backspace interval	• Incorrect words • Corrected words • Backspace use • Backspace sequence length

EEG recordings and analysis

EEG activity was recorded using 20 Sn electrodes (F7, F3, Fz, F4, F8, FCz, C3, Cz, C4, CPz, P7, P3, Pz, P4, P8, PO7, PO8, O1, Oz, O2) attached to an electrode cap. Reference electrodes were placed behind each ear and a common electrode was placed on the collar bone. EOG activity was bipolarly recorded from two Sn electrodes placed at the outer canthi of both eyes for horizontal eye movements and above and below the right eye for vertical eye movements. Electrode impedance was kept below 5 kOhm.

We used independent component analysis (ICA) to identify eye blinks and eye movements (Jung et al., 2000). First, the data was down sampled to a rate of 500 Hz and filtered with a high pass filter of 0.5 Hz and a low-pass filter of 60 Hz. Independent component decomposition was conducted on epochs of 2500ms, starting 1000ms before sentence presentation. Epochs containing artefacts (amplitude < -120 or > 1250 µV) were excluded from ICA decomposition. Channels that contained technical artefacts were replaced with interpolated values of surrounding electrodes (i.e., 0.8 channels on average per participant). In seven participants, resulting ICs did not reflect eye movements and eye blinks. IC decomposition in these participants was performed on 2500ms epochs around the last word of the sentence. Finally, ICs were copied to the original raw data, which was down-sampled to 500 Hz and band-pass filtered between 0.15 and 40 Hz. IC components that reflected eye blinks or eye movements (1 or 2 ICs per participant) were removed from the data.

Subsequently the data was segmented in epochs of 1200ms, starting 200ms before the presentation of the last word of the sentence. Epochs in which the maximal absolute signal amplitude exceeded 110 µV were excluded from analysis. For four participants, settings were slightly adjusted because of larger average amplitudes in these datasets; in one dataset the high-pass filter was set on 0.5 Hz, in two datasets a threshold of 130 µV was used and in one dataset a threshold of 160 µV was used. The data was re-referenced to linked mastoids and aligned to a 200ms pre-stimulus baseline.

The relationship between the independent variables (i.e., time-on-task, congruency, and age) and event-related brain activity (i.e., amplitude and latency of ERP components) was expressed in event-related regression coefficients (ERRCs) for every participant, channel, and time-point in the selected time-window. In other words, an ERRC expresses the relation between ERP amplitude and an independent variable such as time-on-task per participant (see Hauk et al. (2006) and Miozzo et al., (2015) for a further discussion of ERRCs). From these modelled ERPs, the P1 was quantified as the mean amplitude in the 50ms interval surrounding the most positive value between 100 and 160ms after stimulus onset at the averaged signal of O1 and O2 (Boksem et al., 2005), and the N1 as the mean amplitude in the 50ms interval surrounding most negative value between 160 and 210ms post-stimulus at P7 and P8 (Boksem et al., 2005). The P3 was calculated as the mean amplitude surrounding the most positive value between 300 and 500ms in the averaged signal of the electrodes

CPz, Pz, P3, and P4 (Saliasi et al., 2013). Along with P3 amplitude, 50% fractional area latency was calculated in the P3 time-window (Hansen & Hillyard, 1980; Luck, 2005). The mean amplitude of the N400 component was computed on the basis of the averaged signal of the electrodes Pz, Cpz, and Cz in the 250 –500 post stimulus interval (Kutas & Federmeier, 2011).

Procedure

Participants were instructed to abstain from alcohol 24 h before the experiment and from caffeine-containing substances for at least 12 h before the experiment. The experimental session started at 9 pm and took approximately 3.5 h. It was explained to the participants that the study was aimed at investigating differences in information processing between young and middle-aged employees. They were unaware that the analysis also included the factor time-on-task. Before the start of the task electrodes were applied, the participants were asked to hand in their phone and watch, and desk, chair, and screen support were adjusted according to the occupational health and safety guidelines. The participants filled out the AD-ACL and a questionnaire that contained general questions. Written task instructions were provided at the start of the experiment. For the typing task the participants were instructed to read and remember the first part of the sentence that appeared on the screen and type it as fast and accurately as possible after the text on the screen had disappeared. The same instruction applied to the last word of the sentence. Participants were asked to use the backspace key to correct for typing errors. To check whether the participants understood the task, they had to perform three practice trials. Thereafter, the participants performed the experimental tasks for 120 min, after which they filled out the AD-ACL again.

Statistical analysis

Subjective data. For analysing subjective fatigue, reflected in scores on the AD-ACL, we used mixed effects models with time of administration as a within-subject factor, and participant as a random factor. Missing values on the AD-ACL were replaced by the mean scores of the participant on the subscale of the missing value. Separately for both age groups and for each subscale of the AD-ACL, scores were centred around the mean of the first measurement. The raw data were used to calculate the means.

Performance data. For the time-on-task analyses, trials were excluded from the analysis if scores on one of the behavioural measures in that trial was more than two standard deviations below or above the mean score, which was calculated within-subjects and across the whole experiment (*M*= 5.8%, *SD*=2.9). Thereafter, trial data was pooled in 5 minute bins, and average performance was computed for each bin. We used mixed effects models with age as a between-subject factor, time-on-task (ToT) as a within-subject factor, and participant as a random factor (see Table 1). Congruency (Cong) was included as an additional factor in the model used to analyse performance during typing of the last word. In the time-on-task analysis, a varying slope based on level of typing skill (i.e., the ability to type according to the 10-finger system without looking at the keyboard or not) was added to the model to correct for differences in typing skills across participants. The final models (i.e., factors included in the model) were based on model comparison using the likelihood ratio test. Based on these final models we calculated mean values at 20 min and at 120 min after the start of the experiment.

Data transformations were performed in order to realize that residuals of the models were normally distributed and did not show auto-correlation or heteroscedasticity. The dependent variable letter-letter interval was log-transformed. The dependent variables typing time, general interkey interval, stimulus-key interval, incorrect words, backspace sequence length, backspace use, and letter-backspace interval were root-transformed. The dependent variable corrected words was not transformed. To achieve linearity, the log-transform of the independent variable time-on-task was used in the models that included the number of backspaces and the backspace sequence length as dependent variables.

In addition to the previously described time-on-task analyses, we subjected the individual performance data to a linear regression with time-on-task as a fixed factor, which resulted in regression coefficients that described the effect of time-on-task on the dependent variables, separately for each participant. These performance-based regression coefficients were subjected to bivariate Pearson's correlations to examine whether changes in the dependent variables were accompanied by changes in other dependent variables.

ERP data. The mean ERRCs, reflecting the effect of time-on-task and congruency on the mean amplitude of the different components (i.e., P1, N1, P3, and N400), were subjected to statistical analysis. T-tests were conducted to test whether the ERRCs significantly differed from zero, and ANOVAs were used to test whether the ERRCs significantly differed between age groups. To examine the relation between changes in event-related brain activity and changes in performance bivariate Pearson's correlations were computed between the performance-based regressions coefficients and the mean ERRCs of the P3 amplitude and latency in the young and the middle-aged group. If the main analysis indicated a significant interaction (alpha< .05) between factors, follow-up analyses were performed, adjusting error rates according to Bonferroni.

Results

Subjective measurements

Subjective fatigue increased during the experiment, as evidenced by a decrease in general activation scores from the start to the end of the experiment (M_{before}=14.5, M_{after}=11.1; $\chi^2(1)$=40.09, *p*<.001), and an increase in scores on the deactivation/sleep scale (M_{before}=10.1, M_{after}=14.1; $\chi^2(1)$=41.24, *p*<.001). The level of tension, reflected in high-activation scores, did not change during the experiment (M_{before}=6.5, M_{after}=6.9; $\chi^2(1)$=2.08, n.s.), although, participants felt less calm (i.e., decreased scores on the general deactivation scale) after the experiment compared to the beginning of the experiment (M_{before}=17.2, M_{after}=15.2; $\chi^2(1)$=16.41, *p*<.001). Scores on the AD ACL were not significantly influenced by age.

Congruency

Congruency influenced task performance in both the young and the middle-aged group; in both groups stimulus-key interval was longer if the last word was incongruent with the first part of the sentence compared to endings that were congruent (young: M_{cong}= 250ms, M_{incong}=317ms; middle-aged: M_{cong}=431ms, M_{incong}=518ms; Table 2). The effect of congruency was not dependent on time-on-task (Table 2), therefore the data was pooled over congruent and incongruent sentences for analyses concerning time-on-task.

Early visual processing of the final word of a sentence, reflected in the P1 and N1 amplitude, was neither affected by congruency (P1, Cong: *t*(45)=.38, *p*=.707; N1, Cong: *t*(45)=-.45, *p*=.656). ERP components related to the further processing of stimulus information (i.e., P3 and N400), however, were affected by congruency; the P3 amplitude elicited by the incongruent last word was smaller compared to congruent words (P3, Cong: *t*(45)=-5.1, *p*<.001). This difference was more pronounced in the young group than in the middle-aged group (P3, Cong*Age: *F*(1,44)=4.56, *p*=.038), young: *M*=2.7, *SE*=.51, middle-aged: *M*=1.1, *SE*=.52), however, this effect was not influenced by time-on-task (Cong*ToT: *F*(1,44)=1.86, *p*=.070). Moreover, as expected, in sentences that contained a semantic violation, the last word of the sentence elicited a larger N400 amplitude as compared with congruent sentences (Cong: *t*(45)=-3.3, *p*=.002). Processing of semantic violations, reflected in N400 amplitude, did not differ significantly between young and middle-aged participants (Cong*Age: *F*(1,44)=2.96, *p*=.093) and did not change with time-on-task (Cong*ToT: *t*(45)=-1.6, *p*=.112). β-weights, reflecting the N400 for each participant, and β-weights, reflecting the effect of congruency on stimulus-key interval for each participant, were negatively correlated in the young group (*r*(21)=-.49, *p*=.017), indicating that the conflict processing was directly related to reaction speed in the young group. This relation did not reach significance in the middle-aged group (*r*(21)=.19, *p*=.374).

Table 2: *The effects of congruency, time-on-task, age, and their interaction on stimulus-key interval described by the chi-square (χ2) and the β-weights of the final model that describes the effects of congruency, time-on-task, age, and/or their interaction on stimulus-key interval.*

Dependent variable	**Cong χ2(1)**	**Age χ2(1)**	**ToT χ2(1)**	**Cong* Age χ2(1)**	**Cong* ToT χ2(1)**	**Cong* ToT* Age χ2(1)**	**β_{Cong} (SE)**	**β_{Age} (SE)**
Stimulus-key interval	**14.83** ***	**140.48** ***	2.46	0.13	2.45	1.56	1.98 (0.16)	4.96 (1.19)

*: p<.05; **: p<.01; *** p<.001

Behavioural measures

During the first 20 minutes of task performance, the average time to type the first part of the sentence remained stable in the young group (β_{ToT} =0.06, *SE*=0.10, *p*=.53), while the average typing time decreased in the middle-aged group ($\beta_{ToT*Age}$=-0.15, *SE*=0.075, *p*=.04), indicating that in this group the typing task was subject to practice effects. Therefore, data collected during the first 20 minutes of the task was excluded from further analysis.

Typing speed

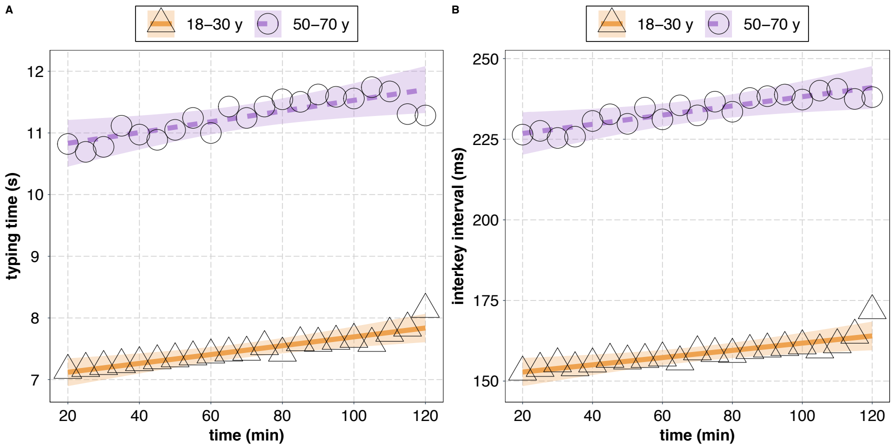
In general, average typing speed was lower in middle-aged participants (typing time$_{sentence}$: 11.3s; typing time$_{last\ word}$: 3.45s; general interkey interval$_{sentence}$: 233ms; general interkey interval$_{word}$: 253ms) than in young participants (typing time$_{sentence}$: 7.36s; typing time$_{last\ word}$: 2.08s; general interkey interval$_{sentence}$: 156ms; general interkey interval$_{word}$: 166ms). In both groups the average typing time necessary to type the first part of the sentence increased with time-on-task (Young: $M_{t=20min}$=7.25s, $M_{t=120min}$=8.03s; Middle-aged: $M_{t=20min}$=11.2s, $M_{t=120min}$=12.0s; Table 3 & Figure 2A). This increase in typing time$_{sentence}$ was, at least partly, due to an increase in the time between two subsequent key presses in the 20-120 min interval (general interkey interval$_{sentence}$: Young: $M_{t=20min}$=152ms, $M_{t=120min}$=162ms; Middle-aged, $M_{t=20min}$=228ms, $M_{t=120min}$=240ms; Table 3 & Figure 2B), as was evidenced by the relation between increased typing time$_{sentence}$ during task performance and increased general interkey interval$_{sentence}$ over the experiment in both young (*r(21)*=.88, *p*<.001) and middle-aged participants (*r(21)*=.60, *p*=.003).

Table 3: *The effects of time-on-task, age, and their interaction on the different dependent variables of speed described by the chi-square (χ2) and the β-weights of the final models that describe the effects of time-on-task, age, and/or their interaction on the dependent variables.*

Dependent variable	**ToT χ2(1)**	**Age χ2(1)**	**ToT*Age χ2(1)**	**β_{ToT} (*SE*)**	**β_{Age} (*SE*)**	**$\beta_{ToT*Age}$ (*SE*)**
Typing time$_{\text{sentence}}$	**9.40****	**38.40 *****	1.61	0.037 (0.004)	20.5 (2.67)	-
General interkey interval$_{\text{sentence}}$	**8.64****	**43.65 *****	1.45	0.003 (0.0004)	2.79 (0.33)	-
Typing time$_{\text{last word}}$	0.42	**45.61 *****	**27.45 *****	-0.01 (0.005)	11.26 (1.62)	0.031 (0.006)
Stimulus-key interval$_{\text{last word}}$	**4.70***	**44.87 *****	**16.54 *****	0.0008 (0.0012)	4.17 (0.56)	0.007 (0.002)
General interkey interval$_{\text{last word}}$	**5.69 ***	**54.62 *****	**7.65 ****	-0.001 (0.002)	2.84 (0.32)	0.004 (0.001)

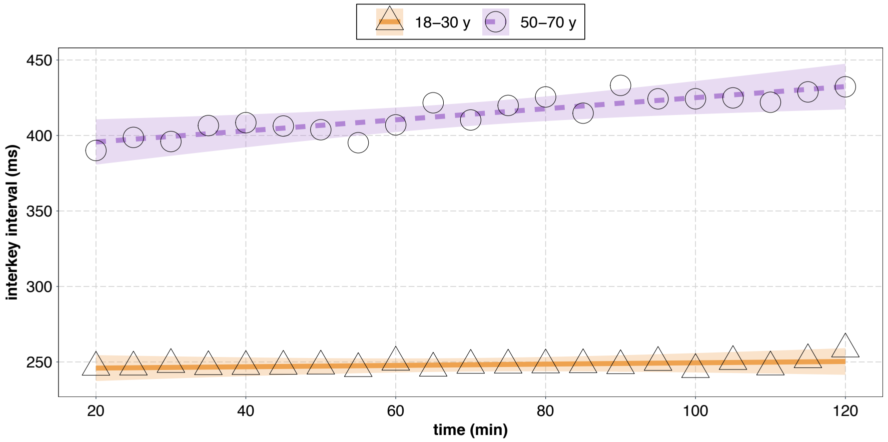
*: p<.05; **: p<.01; ***p<.001

Figure 2: *Mean typing time (A) and general interkey interval (B) during sentence typing as a function of time-on-task in the young and the middle-aged group.*



In the young group, typing speed during the last word of the sentence did not significantly change with time-on-task (typing time: $M_{t=20min}$=2.17s, $M_{t=120min}$=2.09s; general interkey intervallast word: $M_{t=20min}$=165ms, $M_{t=120min}$=167ms; stimulus-key interval: $M_{t=20min}$=244ms, $M_{t=120min}$=246ms; Table 3 & Figure 3). In the middle-aged group, however, typing speed during the last word of the sentence decreased during the 2-h experimental session (typing time$_{last\ word}$: $M_{t=20min}$=3.43s, $M_{t=120min}$=3.62s; general interkey interval$_{last\ word}$: $M_{t=20min}$=245ms, $M_{t=120min}$=265ms; stimulus-key interval: $M_{t=20min}$=397ms, $M_{t=120min}$=430ms; Table 3 & Figure 3). The middle-aged adults not only needed more time to type the last word of the sentence, they also initiated the first keypress later with increasing time-on-task.

Figure 3: *Mean stimulus–key interval as a function of time-on-task in the young and the middle-aged group.*



Typing accuracy

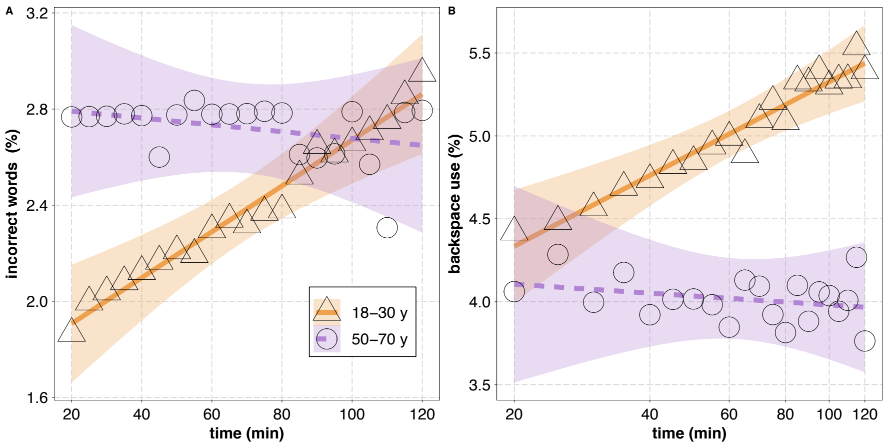
In the young group the percentage of words that contained typing errors (incorrect words) increased from 1.8% to 2.7% over the experimental session, while this increase in typing errors was not observed in the middle-aged group ($M_{t=20min/120min}$=2.5%) (Table 4 & Figure 4A). Note, however, that early during the session the number of incorrect words was higher in the middle-aged group than in the young group. In addition, the average percentage of words that were corrected (corrected words) did not change significantly throughout the experiment in both the young and middle-aged group ($M_{t=20min/120min}$=13.1%; Table 4). Although the effect of time-on-task on the percentage of corrected words in the middle-aged group did not reach the level of significance, we observed that the percentage of error-corrections was positively related to the time to complete a sentence in this age group ($r(21)=.53$, $p=.009$), indicating that in the middle aged participants correcting errors takes time as reflected in increased typing time. In the young group this relation was absent ($r(21)=.28$, $p=.200$). While in the younger adults the percentage corrected words did not change over the experiment, the percentage of keystrokes that were backspaces did increase with increasing time-on-task (young: $M_{t=20min}$=4.2%, $M_{t=120min}$=5.4%; Figure 4B). This increase in backspace use was due to an increase in backspace sequence length (i.e., backspace sequence length: $M_{t=20min}$=1.9 keys, $M_{t=120min}$=2.2 keys, Table 5; $r(21)=.69$, $p<.001$), indicating that young participants detected typing errors later with time-on-task. On average, backspace use did not change over the experiment in the middle-aged group ($M_{t=20min/120min}$=3.8%; Table 4 & Figure 4B). However, backspace sequence length decreased with prolonged task performance in this group ($M_{t=20min}$=2.2 keys, Mt=120min=1.9 keys, Table 5), indicating that middle-aged participants detected typing errors earlier with time-on-task.

Table 4: *The effects of time-on-task, age, and their interaction on the different dependent variables of accuracy described by the chi-square (χ2) and the β-weights of the final models that describe the effects of time-on-task, age, and/or their interaction on the dependent variables.*

Dependent variable	ToT χ2(1)	Age χ2(1)	ToT*Age χ2(1)	β_{ToT} (*SE*)	β_{Age} (*SE*)	$\beta_{ToT*Age}$ (*SE*)
Incorrect words	**4.39***	2.62	**6.16***	0.002 (*0.0007*)	0.24 (*0.15*)	-0.002 (*0.0010*)
Corrected words	1.28	1.57	1.64	-	-	-
Backspace use	1.88	**3.90***	**15.29*****	0.14 (*0.036*)	0.51 (*0.27*)	-0.20 (*0.052*)
Backspace sequence length	0.41	0.07	**7.29****	0.039 (*0.027*)	0.40 (*0.17*)	-0.10 (*0.037*)

*: p<.05; **: p<.01; ***p<.001

Figure 4: *Mean percentage of incorrect words (***A***) and mean backspace use (***B***) as a function of time-on-task in the young and the middle-aged group.*



Key sequence effects

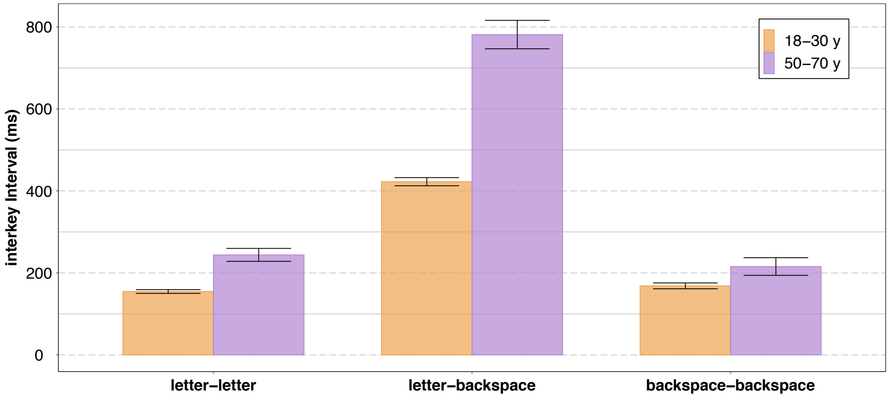
The increase in typing time with increasing time-on-task was reflected in an increase in general interkey interval in both age groups. However, the results showed that average time between two subsequent key presses was dependent on the specific sequence of keys. More specifically, in both age groups the initiation of an error-correction (i.e., letter-backspace interval) resulted in longer interkey intervals as compared to the letter-letter intervals (Table 5 & Figure 5). This effect was more pronounced in the middle-aged group than in the young group (Table 5 & Figure 5). In addition, in the young group typing speed during error-corrections (i.e. backspace-backspace intervals) was slower than during correct typewriting (i.e. letter-letter intervals). In the middle-aged group we observed the opposite effect, that is, typing speed during error-corrections was faster than during correct typewriting (Table 5 & Figure 5).

Table 5: *The effects of key sequence type (i.e., initiation of correction or during correction), age, and their interaction on the interkey interval type described by the chi-square (χ2) and the β-weights of the final models that describe the effects of time-on-task, age, and/or their interaction on the dependent variables.*

Dependent variable	**Type χ2(1)**	**Age χ2(1)**	**Type*Age χ2(1)**	**β_{Type} (*SE*)**	**β_{Age} (*SE*)**	**$\beta_{Type*Age}$ (*SE*)**
Letter-backspace interval	**219.8*****	**54.27*****	**15.52*****	1.00 (0.025)	0.42 (0.054)	0.17 (0.038)
Backspace-backspace interval	0.11	**45.23*****	**27.30*****	0.087 (0.022)	0.45 (0.042)	-0.20 (0.032)

*** p<.001

Figure 5: *Mean letter–letter-(Young: M = 154 ms, SE = 4.38; Middle-aged: M = 240 ms, SE = 16.41), letter–backspace-(Young: M = 420 ms, SE = 9.55; Middle-aged: M = 784 ms, SE = 35.7), and backspace–backspace interval (Young: M = 222 ms, SE = 21.7; Middle-aged: M = 170 ms, SE = 7.03) in the young and in the middle-aged group.*



These results showed that error-corrections resulted in slower typewriting in the younger adults, which suggests that the increase in general interkey interval in this group could, at least partly, be related to These results

These results showed that error-corrections resulted in slower typewriting in the younger adults, which suggests that the increase in general interkey interval in this group could, at least partly, be related to increased backspace use. This relation between backspace use and the interkey interval was further evidenced by the finding that the increase in backspace sequence length with time-on-task was correlated with an increase in general interkey interval in the young group (see accuracy). Since we established that typing speed during error-corrections is lower than typing speed during correct typewriting, these results suggest that the increase backspace sequence length in the young group result in decreased speed with time-on-task, reflected in increased general interkey intervals. Moreover, if we excluded typing speed during backspace use from the time-on-task analysis, the letter-letter interval during correct typewriting still decreased over the experiment in both age groups, however, this effect was now smaller in the young group (Young: $M_{t=20min}$=151ms, $M_{t=120min}$=154ms; Middle-aged: $M_{t=20min}$=232ms, $M_{t=120min}$=246ms; Table 6), thereby providing additional evidence for the supposition that the increase in backspace sequence length with time-on-task is, at least partly, responsible for the increase in general interkey intervals in the young group.

Table 6: *The effect of time-on-task, age, and their interaction on the different types of interkey interval described by the chi-square (χ2) and the β-weights of the final models that describe the effect of time-on-task, age, and their interaction on correct typewriting.*

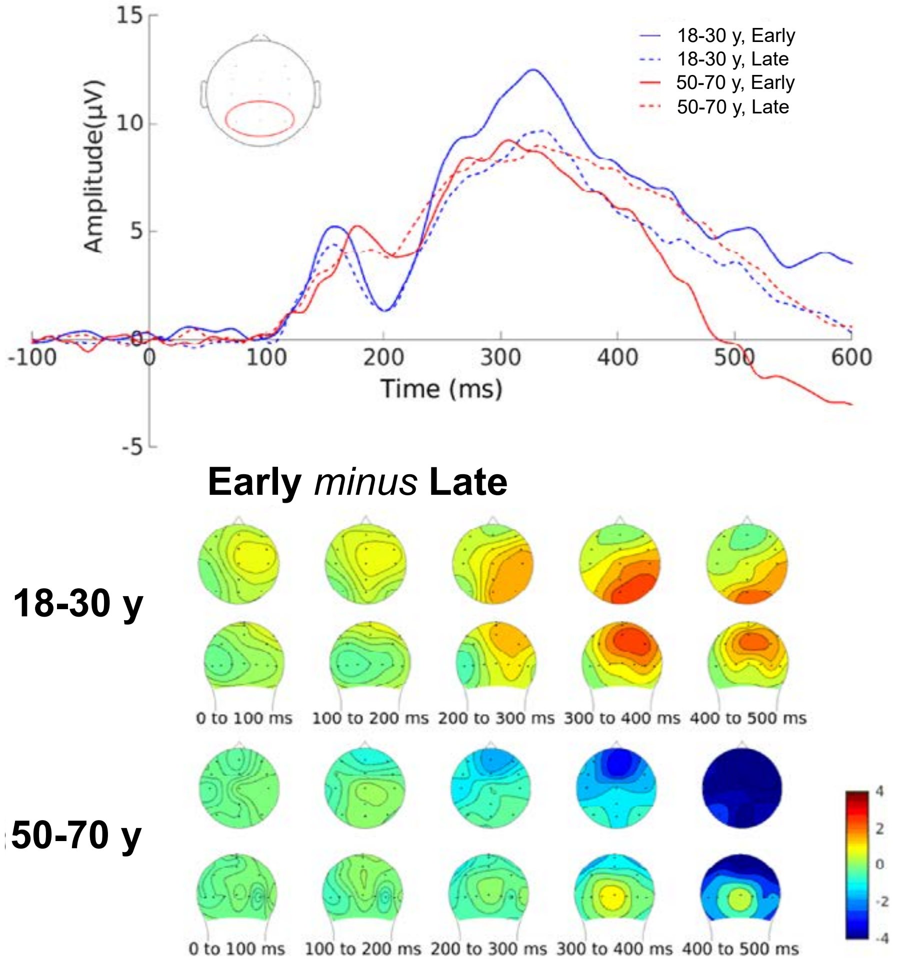
Dependent variable	ToT χ2(1)	Age χ2(1)	ToT* Age χ2(1)	β_{ToT} (*SE*)	β_{Age} (*SE*)	$\beta_{ToT*Age}$ (*SE*)
Letter-letter interval	**8.87****	**50.34 *****	**4.59***	0.0004 (0.00008)	0.43 (0.05)	0.0002 (0.0001)

*: p<.05; **: p<.01; *** p<.001

ERP measures

Early visual processes, reflected in the P1 and the N1 ERP components, were not affected by time-on-task in neither the young nor the middle-aged group (P1, ToT: $t(45)=1.15$, $p=.256$; ToT*Age: $F(1,44)=0.21$, $p=.651$; N1, ToT: $t(45)=-0.71$, $p=.481$; ToT*Age: $F(1,44)=0.01$, $p=.922$). With regard to P3 amplitude, we observed a decrease in P3 amplitude with time-on-task in the young group (ToT*Age: $F(1,44)=7.40$, $p=.009$; Young, ToT: $t(45)=-3.03$, $p=.006$; Figure 6), which was less pronounced in the middle-aged group (middle-aged, ToT: $t(45)=1.50$, $p=.139$). Fractional area latency of the P3 increased with time-on-task in the middle-aged group (ToT*Age: $F(1,44)=4.90$, $p=.032$; middle-aged, ToT: $t(45)=2.78$, $p=.011$), however, this effect was less pronounced in the young group (young, ToT: $t(45)=0.44$, $p=.662$).

Figure 6: *Effect of time-on-task on regression-derived ERP amplitude of the P3 component in young and middle-aged adults. Electrodes Cpz, Pz, P3, and P4 were used to model the ERP data. The modelled ERP early in the session describe the data 20 min after the start of the experiment. The modelled ERP late in the session describe the data 120 min after the start of the experiment.*



Deteriorations of performance in the young group (i.e., increased typing time, general interkey interval, and incorrect words) were accompanied by a decrease in the amplitude of the P3 (typing time: $r(21)=-.62$, $p=.002$; general interkey interval: $r(21)=-.66$, $p<.001$; incorrect words: $r(21)=-.52$, $p=.011$). An opposite trend was observed in the middle-aged group, that is, in middle-aged participants an increase in the amplitude of the P3 was associated with an increase in general interkey interval ($r(21)=.43$, $p=.043$). Decline in other keyboard-based performance measures did not correlate with changes in the P3 amplitude and/or latency in both the young and the middle-aged group.

Discussion

In the present study, we investigated whether and how the effects of mental fatigue are modulated by age, and if we could validate the use of measures based on typing behaviour as an index of the effects of mental fatigue on different aspects of cognition. We aimed to achieve this goal by examining the effects of mental fatigue, induced by prolonged typing performance, in young and middle-aged adults. In line with previous research, our results showed that participants experienced increased feelings of mental fatigue at the end of the experiment compared to the start of the experiment (e.g., Bonnefond et al., 2011; Faber et al., 2012). In addition to the increase of subjectively reported fatigue, we observed that different measures based on typing behaviour were influenced by 2 h of prolonged task performance, confirming that typing behaviour could be used as an index of mental fatigue. In both the young and the middle-aged group a decrease in typing speed was reflected in larger general interkey intervals (i.e., the time between two subsequent keypresses). Although typewriting slowed down in both the young and the middle-aged group as a result of mental fatigue (i.e., increase in typing time and larger general interkey intervals), mental fatigue only influenced accuracy in the young group (i.e., increase in incorrect words and in backspace use in order to correct wrong keypresses). These findings are consistent with previous studies showing that older adults are reluctant to make mistakes and act more error-aversive than younger adults (Starns & Ratcliff, 2010), while younger adults more often adopt a strategy that emphasize speed than older adults (Lucci et al., 2013; Rabbitt, 1979).

The first aim of the present study was to investigate the influence of age on the effects of mental fatigue. In general, the young participants were faster and performed more accurately than the middle-aged participants. Although typing skill has been found to be the largest predictor of differences in typewriting performance between age groups (Salthouse, 1984), the assessed level of skill of our participants (e.g., the ability to type according to the 10-finger system without looking at the keyboard) did not differ between age groups. Moreover, both groups worked a similar number of hours on a computer on a daily basis. Irrespective of these baseline differences, we observed a similar effect of mental fatigue on typing speed in both age groups. More specifically, the time that was needed to type the first part of a sentence increased over the experiment in both the young and the middle-aged groups, indicating that in both groups typing speed decreased as a result of mental fatigue. This effect was at least partly due to an increase in the time between two subsequent keypresses, suggesting that the general interkey interval could be used as an index of mental fatigue in both young and middle-adults.

After typing the first part of the sentence, the final word of the sentence was presented. This word could either be congruent or incongruent with the first part of the sentence. We found that the interval between the presentation of the last word and the first keystroke was longer for incongruent words than for congruent words in both age groups. Moreover, in accordance with results of the study of Hoeks and colleagues (2004), in both the young and in the middle-aged adults incongruent words were accompanied by a more negative deflection over central-posterior brain areas 400ms after stimulus presentation as compared with congruent words. This enlarged N400 indicates that both groups processed semantically conflicting information differently than semantically correct sentences (Kutas & Petten, 1988). In addition to the effects of congruency, we found that the time necessary to type the final word of the sentence increased with increasing time-on-task. The observed effects of time-on-task on stimulus-key interval, general interkey interval and typing time were most pronounced in the middle aged participants. In both age groups, mental fatigue did not influence conflict processing either on a behavioural or neural level, making it unlikely that the congruency manipulation contributed to these differences in typing speed between the young and the middle-aged group during the final word of the sentence. In line with Hess and Ennis (2011), one might argue that these differences are related to the amount of effort that middle-aged and young participants invested in the task, that is, middle-aged participants might have invested greater effort to successfully perform the task than the younger adults. This enhanced cognitive engagement might have resulted in higher costs associated with task performance in the middle-aged group (Cappell et al., 2010; Ennis et al., 2014). As a result typing speed during the final word of the sentence could have been more prone to mental fatigue in this group compared to the young group.

Additional to the effects of mental fatigue and age on typing speed, we observed effects of prolonged task performance on typing accuracy. We found that with time-on-task the young participants showed an increase in words containing typing errors which was correlated with a decrease in P3 amplitude, an ERP component that has been associated with attention and memory operations (Polich, 2007). These results are in line with findings of Hopstaken and colleagues (2015), who established a strong link between task disengagement, as measured by a decrease in P3 amplitude, and mental fatigue. As a result of fatigue-related diminished cognitive engagement in the young adults cognitive control processes requiring attention and memory seem to deteriorate, resulting in less efficient performance, as reflected in the increase in incorrectly typed words. Although participants were explicitly instructed to correct mistakes by using the backspace key, the number of error-corrections over the experiment remained stable, indicating that young participants did not compensate for reduced typing accuracy. Moreover, in the young group we observed an increase in the length of a backspace sequence with prolonged task performance, which indicates that young participants detected their mistakes later as they became fatigued. Despite the fact that middle-aged participants started out performing with lower accuracy than young participants, we did not find an additional decline in accuracy with increasing time-on-task in this group. More specifically, both the mistakes that were corrected and the mistakes that were not corrected by the backspace key remained stable over the experiment in the middle-aged group. In addition, we found a decrease in the backspace sequence length over the experiment in the middle-aged group, indicating that middle-aged participants detected their mistakes earlier with time-on-task. These results provide further support for the relation between error-aversive behaviour and age (Hertzog et al., 1993; Starns & Ratcliff, 2010).

In the present experiment, backspace use indirectly affected typing speed, as it was positively correlated with the number of letters that had to be typed, and therefore, with the time that was needed to complete a sentence. Besides this obvious effect of backspace use on typing speed, we observed that young participants slowed down before error-corrections (i.e., letter-backspace interval between the last key press before a backspace and the first backspace) and during error-corrections (i.e., backspace-backspace interval between backspaces), which is in line with findings of Kalfaoğlu and Stafford (2014), who also found a decrease in typing speed before and during error-corrections. Both these effects of backspace use on typing speed are important for the interpretation of the effects of mental fatigue on typing speed observed in the present experiment. Firstly, the young group, and not the middle-aged group, used the backspace key more often with time-on-task because they detected their mistakes later, and subsequently had to delete and re-type more letters, resulting in increased time necessary to type the sentence. Secondly, we observed decreased typing speed with time-on-task in the young group as a result of the increase in backspace sequence length, because typing speed during error-corrections is lower than during correct typewriting[1]. These findings suggest that backspace use, besides general slowing, is responsible for the decrease in typing speed during the experiment in the young group, while general slowing, by itself, seems to explain the decrease in typing speed with increasing time-on-task in the middle-aged group.

In sum, our results revealed that changes in typewriting performance with increasing time-on-task seemed to have a different cause dependent on age. Altogether, the effects of time-on-task on different measures of typing performance confirm that middle-aged adults are reluctant to make mistakes and act more error-aversive than younger adults (Starns & Ratcliff, 2010), while younger adults more often adopt a strategy that emphasize speed than older adults (Lucci et al., 2013; Rabbitt, 1979), and that these differential strategies are visible in the ecologically sensitive typing task.

The second aim of our study was to investigate if we could validate the use of measures based on typing behaviour as an index of the effects of mental fatigue on different aspects of cognition. Although the task challenged mental capabilities, typewriting also demands physical activity from the participant. Previous research showed that prolonged physical activity can result in physical fatigue. There are several reasons, however, why physical fatigue probably did not play a role in our experiment. Firstly, previous research showed that maximal voluntary contractions reach levels of only 6% to 20% during typewriting (Martin, 1996), which has been found to be too low to interfere with task performance (Lorist et al., 2002). Secondly, Kimura and colleagues (2007) established that muscle fatigue, measured by changes in the trapezius muscle, did not occur during 2 h of prolonged typing performance, except if 1kg weight was added to each wrist. Lastly, the typewriting task alternated with a mouse targeting task during which participants did not use the keyboard, resulting in decreased physical load as compared with continuous typing. Combining previous literature with the characteristics of our design, we can conclude that physical fatigue most likely did not play a role in our experiment. Therefore, we would suggest that different measures based on typewriting performance can provide information about the influence of mental fatigue.

---

[1] *The second result was supported by our finding that the effect of mental fatigue on the interkey interval decreased in the young group if we excluded error-corrections from the analysis.*

Brain activation, and specifically ERPs, can provide more insight in the cognitive processes underlying the described decline in typing performance as a result of mental fatigue. For example, the supposition that the general interkey interval could be used as an index of mental fatigue is supported by the relation between typing time, the general interkey interval, and brain activity in the young group. More specifically, we observed a decrease in the P3 amplitude over the experiment in the young group, which was related to an increase in typing time and longer general interkey interval. These effects were not related to a change in early visual processes (i.e., no change in P1 and N1 amplitude with time-on-task) in both age groups, indicating that sensory processes were not affected by mental fatigue, and suggesting that both groups were still sufficiently engaging in the task at the end of the experiment to be able to perceive visual information presented on the screen. Findings concerning the P3 amplitude were in line with Hopstaken an colleagues (2016), who observed a decrease in the P3 amplitude with time-on-task, which was not correlated with eye-movements away from task-relevant areas of the screen, suggesting that the decrease was related to lowered attention and memory processes involved in the further processing of task relevant stimuli. Besides a decrease in the P3 amplitude with time-on-task, Hopstaken et al. found an increase in the P3 amplitude if the benefits of finishing the task improved (i.e., increasing monetary rewards) at the end of the experiment, indicating that changes in the P3 amplitude are related to a cost-benefit evaluation of effort (Boksem & Tops, 2008; Hockey, 2011). They concluded that young adults are able to disengage from costly fatiguing tasks to prevent excessive use of cognitive resources, and save these resources for more rewarding situations.

Although the P3 amplitude showed a similar trend over time in the young as in the middle-aged group, the decrease did not reach significance in the middle-aged group. In contrast, the fractional area latency increased in the middle-aged group, while this effect was less pronounced in the young group. These differential effects could indicate that cognitive processes were affected differently by mental fatigue in the young and middle-aged group, that is, young adults might show more decline in the quality of processing, while processing speed declines more in middle-aged adults. Wascher and Getzmann (2014) also investigated the interaction between age and mental fatigue in age groups comparable to the groups we examined in the present experiment. In their study participants performed a visual spatial inhibition of return task. Wascher and Getzmann found similar results concerning the P3 amplitude: the P3 amplitude decreased over the experiment in the young group, however, not in the middle-aged group. In addition, they found that only middle-aged participants showed a decline in performance as a result of mental fatigue. Based on these findings they argued that adaptation mechanisms might differ between young and middle-aged adults. Behaviourally, our results showed a slightly different pattern, that is, young participants slowed down as a result of mental fatigue. However, in line with Wascher and Getzmann, different adaptation mechanisms were at play: speed was affected more in the middle-aged group, whereas accuracy only declined in the young group.

Irrespective of the effects of mental fatigue, the results of the present study also showed a relation between typing performance and the N400 component, reflecting conflict processing. As expected, typing time was affected by the processing of semantically conflicting information; participants slowed down when the text they typed was not fitting well with the context.

In sum, top-down cognitive control (e.g., attentional and working memory processes) is most susceptible to the effects of mental fatigue in both young and middle-aged adults. These results are in line with previous research that mental fatigue mainly has an effect on cognitive control processes (e.g., Faber et al., 2012; Lorist et al., 2005), requiring the investment of effort (Dehaene et al., 1998), whereas bottom-up processes, requiring less effort, are relatively unaffected by mental fatigue. Although both age groups show a decline in cognitive control, young adults seem to demonstrate a larger decline in the quality of processing, whereas middle-aged adults show a larger decline in processing speed. These effects directly relate to the behavioural patterns found in the present study, which illustrate that middle-aged adults are reluctant to make mistakes and act more error-aversive than younger adults (Starns & Ratcliff, 2010), while younger adults more often adopt a strategy that emphasizes speed than older adults.

Chapter 3

Dynamics in typewriting performance reflect mental fatigue during real-life office work

This chapter is based on:

de Jong, M
Bonvanie, A.M.
Jolij, J.
Lorist, M.M. (2020)

Dynamics in typewriting performance reflect mental fatigue during real-life office work. PLOS ONE 15(10): e0239984. https://doi.org/10.1371/journal.pone.0

Abstract

Mental fatigue has repeatedly been associated with decline in task performance in controlled situations, such as the lab, and in less controlled settings, such as the working environment. Given that a large number of factors can influence the course of mental fatigue, it is challenging to objectively and unobtrusively monitor mental fatigue on the work floor. We aimed to provide a proof of principle of a method to monitor mental fatigue in an uncontrolled office environment, and to study how typewriting dynamics change over different time-scales (i.e., time-on-task, time-of-day, day-of-week). To investigate this, typewriting performance of university employees was recorded for 6 consecutive weeks, allowing not only to examine performance speed, but also providing a natural setting to study error correction. We show that markers derived from typewriting are susceptible to changes in behaviour related to mental fatigue. In the morning, workers first maintain typing speed during prolonged task performance, which resulted in an increased number of typing errors they had to correct. During the day, they seemed to readjust this strategy, reflected in a decline in both typing speed and accuracy. Additionally, we found that on Mondays and Fridays, workers adopted a strategy that favoured typing speed, while on the other days of the week typing accuracy was higher. Although workers are allowed to take breaks, mental fatigue builds up during the day. Day-to-day patterns show no increase in mental fatigue over days, indicating that office workers are able to recover from work-related demands after a working day.

Keywords: *Work environment, behaviour, time-on-task, time-of-day, day-of-week*

Introduction

In order to be able to interact with the dynamically changing world around us, we continuously adapt our behaviour. During prolonged task performance, however, this adaptation is often insufficient to counter for the increasing demands placed on our information processing system. This typically is reflected as a decline in task performance over time: people perform more slowly, make more mistakes, and are less able to correct for these mistakes (Boksem et al., 2006; Lorist et al., 2005). This decline in performance, commonly known as mental fatigue, occurs in many settings, and might not only have implications for productivity, but also for the safety of employees and their environment (McCormick et al., 2012; Ricci et al., 2007). More specifically, long working hours and experiencing work-related fatigue have shown to be predictors of health complaints and absence due to sickness (Sluiter et al., 2003; Sparks et al., 1997). Additionally, there is an increased risk of fatigue-related accidents when employees engage in traffic after a long day of work, thereby endangering not only themselves but also others (Gander et al., 2007). Considering the impact of mental fatigue in the working environment, it is important to conduct research on how to detect, deal with or even prevent the effects of mental fatigue.

Ever since the beginning of the 19th century researchers have studied mental fatigue and its effects on performance in controlled experimental settings and in real-life situations, such as the workplace (Bartley & Chute, 1949; Hockey, 2013; Muscio, 1921; Thorndike, 1900). Several theories regarding the cognitive mechanisms behind its manifestation have been proposed, eventually leading to the widely accepted theory suggesting that mental fatigue develops as a result of a cost-benefit evaluation of effort (Boksem & Tops, 2008; Hockey, 2011). According to this theory, if the costs of performing a task exceed the benefits of finishing the task, people will come to experience subjective feelings of mental fatigue (e.g., aversion against task performance, low vigilance) and performance deteriorates (i.e., people become slower and less accurate).

Kreapelin was the first to attempt to quantify the course of mental fatigue during task performance. It soon became clear, however, that there was no such thing as a typical decline in performance over time. Mental fatigue and its effects on behaviour depend on several personal and environmental factors. For example, people are able to overcome the effects of mental fatigue if they are sufficiently motivated when they receive a monetary reward based on their performance (Boksem et al., 2005) or if they drink a cup of (caffeinated) coffee (Lorist et al., 1994a). Given that it is hard to define a specific course of mental fatigue over time, and employees themselves are poor at detecting when they are not capable of performing a task at an adequate level anymore (Zhang et al., 2011), it is challenging to effectively monitor and prevent mental fatigue in the working environment. In order to detect this decline in performance, it is necessary to continuously monitor behaviour dynamics without interference of work.

Developments in information technology, however, have made it possible to monitor behaviour in novel ways, without interfering with regular work activities. For example, Pimenta and colleagues (2014) developed a method for non-invasive measurement of mental fatigue by monitoring a very common behaviour for office workers: typewriting performance. They found that several markers of typing performance were susceptible to the effects of time-of-day. To validate whether changes in these markers were due to mental fatigue specifically, Jong and colleagues (2018) conducted an experiment in which brain activity using electroencephalography (EEG) was recorded during a 2-hour typewriting task. They were specifically interested in the P3 brain potential of which the amplitude has been known to decrease with increasing mental fatigue (Hopstaken et al., 2015; Lorist & Jolij, 2012; Wascher & Getzmann, 2014). The study showed that both typing speed, reflected in the time between two subsequent keypresses (interkey interval), and typing accuracy, reflected in overall backspace use and incorrectly typed words, declined with prolonged task performance. Moreover, these deteriorations in typewriting performance with time-on-task correlated with neural makers signalling mental fatigue, indicating that monitoring typewriting markers can provide information about the level of mental fatigue, at least in a controlled setting.

Although changes in typewriting have been found to reflect mental fatigue under these standardized conditions, there are many other variables that could influence behaviour dynamics under less controlled conditions. For instance, at the workplace, where deteriorations in task performance are particularly problematic, the effects of time-of-day and day-of-week have found to influence performance, as well. While time-on-task effects have mostly been studied in experimental settings, studies concerning the effects of time-of-day and day-of-week are generally performed in real-life settings, investigating self-reports. A study of Linder and colleagues (2014), for example, showed that clinicians prescribe unnecessary antibiotics more often in the afternoon as compared to the beginning of the day. Similar effects have been found over the different days of the week, where employees have been found to feel more energized after the weekend, resulting in better reported performance at the beginning compared to the end of the week (Binnewies et al., 2010). Although these effects work on different time-scales, they all resulted in changes in (self-reported) performance levels. Moreover, continuously performing a task, and engaging in work for multiple hours or days, requires rest to restore performance to its former level (Kühnel et al., 2017). More specifically, time-on-task effects can be reversed by taking a short (coffee) break (Steinborn & Huestegge, 2016), time-of-day effects can be reversed by a night's rest (Hooff et al., 2007), and day-of-week effects can be reversed by a weekend break. Although there is substantial evidence that prolonged task performance, manipulated by time-on-task, time-of-day, and day-of-week, separately influence performance, interestingly, it is not yet known how these factors interact and subsequently influence behaviour dynamics.

Previous experimental studies on mental fatigue mainly focused on the effects of time-on-task on behavioural performance, investigating isolated effects of prolonged task performance on specific cognitive processes (e.g., error processing; Lorist et al., 2005). In addition, studies in real-life settings focused on specific professions, especially those involving shift-work (Brown et al., 2020; Kecklund & Axelsson, 2016), where the manifestation of fatigue was expected to be potent or even dangerous, given its relationship with serious accidents (Chan, 2011; Swaen et al., 2003). There seems to have been done little research on the manifestation of mental fatigue during regular 9 to 5 jobs.

Present study

In order to gain more insight in the manner in which behavioural dynamics in the workplace are influenced by time-on-task, time-of-day and day-of-week, we first focused on validating a potentially useful method to study mental fatigue on the work floor without interfering with regular working activities. To this end, markers in typewriting that were found to be sensitive to mental fatigue in a lab setting (i.e. interkey interval and backspace use) were recorded for six consecutive weeks during regular office work (de Jong et al., 2018). Second, we investigated the influence of mental fatigue on these markers at different time-scales (i.e., time-on-task, time-of-day, and day-of-week).

In line with findings in an experimental setting (de Jong et al., 2018), we hypothesized that there would be a main effect of time-on-task on both the interkey interval and the percentage of backspaces, where we expected that both measures would increase with time-on-task. Secondly, we hypothesized that the magnitude of the effect of time-on-task on these performance measures would depend on time-of-day (main effect). That is, we expected a larger increase in both the interkey interval and the percentage of backspaces with time-on-task in the afternoon (interaction) than in the morning. Lastly, we hypothesized that typewriting patterns would change over the course of the week. We expected these changes to manifest in two ways. First, we hypothesized that employees became slower and less accurate over the week (main effect), and second, we expected a larger decline of performance (interkey interval and backspace use) with time-on-task over the week (interaction).

Materials and methods

Participants

Forty-five office workers gave their written informed consent to participate in a study that was approved by the Ethics Committee of the Faculty of Economics and Business in Groningen. This research complied with the tenets of the Declaration of Helsinki. Participants were employees of the Faculty of Economics and Business of the University of Groningen and were recruited via the health and safety coordinator of the faculty. They were included if they worked for at least 0.8 full-time equivalent (32 h a week) and typewriting activities were part of their work. Only datasets which contained more than 30 subsets of more than 45 min of continuous typing were included in order to perform reliable statistical analyses. From now on we will refer to these subsets as tasks. As a result, data of 23 employees was excluded from the analysis, leaving data of 22 employees (12 females, *M*=48.1 year, *SD*=13.4). There was variation in function profile across participants that were included in the study (i.e., scientific staff, support staff). Participants that were excluded from the analyses performed working activities during the measurement period that did not include the required amount of typing activities (e.g., teaching and collecting research data), which was specifically the case for Ph.D. students and (Postdoctoral) researchers. In addition, a number of participants worked on multiple workstations during the 6 weeks of data collection, which was reflected in a limited amount of typewriting data that was recorded from these participants at the workstation on which the recording software was installed. Data of these participants were excluded from the analyses, as well.

Apparatus and Materials

The experiment was conducted in the natural working environment of the participants at the faculty of Economics and Business of the University of Groningen. The experimental setup consisted of an office chair behind an adjustable desk, a windows computer with a QWERTY keyboard, and screen support. The working environment was adjusted according to the occupational health and safety guidelines of the faculty. The percentage of backspaces and the interkey interval was acquired using keylogging software (aXtion).

Typing performance

Previous research of de Jong et al. (2018), found backspace use and the interkey interval to be susceptible to the effects of mental fatigue in a controlled lab setting. In order to monitor these typing indices, keylogging software, installed on the workstations, registered a timestamp at the start of each keystroke. To safeguard the confidentiality of the typed text during the study, only the backspace key was given a unique marker. Each minute, the average interkey interval (the time between two subsequent keystrokes) and the percentage of backspaces of the preceding 15 min was calculated and registered for offline analysis. If the time between two subsequent keystrokes was longer than 5 s, the interkey interval was not included in the average. A series of average values was included in subsequent analysis if more than 45 successive averages were recorded. In the present study, continuous typewriting was defined as typewriting during a block of at least 45 minutes.

Procedure

Typing performance was monitored for 6 weeks in the natural working environment of the participants. Data collection of the first cohort started on the first Monday of May and the second cohort started on the first Monday of November. A week before the start of the monitoring period, the keylogging software was installed on the computers of the participants and the office environment was confirmed to be or adjusted according to occupational health and safety guidelines of the faculty. During this week, participants also filled out a questionnaire with demographic and work-related questions (Appendix II). Each Monday, starting in the second week of the experiment, participants filled out a questionnaire with general questions about how they experienced the week before (Appendix III). Each working day, participants received real-time feedback on their performance provided via text messages on their mobile phones and via email. An overview was provided via email at the end of the day.

Statistical analysis

Statistical analysis was conducted in R version 3.4.4 (R.Studio Team, 2016; R Development Core Team, 2017). For statistical significance testing, we used a mixed-modelling approach using the lme4 package version 1.1-21(Bates, Mächler, et al., 2015). The package lmertest version 3.0-1 was used to obtain statistical significance by approximating the degrees of freedom using the Satterthwaite approximation (Kuznetsova et al., 2017). The data provided to the models included the interkey interval and the percentage of backspaces. The models contained a varying intercept per participant. In addition, a varying slope for time-on-task and time-of-day by subject was added to the model if the fit of the model improved as indicated by the Akaike Information Criterion (AIC) (Bozdogan, 1987). The models used to statistically test the effects of time-on-task (120 min of continuous typewriting), time-of-day (morning and afternoon), and day-of-week (Monday, Tuesday, Wednesday, Thursday, and Friday) on the dependent typewriting variables (i.e., interkey interval and percentage of backspaces) are listed in Table 1.

Table 1. *Models used to statistically test the effects of time-on-task, time-of-day, and day-of week on the interkey interval and the percentage of backspace keystrokes.*

Dependent variable	Equation
Interkey interval$_n$ and Backspace use$_n$	$\beta_{0.j} + \beta_1 \mathit{time\ on\ task}_n + \beta_2 \mathit{time\ of\ day}_n + \beta_3 \mathit{day\ of\ week}_n + \beta_5 \mathit{time\ on\ task}_n \times \mathit{time\ of\ day}_n + \beta_6 \mathit{time\ on\ task}_n \times \mathit{day\ of\ week}_n + \beta_7 \mathit{time\ of\ day}_n \times \mathit{day\ of\ week}_n + \beta_8 \mathit{time\ on\ task}_n \times \mathit{time\ of\ day}_n \times \mathit{day\ of\ week}_n + \epsilon_n$

n reflects a time-block (minute) and *j* reflects a participant. β_0 reflects the intercept of the model, $\beta_{1\text{-}8}$ reflect the regression coefficients, and ϵ reflects the error term. The notation for these models allowed for a varying intercept per participant (as indicated by *j*).

Post-hoc tests were performed to assess the main and interaction effects, adjusting error rates according to Bonferroni. First, to estimate the difference between the effect of time-on-task between the morning and the afternoon, polynomial contrasts were compared using pairwise comparisons. Second, pairwise comparisons were administered to compare the interkey interval and backspace use on the different days of the week. Finally, polynomial contrasts were used to estimate the linear and quadratic trends with time-on-task in the morning and the afternoon, and over the different days of the week. Statistical tests were considered significant at $p < .05$.

Speed and Accuracy

In order to investigate the relationship between typing speed (interkey interval) and typing accuracy (backspace use) during continuous typewriting, we calculated the regression coefficients that described the effect of time-on-task on the dependent variables in the morning and in the afternoon for each participant. For these personalized regression coefficients, we calculated Pearson's correlations to identify whether changes in typing speed and accuracy were related.

Results

In order to systematically discuss the results, we first report the effects of time-on-task on the interkey interval, reflecting typing speed, and backspace use, reflecting accuracy. Thereafter, we go into the effects of time-of-day and the interaction of time-of-day with time-on-task on the same measures. Lastly, we report how these typewriting patterns change over the different days of the week. The models that were used to statistically test the effects of prolonged task performance on the different time-scales (i.e., time-of-day, time-of-day, and day-of-week) can be found in Table 1. An overview of the main and interaction effects is provided in Table 2 and Table 3, respectively.

Table 2: *The main and interaction effects of time-on-task, time-of-day, and day-of week on the interkey interval.*

Main and interaction effects	***F*-value**	***Dfn***	***Dfd***	***p*-value**
Time-on-task	17.75	2	94707	<.001
Time-of-day	1.07	1	94719	.302
Day-of-week	74.15	4	94707	<.001
Time-on-task × time-of-day	6.90	2	94703	.001
Time-on-task × day-of-week	47.57	8	94701	<.001
Time-of-day × day-of-week	54.69	4	94701	<.001
Time-on-task × time-of-day × day-of-week	72.08	8	94701	<.001

Table 3: *The effects of time-on-task, time-of-day and day-of week and their interaction on the percentage of backspaces.*

Main and interaction effects	***F*-value**	***Dfn***	***Dfd***	***p*-value**
Time-on-task	284.28	2	91647	<.001
Time-of-day	36.55	1	92656	<.001
Day-of-week	53.37	4	91637	<.001
Time-on-task × time-of-day	12.93	2	91634	<.001
Time-on-task × day-of-week	21.85	8	91633	<.001
Time-of-day × day-of-week	36.93	4	91635	<.001
Time-on-task × time-of-day × day-of-week	9.54	8	91633	<.001

Time-on-task

The results showed that both the interkey interval ($F(2, 94707) = 17.75$, $p<.001$) and the percentage of backspace keystrokes ($F(2, 91637) = 284.28$, $p<.001$) changed with prolonged task performance (i.e., subset of > 45 minutes of continuous typewriting). That is, in general, we observed an increase in both the interkey interval and the percentage of backspaces, reflecting a decrease in typing speed and a decline in typing accuracy with time-on-task. However, as expected, these effects were modulated by time-of-day and day-of-week. These modulations will be discussed below.

Time-of-day

Although mean interkey interval (main effect **time-of-day**: $F(1, 94719) = 1.07$, n.s.) did not differ between the morning and the afternoon, the effect of **time-on-task** on the interkey interval was modulated by **time-of-day** (interaction effect **time-on-task** × **time-of-day**: *F*(2, 94703)=6.90, *p*=.001; **afternoon**

minus **morning**$_{linear}$: *z*=3.29, *p*=.006; **afternoon** *minus* **morning**$_{quadratic}$: *z*=-3.51, *p*=.002). That is, post-hoc tests revealed that the interkey interval remained stable during continuous typewriting in the morning, but in general it increased with 11.6 ms during two hours of continuous task performance in the afternoon (see Table 4 and Fig 1A).

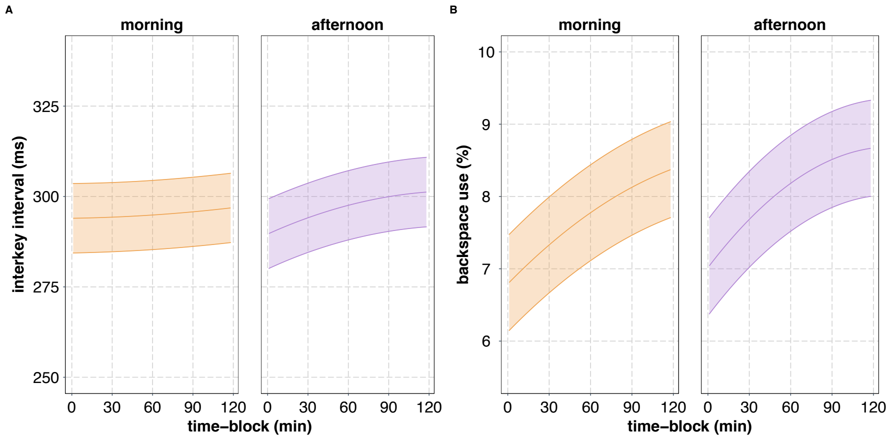
Table 4: *The effect of time-on-task on the interkey interval and backspace use in the morning and the afternoon. Mchange reflects the average change in the dependent variable from the 1st to the 120th minute of continuous typewriting.*

Dependent variable	**Time-of-day**	**Polynomial**	***z*-value**	**M$_{change}$ (SE)**
Interkey interval (ms)	Morning	Linear	-1.92	2.94 (1.53)
		Quadratic	1.73	
	Afternoon	Linear	**5.42*****	11.51 (2.12)
		Quadratic	**-3.07***	
Backspace use (% of backspace keystrokes)	Morning	Linear	**19.74*****	1.57 (0.08)
		Quadratic	**-11.11*****	
	Afternoon	Linear	**14.99*****	1.63 (0.11)
		Quadratic	**-9.84*****	

Bolded values are significant; *∗p<.05; ∗∗p<.01; ∗∗∗p <.001*

Backspace use increased from the morning to the afternoon (main effect **time-of-day**: *F*(1, 91656)= 36.55, *p*<.001). Additionally, the effect of time-on-task on backspace use differed between the morning and the afternoon (interaction effect **time-on-task** × **time-of-day**: *F*(2, 93467)=9.54, *p*<.001). That is, although the percentage of backspace keystrokes increased with ~1.6% during two hours of prolonged task performance, both in the morning and in the afternoon (**afternoon** minus **morning**$_{linear}$: *z*=-0.46, *p*=1.0), the increase followed a more quadratic function in the afternoon compared to the morning (**afternoon** *minus* **morning**$_{quadratic}$: *z*=-3.95, *p*<0.001; see Table 2 and Fig 1B).

Figure 1: *The effect of time-on-task on typewriting changes from the morning to the afternoon. Time-blocks are calculated based on the preceding 15 min, see method section. (***A***) the interaction between time-on-task and time-of-day on the average interkey interval. (***B***) The interaction between time-on-task and time-of-day on backspace use. The confidence intervals reflect the standard errors of the mean.*



Day-of-week

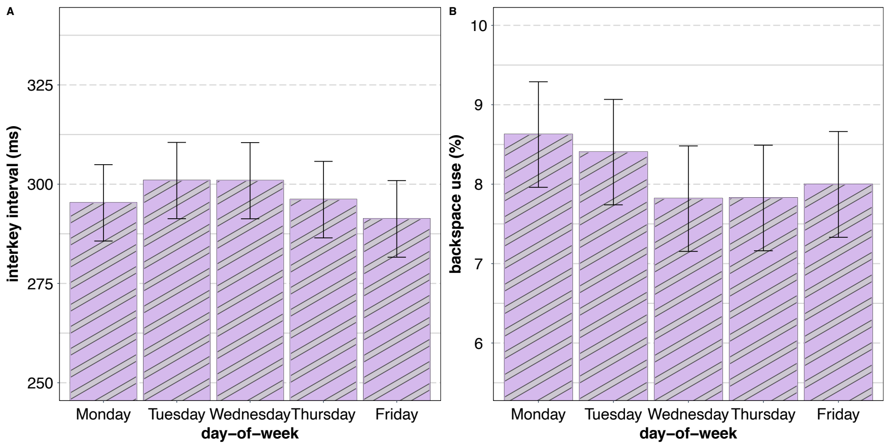
In addition to the effects of time-on-task and time-of-day, we also looked into changes in typewriting patterns over the workweek. First, we hypothesized that typing performance would decline over the workweek, reflected in an increase in the interkey interval and the percentage of backspaces. Contrary to our expectations, we observed an increase in the interkey interval from 295ms on Monday to 301ms on both Tuesday (Mon-Tue: z=-3.93, p<.001) and Wednesday (Mon-Wed: z=-3.99, p<.001), after which the interkey interval decreased to 296ms on Thursday and 291ms on Friday (z_{linear}=-6.09, p<.001), during which employees' typewriting was fastest (see Table 5 and Fig 2A).

Backspace use was highest on Monday with 8.6% of the keystrokes were backspace keystrokes, followed by Tuesday with 8.4% compared to the other days of the week (Mon-mean(Wed, Thu, Fri): z=11.39, p<.001; Tue-mean(Wed, Thu, Fri): z=8.91, p<.001; see Table 5 and Fig 2B). Backspace use did not significantly differ between Wednesday, Thursday and Friday (Wed-Thu: z=-0.11, p=1.0; Wed-Fri: z=2.11, p=.351; Thu-Fri: z=1.88, p=.604).

Table 5: *Average typing performance on the different days of the week, reflected by the interkey interval (ms) and backspace use (% of backspace keystrokes).*

Day-of-week	**Mean interkey interval in ms (*SE*)**	**Mean percentage of backspace keystrokes (*SE*)**
Monday	295 (9.62)	8.62 (0.67)
Tuesday	301 (9.60)	8.40 (0.66)
Wednesday	301 (9.60)	7.82 (0.66)
Thursday	296 (9.63)	7.83 (0.66)
Friday	291 (9.65)	8.00 (0.67)

Figure 2: *The course of typewriting performance over the different days of the week.* (**A**) *The effect of day-of-week on the interkey interval.* (**B**) *The effect of day-of-week on the percentage of backspace keystrokes. The confidence intervals reflect the standard errors of the mean.*



Additionally, we observed that the effects of time-on-task on typing speed differed over the days of the week (see Table 2 for an overview of the main and interaction effects). That is, on Monday afternoon (z=-7.10, p<.001), and Friday morning (z=-8.33, p<.001) the interkey interval decreased with time-on-task. On the other days of the week, the interkey interval either remained stable or increased, the last one reflecting a decrease in typing speed with time-on-task.

We also observed changes in the effect of prolonged task performance on backspace use over the working week (see Table 3 for an overview of the main and interaction effects). On all days, except for Friday afternoon, the percentage of backspace keystrokes increased with time-on-task in the morning and in the afternoon. On Friday in the afternoon no change in backspace use was observed (z=0.82, *n.s.*).

Speed and accuracy

In order to investigate the relationship between typing speed (interkey interval) and typing accuracy (backspace use) between the morning and the afternoon, we calculated the effect of time-on-task in the morning and in the afternoon. The correlations between these coefficients revealed that the relation between speed and accuracy differed between the morning and the afternoon. In the morning, we observed no correlation between the effect of time-on-task on typing speed and the effect of time-on-task on the percentage of backspace keystrokes (r=-0.04, *n.s.*). In the afternoon, however, there was a positive relationship between the increase in the interkey interval and the increase in the percentage of backspaces. More specifically, participants that showed a larger increase in the interkey interval with time-on-task also showed a larger increase in the percentage of backspace keystrokes with time-on-task (r=0.622, p<.001), indicating that changes in typewriting with time-on-task do not reflect changes in speed-accuracy trade-off.

Discussion

In the current study we evaluated novel, non-invasive measures based on typewriting to continuously monitor behaviour in a working environment. Our aims were, first, to provide a proof of principle of this method, and, second, to study how typewriting dynamics during regular office work change over different time-scales (i.e., time-on-task, time-of-day, day-of-week). Based on earlier findings observed in a controlled environment, we focused on interkey interval and backspace use as indices of behaviour. To investigate these aims, the typewriting markers were recorded for six consecutive weeks during regular office work performed in a university environment. We confirmed that typewriting behaviour contains sensitive markers that reflect changes in behaviour over time. In addition to general changes in speed and accuracy with time-on-task, we found that the effects of time-on-task as indexed by our typewriting measures changed throughout the day. More specifically, the effect of time-on-task on typing speed (i.e., interkey interval) was more pronounced in the afternoon than in the morning. Moreover, on average, office workers used the backspace key more often in the afternoon compared to the morning, although the effect of time-on-task on backspace use, reflecting task accuracy, was smaller in the afternoon. Finally, an analysis of time of week effects provides no evidence for a general decline in performance over the week.

With regard to our first aim, as hypothesized, the length of the interkey interval and the percentage of backspace keystrokes both increased with time-on-task, replicating previous work suggesting that changes in markers derived from typewriting are sensitive to mental fatigue elicited during continuous task performance (de Jong et al., 2018; Pimenta et al., 2013). Previously, this type of research was mainly conducted in simulated office environments (Hockey & Earle, 2006), or focused on self-reported behaviour of employees (Smolders et al., 2012), using measures that either interrupted regular office work or relied on subjective measures influenced by the observer's personal judgment (Norbert, 1999). Our findings show that our measures based on typing behaviour have practical potential to objectively monitor performance efficiency without disturbing regular work-related activities.

A similar pattern of results was observed during simulated office work (de Jong et al., 2018) as in the present, real-life environment. Moreover, the changes with time-on-task in typewriting performance were even found to be more pronounced in the present study. Relevant in this perspective is that in the present study, compared to a relatively controlled experimental environment, many uncontrollable factors may have influenced performance efficiency due to the dynamic nature of the actual office environment (Davis, 1946; Smith-Jackson & Klein, 2009). On the one hand, factors such as interruptions related to the presence of others (Galván et al., 2013) and uncontrollable requests for actions from electronic devices (e.g., online activity, telephone calls), might increase task demands (Steege et al., 2015), which could in turn increase the effects of mental fatigue on performance efficiency. However, on the other hand, work motivation (Boksem et al., 2006) and enhanced autonomy with regard to setting one's own schedule and planning work-breaks if needed (Hockey & Earle, 2006), might reduce experienced task demands and related levels of mental fatigue during regular office work compared to the lab setting. Interestingly, despite these noisy conditions, we observed significant changes in the typewriting indices during prolonged task performance. To summarize, with regard to our first aim, we provided a proof of principle of the sensitivity of these measures, confirming that typewriting markers are susceptible to changes in behaviour related to the effects of mental fatigue, not only in a controlled experimental setting, but also in an uncontrolled office environment.

Under real life conditions, factors that influence our behaviour vary every day and even from hour to hour, and therefore a substantial variability in performance and the effects of mental fatigue might be expected over time. Our second aim was to investigate how typewriting dynamics during regular office work changed over different time-scales. In general, performance efficiency during real life activities, such as typewriting, depends on two dimensions: speed and accuracy. In the present study, the interkey interval served as an indicator of typing speed and the percentage of backspace keystrokes was used as an indicator of typing accuracy (de Jong et al., 2018). Using the backspace key is an indirect measure of typing accuracy, given that it is used to correct mistakes in typewriting. Therefore, while interpreting the results, it is important to keep in mind that an increase in backspace use could originate from different types of behaviour. That is, in our study, participants could have corrected more (in)correctly typed letters and/or detected their errors later, which, as a result, required more consecutive backspaces to correct one incorrect keystroke.

The results of the present study showed that, in the morning, typing speed remained relatively stable over time. Simultaneously, typing accuracy declined, which was revealed by an increase in backspace use. In the afternoon, we observed a decline in both dimensions of typing performance. More specifically, typing speed decreased over time, reflected by an increase in the interkey interval, and additionally the quality of typing was reduced, indicated by the increase in the percentage of backspace keystrokes with prolonged task performance. This pattern shows similarities with previous research investigating the effects of mental fatigue on task control in a lab setting (Lorist et al., 2000). Lorist and colleagues showed that if participants were instructed to perform fast, accuracy steadily declined from the start of the experiment, while participants kept responding at a stable speed. After a while, participants seemed to adjust their strategy. That is, over time, participants performed at a slower pace as well, which was observed in an increase in RTs.

In daily life, people adoptively invoke qualitatively different performance strategies. Adopting a strategy that focuses on speed generally results in a larger number of errors, while adopting a strategy that focuses on high accuracy results in slower performance (Wickens & Hollands, 2000). People tend to moderate this speed-accuracy trade-off based on external conditions and the time available to complete their work. The results suggested that, in the present study, office workers first tried to maintain typing speed, which resulted in an increased number of typing errors they had to correct. In the afternoon, however, they seemed to readjust their strategy, which resulted in a decline in both dimensions of performance. In comparison to the study of Lorist and colleagues (2000), who measured prolonged performance during a 2h session, the pattern we found stretched out over the day, indicating that mental fatigue might build up during a working day. These findings imply that, although office workers are entitled to have breaks during a working day, the scheduled breaks might not have been enough to fully recover from the demands they encountered during the day.

Previous research on the effects of mental fatigue on typewriting strategies showed that people tend to make more typing errors during prolonged task performance (de Jong et al., 2018). However, when given the option to correct their mistakes, they at least partly correct these mistakes. Although several studies observed a similar increase in correction behaviour during prolonged task performance, people are not able to correct for the total increase in mistakes, even if these mistakes are in plain sight, that is, clearly visible on the screen in front of them (Boksem et al., 2006; Lorist et al., 2005). In order to safeguard the privacy of the employees, however, we did not monitor the identity of the keys apart from the backspace keys. For this reason, we were not able to identify and analyse the mistakes that were not corrected during typewriting.

Error-corrections indirectly reflect accuracy on a given task, and using error-corrections during typewriting as a measure of mental fatigue might therefore provide information on the effects of mental fatigue on underlying cognitive processes. Monitoring performance requires higher-order mental functions, which are prone to the effects of mental fatigue. Experimental studies showed that erroneous responses are usually followed by specific brain activation patterns, called the error-related negativity (ERN) (Falkenstein et al., 1991; Gehring et al., 1993), and result in decreased response speed on the next trial (i.e. post-error slowing) (Rabbitt, 1979). Lorist and colleagues (2005) investigated these behavioural and brain activity patterns during prolonged performance on an Eriksen flanker task. They found that performance monitoring declined over time, which was reflected in a significant decrease of brain activity patterns related to error processing (ERN), and was accompanied by a decrease of post-error slowing. This decreased ability to monitor behaviour and adapt performance concurrently might have resulted in a later detection of errors and therefore in an increase in error-corrections during the present study, as was also shown in the controlled lab study of de Jong and colleagues (2018).

The present study focused on the dynamics in typewriting during prolonged task performance. Previous research repeatedly showed that, in addition to a tonic decline in speed and accuracy over time, fatigued participants also experience short-term lapses in performance during which they are unable to process any information (Bertelson & Joffe, 1963; Bills, 1931). These phasic lapses in performance, so-called mental blocks, are characterised by extremely long reaction times during experimental tasks and can be detected by studying the distribution of reaction times (Steinborn & Huestegge, 2016). In the present study, we excluded lapses of attention by excluding interkey intervals that were longer than 5 seconds. However, for future research it would be interesting to investigate whether the effects of prolonged task performance on length or number of short-term lapses in performance during prolonged typewriting follow a similar pattern as the tonic effects of prolonged task performance on typewriting.

Besides the effects of mental fatigue on typewriting dynamics during the day, the present study also provides direct insight into typewriting dynamics during a working week. First, we found no evidence for a general decline in typewriting performance with day-of-week, given that backspace use remains stable on Wednesday, Thursday and Friday, and the interkey interval decreases, reflecting an increase in typing speed, from Wednesday to Friday. Second, we found that the effects of prolonged task performance on typing speed and accuracy followed a similar pattern over the different days of the week, suggesting that mental fatigue elicited on the previous day, as reflected in the effects of time-on-task and time-of-day, did not influence the course of performance during prolonged task performance on the next day. These results provide proof that mental fatigue does not accumulate across the days of the week. Although previous literature does not paint a consistent picture, Persson and colleagues (2006) indicated that alertness of construction workers did not increase during a working week. This pattern was shown in construction workers with a regular working schedule (7-15 h), but also in workers with an extended schedule (six days in a row, one day off, five days in a row, nine days off) that stayed at accommodations at the construction site during the working week.

Our findings provide further evidence for office workers' ability to recover from work-related demands during the week. Nonetheless, typewriting dynamics were subject to daily variations. That is, on Mondays and Fridays, office workers adopted a typewriting strategy that maximized typing speed, while on the other days of the week they either adopted a strategy that maximized accuracy (Wednesdays), or performed both fast and accurately compared to the other days of the week (Thursdays). Although some studies showed that employees recover from regular work-related demands after engaging in pleasurable activities during the weekend (Binnewies et al., 2010; Fritz & Sonnentag, 2005), other studies revealed that Mondays serve as a transition day from pleasurable activities to the structured demanding work week, which is reflected in a more negative mood (Croft & Walker, 2001), increased stress levels (Devereux et al., 2011), and decreased ability to recover from work-related demands on Mondays (Rook & Zijlstra, 2006). These factors might have also led to the observed behavioural patterns in the present study. Similarly, it could be argued that Fridays also serves as a transition day from the work week to the weekend. In contrast to Mondays, however, Fridays have previously been associated with improved mood compared to the rest of the week (Reis et al., 2000).

This study has implications for real-life working environments, given that a large part of the working population regularly performs computer work. In the Netherlands, for example, 40% of the employees perform computer work more than 6 h every day (Hooftman et al., 2019). There are several ways in which monitoring typewriting could support employees during their work. First, personalized real-time feedback based on changes in typing behaviour could be provided to the users in order to help them detect when lapses in performance occur and a short break might be beneficial. However, real-time feedback might be biased due to dynamics in typewriting performance that are not related to lapses in performance. One of the characteristics of our working environment is the large variability in working conditions, due to changes in work-related tasks, noise in the working environment, and changes in general persons state, among others. Our method also allows monitoring performance over a longer period of time enabling us to detect regularities in working activities. Related to this, a second possibility of our method is to provide feedback on an individual level to help employees realize a more optimal work-break schedule that is complementary with their individual state and specific work-related demands. By comparing behaviour dynamics over several weeks, typing behaviour could help decide when, during the workday or -week an employee should work on tasks that need high accuracy or when it is better to work on less demanding tasks. A third option is to use changes in typing behaviour to evaluate interventions in the working environment. For instance, it might provide relevant information with regard to performance efficiency for evaluating the effectiveness of a 6-hour workday instead of our regular 8-hour workday. Previously, researchers already used questionnaires to evaluate this specific intervention (Akerstedt et al., 2001), however, measuring performance, and importantly, doing so without interrupting regular activities, could enhance our knowledge of its effects on performance and productivity more objectively.

Conclusions

The typing indices that were used to describe behaviour dynamics reflect subtitle changes in both speed and accuracy during regular office work, not only during the day but also over the week. These findings might be relevant to consider when scheduling different tasks over the day, but could also provide information about the number of hours that employees can or should work during a day.

Chapter 4

Caffeine boosts preparatory attention for reward-related stimulus information

This chapter is based on:

de Jong, M. †
van den Berg, B. †
Woldorff, M. G.
Lorist, M. M. (2021)

† These authors have contributed equally to this work

Caffeine Boosts Preparatory Attention for Reward-related Stimulus Information. Journal of cognitive neuroscience, 33(1), 104–118. https://doi.org/10.1162/jocn_a_01630

Abstract

The intake of caffeine and the prospect of reward have both been associated with increased arousal, enhanced attention, and improved behavioural performance on cognitive tasks, but how they interact to exert these effects is not well understood. To investigate this question, we had participants engage in a two-session cued-reward cognitive task while recording their electrical brain activity using scalp electroencephalography. The cue indicated whether monetary reward could be received for fast and accurate responses to a colour-word Stroop stimulus that followed. Before each session, participants ingested decaffeinated coffee with either caffeine (3 mg/kg bodyweight) or placebo (3 mg/kg bodyweight lactose). The behavioural results showed that both caffeine and reward-prospect improved response accuracy and speed. In the brain, reward-prospect resulted in an enlarged contingent negative variation (CNV) and reduced posterior alpha power (indicating increased cortical activity) prior to stimulus presentation, both neural markers for preparatory attention. Moreover, the CNV enhancement for reward-prospect trials was considerably more pronounced in the caffeine condition as compared to the placebo condition. These interactive neural enhancements due to caffeine and reward-prospect were mainly visible in preparatory attention activity triggered by the cue (CNV). In addition, some interactive neural enhancements in the processing of the Stroop target stimulus that followed were also observed. The results suggest that caffeine facilitates the neural processes underlying attentional preparation and stimulus processing, especially for task-relevant information.

Keywords: contingent negative variation, neural arousal, caffeine, reward, oscillatory alpha, preparatory attention, behavioural relevance

Introduction

Humans have always searched for ways to enhance their cognitive capacities in order to efficiently cope with the overload of information they encounter in everyday life. One of the most useful set of mechanisms humans have available to distinguish between relevant and irrelevant information is our attention system (Petersen & Posner, 2012). In particular, this system selectively guides attention towards environmental stimuli and events that are associated with the possibility of gaining reward (Aarts et al., 2011; Engelmann & Pessoa, 2007; Hickey et al., 2010; Hickey & van Zoest, 2012; Roelfsema et al., 2010; Van Den Berg et al., 2019). In addition, caffeine-containing substances (e.g., coffee, tea), arguably the most widely used psychoactive cognitive enhancer in the world, are being used on a daily basis to boost this attention system (Lorist & Tops, 2003; Saville et al., 2018; Wilhelmus et al., 2017).

The enhancing effects of reward-prospect and caffeine on attention have both been linked to their putative influence on the dopaminergic systems of the brain. More specifically, the prospect of reward has been found to increase dopamine release directly and acutely (Schultz, 2000, 2015). Caffeine, on the other hand, influences the central nervous system mainly through antagonizing adenosine receptors (Dunwiddie & Masino, 2001; Fredholm et al., 1999). Adenosine A1 receptors are found throughout both cortical and subcortical brain areas. Adenosine A2A receptors, on the other hand, have been found to be specifically present in dopamine-rich areas of the brain, with highest levels at postsynaptic neurons in the striatum where they are co-localized with dopamine D2 receptors (Ferré, 2008; Fredholm et al., 2001; Gevaerd et al., 2001). The dosage of caffeine in one cup (200 ml) of coffee (typically regarded as 90 mg [European Food Safety Authority]) has been shown to work as an antagonist at both A1 and A2A receptors, and to facilitate the signalling of dopamine in the brain (Davis et al., 2003; Ferré, 2016; Okada et al., 1996; Quarta et al., 2004), particularly in the striatum, which is an essential part of the reward system (Schultz, 2000, 2015). This inverse relationship between adenosine and dopamine on the receptor level might underlie findings showing non-specific adenosine antagonists, such as caffeine, can increase dopamine levels (Ferré et al., 2001; Ferré, 2008; Ferré et al., 2018)

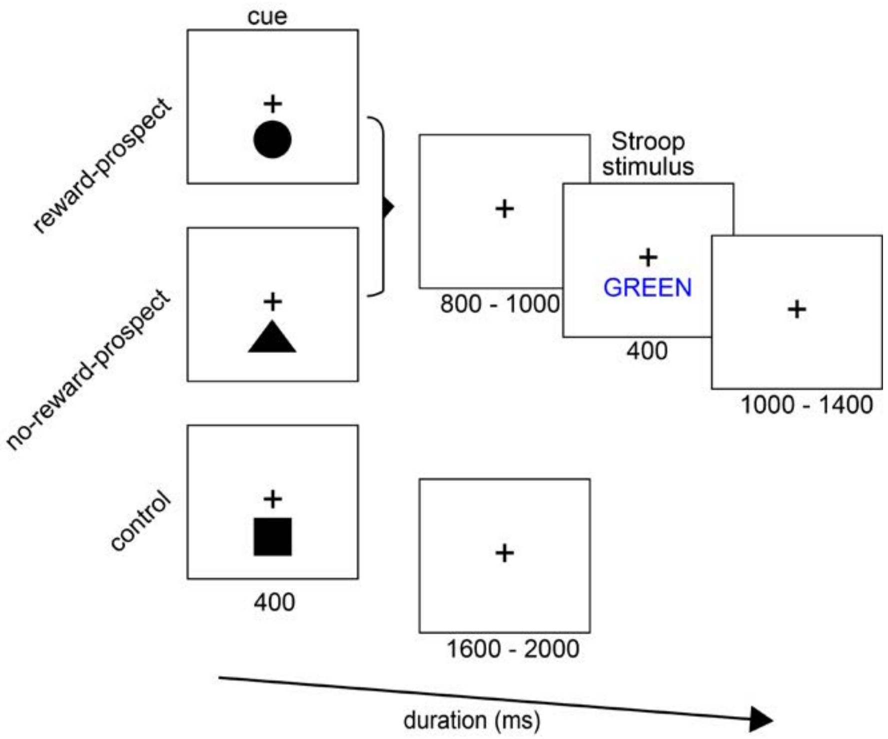
Both the prospect of gaining reward and the consumption of caffeine-containing substances have been found to improve behavioural performance. For instance, cueing of the prospect of gaining monetary rewards has been found to guide attention selectively towards the potentially rewarding events or stimuli (Schevernels et al., 2014; van den Berg et al., 2014). Neurally, the prospect of gaining reward on an upcoming visual task triggers increased activity in cortical regions involved in attentional control and improved processing of relevant information in the visual cortices on that task, thereby improving behavioural performance (Hickey et al., 2010; Roelfsema et al., 2010; van den Berg et al., 2019). Similarly, caffeine, compared to placebo, has also been found to increase neural activation in cortical areas involved in the selection of relevant information. Specifically, caffeine biases the processing of relevant stimuli over irrelevant ones, by modulating task-related brain activation related to the processing of relevant- and/or irrelevant stimuli (Kenemans & Lorist, 1995; Lorist et al., 1994; Lorist et al., 1995; Ruijter et al., 2000). In addition to these specific effects of caffeine on attention, several studies have also showed a broad sustained increase in neural arousal, after the intake of caffeine (vs. placebo), consistent with the more general stimulating effects of caffeine on behaviour (Kenemans & Lorist, 1995). Consistent with these neural modulations, caffeine doses as low as the dose in half a cup of coffee speed up reaction times (RTs) and improve accuracy (Lieberman et al., 1987).

Although both reward-prospect and caffeine intake have substantial beneficial effects on behavioural performance, the nature of their interaction has remained elusive. Given that both reward-prospect and caffeine enhance the dopaminergic system, one might expect interactive properties in terms of cortical modulation between tonic modulation of caffeine and acute modulations by reward-prospect. There is substantial evidence, however, that both reward-prospect and caffeine influence behaviour through an effect on attentional preparation. In studies employing high spatial resolution fMRI and ones using high temporal resolution EEG recordings, attentional preparation elicited brain activity in regions of a fronto-parietal attentional control network (Corbetta & Shulman, 2002). This set of brain regions is thought to be the main contributor to the fronto-centrally distributed contingent negative variation (CNV) (Grent-'t-Jong & Woldorff, 2007), a negative-polarity slow-wave event-related potential (ERP) that is elicited before an imperative stimulus is presented (Brunia et al., 2012; Walter et al., 1964). Importantly, research has suggested a link between dopamine levels and the processes underlying the CNV (Linssen et al., 2011). The amplitude of the CNV has further been found to predict behavioural performance (Hillyard, 1969; van den Berg et al., 2014, 2016). Although all aspects of the cognitive functions reflected by the CNV are not fully understood, it is generally thought that a larger CNV indicates increased arousal or stronger anticipation of expected upcoming task-relevant events or stimuli (Brunia et al., 2012; Hillyard, 1969; van den Berg et al., 2014, 2016).

Another neural marker that has been associated with preparatory attentional processes is oscillatory activity in the alpha-band frequency range (8 – 14Hz). Decreases of the power in this frequency band has been related to increased cortical activity and enhanced selective attention (Scheeringa et al., 2012; Worden et al., 2000), indicative of heightened information sensitivity in visual regions (Jensen & Mazaheri, 2010). In addition, power in the alpha frequency band has been shown to predict behavioural performance (van den Berg et al., 2016; van Dijk et al., 2008), and has previously been found to be modulated by reward-prospect as well as caffeine (Ashton et al., 1974; Kenemans & Lorist, 1995; Tieges et al., 2007; van den Berg et al., 2014).

The aim of the present study was to examine whether and how the prospect of reward and the consumption of caffeine-containing substances, two factors that both enhance attention in everyday life, interact on the behavioural and/or neural level. To investigate this interaction, participants participated in a two-session study in which they either received coffee with caffeine or with lactose (placebo) prior to the experimental session, in which they performed an adapted version of the cued-reward task of van den Berg and colleagues (2014). In this task (**Figure 1**), participants were instructed to respond as fast and accurately as possible to target Stroop stimuli. At the beginning of each trial, participants were presented with a cue that indicated that there was prospect of receiving a reward on that trial or there was no such prospect. During reward-prospect trials participants could earn money if they responded accurately and sufficiently fast.

Figure 1: Task and stimuli. Each trial started with a cue, consisting of a symbol that indicated reward-prospect, no-reward-prospect, or a control cue indicating that no Stroop stimulus would appear on that trial. Following a fixation interval, the Stroop stimulus would appear after the first two cue types. In case of the control cue, no Stroop stimulus follows, but instead the next trail would begin after a variable delay.



Based on previous findings, our first main hypothesis was that both reward-prospect and caffeine would improve behavioural performance. Neurally, we hypothesized they would both enhance preparatory cortical activity, as indexed by the CNV and alpha power. In addition, we had two main competing hypothesis in terms of potential interactions of these factors. On the one hand, previous research has indicated that caffeine specifically improves selective attention towards relevant information (Lorist et al., 1994; Lorist et al., 1995; Ruijter, Lorist, et al., 2000). Based on these studies, one might expect that caffeine would lead to enhanced attentional preparation for more important stimuli (i.e., in reward-prospect trials) as compared to less important ones (i.e., in no-reward-prospect trials). On the other hand, there is evidence that the stimulating effects of caffeine are most pronounced in situations when attentional control of perceptual functions is reduced, such as in the presence of mental fatigue or a lack of motivation (Koelega, 1993; Lorist et al., 1994; Ruijter et al., 1999; Weiss & Laties, 1962). Based on these findings, an alternative hypothesis is that the effects of caffeine would be most pronounced in the no-reward-prospect condition compared to the reward-prospect one, where the attentional system is already triggered by the anticipation of reward. Finally, we inspected the effect of the factors caffeine and reward on the processing of the Stroop stimulus as indicated by the late positive complex (LPC), a component that indicates the level of processing of target stimulus information (Kappenman & Luck, 2010; van den Berg et al., 2014). Because the LPC has generally been found to be related to response speed, for the LPC we expected the amplitude to parallel the effects of reward and caffeine on behavioural performance.

Methods

Participants

Thirty-one healthy adults (10 males), ranging in age from 18 to 31 (*M* = 22.0 year, *SD* = 3.6), participated in the two-session experiment. Participants received either course credits or 7 euros per hour for participation. In addition, they received a monetary reward that depended on their performance (*M* = 10.6 euros, *SD* = 3.9). All participants were native Dutch speakers, right-handed, regular coffee-drinkers who ingested a minimum of 2 cups per day (*M* = 4.2 cups/day, *SD* = 1.7). We did not further explore relationships between potential neural and behavioural effect of caffeine and reward and habitual caffeine intake. They had a regular sleep schedule and had normal or corrected-to-normal-visual acuity. Participants indicated that they were not lactose-intolerant and that they did not smoke. Data from two participants were excluded due to technical issues during recording, while data from another three were excluded due to excessive noise in the EEG (i.e., > 30% of the EEG epochs rejected due to artefacts, see EEG pre-processing below). The experiment was approved by the Ethics Committee of the Psychology Department of the University of Groningen, and participants gave their written informed consent before the start of the first experimental session.

Apparatus

The experiment was conducted in a sound- and light-attenuated room with the stimuli being presented on a 100Hz LCD monitor with a resolution of 1920 × 1080 (Iiyama ProLite G2773HS). Participants sat in a comfortable desk chair at a viewing distance of 70 cm from the monitor and gave behavioural responses using a gamepad with four bumper buttons, using the index and middle finger of their left and right hands (Logitech Rumblepad, http://www.logitech.com/). The experimental task was programmed using the Presentation software package (version 18.1

06.09.15, http://www.neurobs.com/). Stimuli were randomized using the R statistical programming software package (R Development Core Team, 2013).

Task and Stimuli

During the entire task, a central fixation cross was continually visible in the middle of the screen. At the start of each trial (**Figure 1**), a cue stimulus (circle, triangle, or square [visual angle: 1.23°]) was presented, 1 cm below the fixation cross, for 400 ms. This cue stimulus indicated whether the trial was a reward-prospect trial (40 % of the trials), in which a monetary reward would be given if a response to the subsequent imperative stimulus (a Stroop stimulus) was both correct and met a pre-defined response time (RT) criterion (see below), or whether was a no-reward-prospect trial (40 %) or a control trial (20 %), in which no imperative Stroop stimulus followed the cue. The meaning-shape mapping of the cue stimuli was counterbalanced across participants. Following a fixation screen presented for 800 to 1000 ms, in the reward-prospect and no-reward-prospect trials, a Stroop colour-word stimulus (i.e., Dutch words for “RED”, “GREEN”, “BLUE”, and “YELLOW” [visual angle: 1.23° by 4.91°]) was presented below the fixation mark for 400 ms. Participants were instructed to indicate the font-colour of the Stroop stimulus fast and accurate. In half of the trials, the Stroop stimulus was congruent (i.e., word meaning matched the font colour) and in the other half the stimuli were incongruent (i.e., word meaning did not match the font colour). The interval between the offset of the Stroop stimulus and the next cue varied randomly between 1000 and 1400 ms. The interval between the control cue (i.e., the one indicating that no Stroop target stimulus would follow) and the cue for the next trial was varied randomly between 1600 and 2000 ms.

Participants were instructed to respond to the font-colour of the Stroop stimulus as fast and as accurately as possible, by pressing one of the four buttons corresponding to the font-colour on the gamepad. After receiving instructions, the participants first performed a practice block of 30 trials, in which they received positive feedback (‘correct’) if their response was correct and faster than 900 ms, or negative feedback (‘incorrect’) if the response was incorrect or slower than 900 ms. If participants did not achieve a hit-rate of over 80 %, they performed a second practice block of 30 trials; otherwise the experimental task started.

After the practice trials, participants performed a block of 30 trials, which was used to calculate the RT criterion (RT_{crit} = mean RT + 200 ms) in the reward-prospect condition. The number of points participants could earn was based on the RT on each individual trial (RTi in ms) according to the formula: RT_{crit} - RTi. These points were converted to euros (i.e., 3000 points represented 1 euro). Participants did not receive a penalty if they responded too slowly or incorrectly, and hence they could only gain money and not lose any.

Thereafter subjects performed 6 experimental blocks of 200 trials each. After each block participants could take a self-timed break. Within each block there were 5 s breaks every 15 trials and a 30 s break every 100 trials. After every 30 trials a screen was presented for 2000 ms that provided feedback, namely indicating the total amount of money made thus far.

Procedure

The experiment consisted of two sessions that were scheduled exactly one week apart. Both sessions started at 9:00 a.m. and each took approximately 3.5 hours. Participants were instructed to abstain from alcohol and caffeine-containing substances for at least 12 h before each session. Participants were told that they would receive a cup of coffee at the start of an experimental session. Approximately 45 min before the start of the task, participants received a cup of decaffeinated coffee, to which either caffeine or lactose (both 3 mg/kg bodyweight [bodyweight was reported by the participant]) had been added. The experimenter was blind to whether caffeine or lactose had been added. In addition, participants were not informed that the coffee could contain either caffeine or lactose to avoid potential anticipation effects (Mills et al., 2017). The order of both conditions was counterbalanced over sessions across participants.

EEG recording and data analysis

EEG was recorded using a 64-channel ANT waveguard electrode cap (10-10 system), using an online average reference. The sampling rate was 512 Hz and the data were filtered during recording using a FIR filter with a corner frequency at 102 Hz (0.2 x sampling rate). Additional electrodes were placed on both the left and right mastoids, and vertical and horizontal EOG activity was recorded from two electrodes placed above and below the right eye and from two electrodes placed lateral to the outer canthi of the two eyes, respectively. Electrode impedances were kept below 5 kΩ. The analyses were performed using custom Matlab scripts (MATLAB - Release 2015b) in combination with the EEG analysis toolboxes Fieldtrip (Oostenveld et al., 2011) and EEGlab (Delorme & Makeig, 2004).

Data were offline re-referenced to the algebraic average of the mastoid electrodes. Channels that contained excessive noise were replaced by interpolated values of the surrounding electrodes (spherical spline interpolation). Eye blinks and eye movements were corrected using independent component analysis (ICA) to reconstruct the data excluding those components that reflected eye blinks. The data were filtered using a 0.01 Hz high pass filter. Epochs were extracted from -1500 to 2500 ms surrounding the onset of the cue stimulus and from -1500 to 2500 ms surrounding the onset of the Stroop stimulus. Epochs containing any remaining artefacts (amplitude > 150 µV, -500 to 1500 ms surrounding cue and stimulus onset) were excluded from the analysis (average epochs rejected per subject per session, cue epochs rejected: *M* = 7.2 % (*SD* = 7.6 %); target epochs rejected: *M* = 5.9 %, (*SD* = 6.3 %)).

After artefact rejection, the mean number of epochs per condition for the cue was as followed: on average each session consisted of 455 (*SD* = 41) reward-prospect trials, 453 (*SD* = 39) no-reward-prospect trials, and 219 (*SD* = 21) control trials. The mean number of epochs per condition for the target was as followed: on average each session consisted of 461 (*SD* = 39) reward-prospect trials, and 458 (*SD* = 37) no-reward-prospect trials.

Event Related Potentials

To statistically examine cue-evoked brain activation, the mean amplitude of the CNV was derived from a fronto-central region of interest (ROIfc: FCz, FC1, FC2, Cz, C1 and C2), consistent with previous literature (Brunia et al., 2012), measured in the 700 – 1100 ms interval after cue presentation for every trial. Stimulus-evoked brain activation was examined in two ROIs to investigate the 'early' reward related frontal positivity (van den Berg et al., 2014), and the influence of reward and caffeine on the late positive complex, while at the same time considering the effect of Stroop stimulus conflict on electrical brain activity. The first ROI consisted of fronto-central channels (ROIfc: FCz, FC1, FC2, Cz, C1 and C2), in which brain activity was examined between 400 and 500 ms post-stimulus. Note that this ROI and latency interval is consistent with the location and latency range in which the incongruency related negativity is usually found; a central negative deflection (incongruent vs. congruent) that that has been related to the processing of stimulus conflict (Liotti et al., 2000; van den Berg et al., 2014; West & Alain, 2000; West & Bell, 1997). The second ROI, defined to investigate the LPC, consisted of parietal channels (ROIp: Pz, P3, P4, POz, PO3 and PO4), from which the ERP signal was extracted between

700 and 800 ms post-stimulus. Note that the second interval overlapped with the later stage of conflict processing; a larger positive deflection for incongruent compared to congruent stimuli. As such, these ROIs allow to investigate the influence of reward and caffeine on the processing of the Stroop stimulus, while taking into account the different neural process that are evoked by stimulus conflict.

Power Analysis

We used two different approaches to estimate oscillatory power in the EEG data surrounding the cue. Although we were primarily interested in alpha band (8-14Hz) activity, we estimated power for a wider range to ensure that potential effects were specific for the alpha frequency band. First, to obtain sustained/tonic changes in power, which are expected as a result of caffeine-induced changes in cortical arousal, the frequency contents of cue-locked epochs were analysed. The data were first multiplied by a Hanning window and subsequent power estimates were obtained using a Fourier transformation. As these epochs ranged from -1500ms until 2500ms post-cue, the resulting activity contained pre-cue activity, cue-evoked activity, and stimulus evoked activity. We measured alpha power (8-14Hz) (log10 transformed power) on each epoch in an occipital ROI (ROIo: PO7, PO8, PO3, PO4, O1, O2).

Second, to examine changes in power specifically related to cue-evoked activity, time-frequency decomposition was done by multiplying the cue-epoched data with a moving (steps of 50ms) Hanning window. The temporal width of the Hanning window decreased with increasing frequency (7 cycles per frequency from 5 to 40Hz; for 5Hz [200ms per cycle], the width of the window is 1400ms). Finally, obtained power was Log10-transformed. No baseline correction was performed for the power analyses. To assess the effects of reward and caffeine on alpha activation, we measured alpha power (8-14Hz) (log10 transformed power) between

700 and 1100ms after cue presentation on each trial in an occipital ROI (ROIo: PO7, PO8, PO3, PO4, O1, O2), consistent with previous studies (van den Berg et al., 2014; Worden et al., 2000).

Statistical analysis

For statistical significance testing, we employed a mixed-modelling approach using the lme4 statistical package (Bates, Mächler, et al., 2015). The data provided to the model included the response time, and a mean amplitude of the various neural measurements on each trial (excluding trials market as artefacts, for the cue interval, the data consisted of ~62400 observations [2 sessions × 1200 trials × 26 subjects]. For the target interval [excluding trials in which the participant responded incorrectly to the Stroop stimulus] this was ~41000 observations). A mixed modelling approach has the advantage of taking into consideration the individual data points (as opposed to binning trials into various conditions per subject as, for example, in standard GLM-ANOVA) and therefore allows to account for variance that is related to between session effects, introduced in the present study by counterbalancing the session order of administration of caffeine across participants. In other words, because we counterbalanced the session order of administration of caffeine across participants, a mixed modelling allows for the accounting of session related variance. To establish the random-effects structure we used a procedure in which we started with a full model (containing random slopes per subject for all corresponding fixed effects; including the interaction terms) and we subsequently reduced model complexity by stepwise removing random factors (starting with the interaction terms) until the model was not singular (Bates, Kliegl, et al., 2015). This stepwise procedure has been shown to result in a 'hybrid' model that avoids overfitting the data while containing relevant random effects to control for the Type 1 error rate (Matuschek et al., 2017).

The formulas in **Table 1** reflect the resulting models used to statistically test the effects of congruency, caffeine, and reward (and their interactions) on the behavioural and neural dependent variables. For all models (behavioural-, cue- and target interval models), we used a random intercept for each subject and a random slope per subject for session and reward-prospect. For the behavioural and target interval models a random slope by subject for congruency was added. To obtain information about statistical significance, the degrees of freedom were approximated using the Satterthwaite approximation of effective degrees of freedom as calculated by the R package "lmertest", which has been show to control relatively well for Type-1 error rate (Kuznetsova et al., 2017; Luke, 2017). In addition to the mixed-model single-trial statistical analysis we provide GLM-ANOVA tables, based on the means per condition per subject (collapsed over session), as Supplementary Materials (Appendix IV). We also incorporated the alternative GLM-ANOVA approach proposed by Kenemans et al. (1999) to examine session effects in Appendix IV, in which order (i.e., caffeine session followed by placebo session, or vice versa) was added as a between-subject factor. Finally, we tested if the additional effect of order improved the mixed model fit. The results presented in Table (**S2**) show that the addition of this factor did not substantially improve the fit of any of the models tested, providing support for exclusion of this factor.

Table1: Models used to statistically test the effects of congruency, caffeine, and reward-prospect on the behavioural and neural dependent variables. Within the formulas the subscript n indicates single trials, the subscript j indicates the individual subject. For instance, the notation β0.j indicates the intercept, which varies by subject (the random effect).

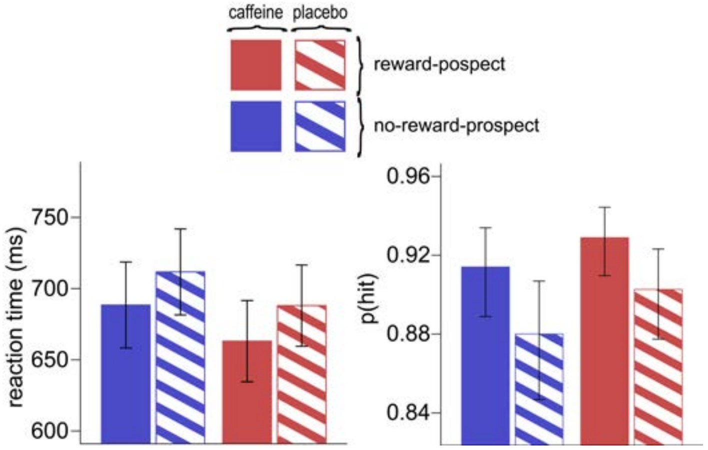
Behavioural analysis	$p(hit)_n = logit^{-1}(\beta_{0.j} + \beta_1 congruency_{j[n]} + \beta_2 reward_{j[n]} + \beta_3 caffeine_n + \beta_4 session_{j[n]} + \beta_5 congruency_n \times reward_n + \beta_6 congruency_n \times caffeine_n + \beta_7 reward \times congruency_n + \beta_8 congruency \times reward \times caffeine_n + \epsilon_n)$
	$RTs_n = \beta_{0.j} + \beta_1 congruency_{j[n]} + \beta_2 reward_{j[n]} + \beta_3 caffeine_n + \beta_4 session_{j[n]} + \beta_5 congruency_n \times reward_n + \beta_6 congruency_n \times caffeine_n + \beta_7 reward_n \times caffeine_n + \beta_8 congruency_n \times reward_n \times caffeine_n + \epsilon_n$
Cue interval analysis	$Marker_n = \beta_{0.j[n]} + \beta_1 reward_{j[n]} + \beta_2 caffeine_n + \beta_3 session_{j[n]} + \beta_4 reward \times caffeine_n + \epsilon_n$
Target interval analysis	$\mu V_n = \beta_{0.j} + \beta_1 congruency_{j[n]} + \beta_2 reward_{j[n]} + \beta_3 caffeine_n + \beta_4 session_{j[n]} + \beta_5 congruency_n \times reward_n + \beta_6 congruency_n \times caffeine_n + \beta_7 reward_n \times caffeine_n + \beta_8 congruency_n \times reward_n \times caffeine_n + \epsilon_n$
Sustained alpha power	$alpha\ power_n = \beta_{0.j} + \beta_1 caffeine_n + \beta_2 session_{j[n]} + \epsilon_n$

Results

Behavioural performance

Participants responded on average 58 ms (*SE* = 5 ms) more slowly and 2.9 % (*SE* = 0.5 %) less accurately to incongruent than to congruent Stroop stimuli, replicating a multitude of studies of the behavioural effects of Stroop incongruency (MacLeod, 1991)(main effect of congruency; RT: $F(1,25) = 192.6$, $p < .001$; accuracy: $\chi2(1) = 72.7$, $p < .001$). In addition, they responded more quickly (~25ms) and more accurately to the font-colour of the Stroop stimulus if the cue for that trial indicated reward-prospect as compared to the no-reward-prospect (main effect of reward; RT: $F(1,25) = 42.3$, $p < .001$; accuracy: $\chi2(1)=12.9$, $p < .001$; Figure 2). Moreover, we found that RTs decreased (~23ms) and the proportion of hits increased in the caffeine condition compared to the placebo condition (main effect caffeine; RT: $F(1,24) = 7.1$, $p = .014$; accuracy: $\chi2(1) = 13.3$, $p < .001$; Figure 2). On the behavioural level, no significant interactions were observed between the independent variables of congruency, reward, and caffeine.

Figure 2: *Behavioural results. Participants responded faster and more accurately when cued with reward-prospect vs. no-reward-prospect and when receiving caffeine vs. placebo. Error bars reflect 95 % CIs.*

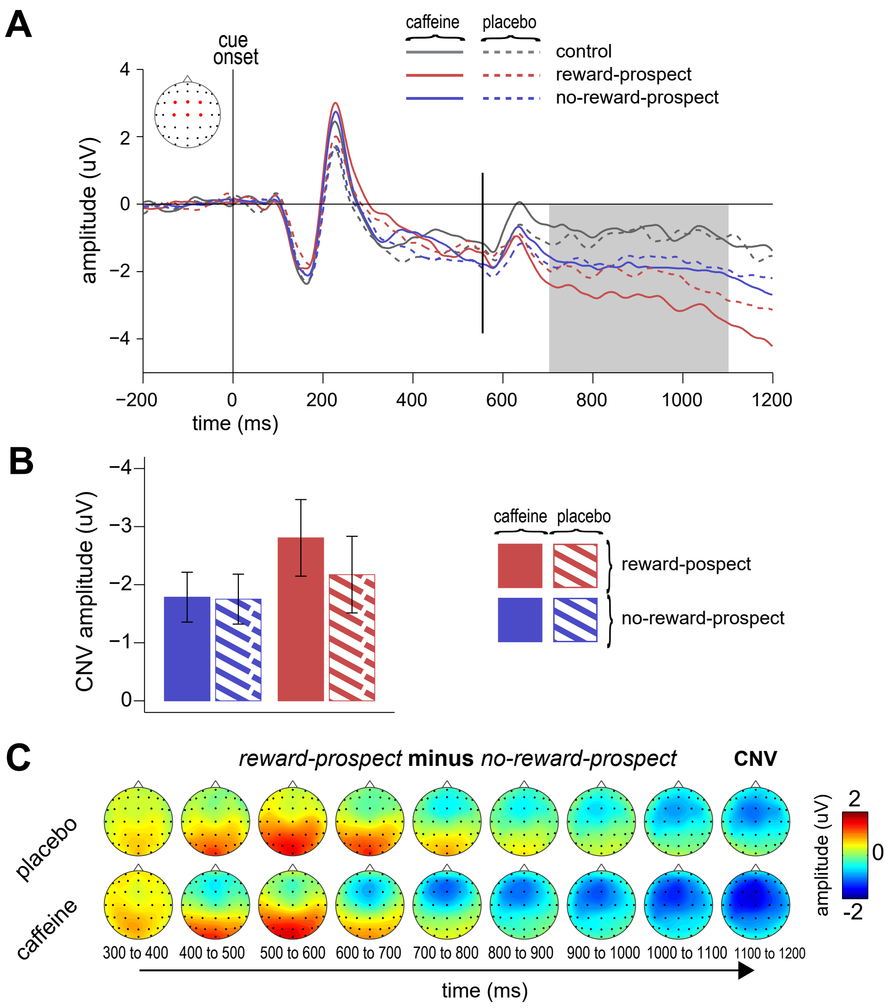


EEG results

Cue-evoked brain activation: Influence of reward and caffeine on the CNV

Visual inspection of Figure 3 shows that the ERPs evoked by both the reward-prospect and no-reward-prospect cues started to diverge from the control trials around 700ms after cue onset. In the conditions for which the participants were cued that an upcoming Stroop stimulus would appear, a more pronounced fronto-central slow-wave negative deflection (the CNV) was observed compared to the control condition (in which the participant knew no Stroop stimulus would be coming). The amplitude of the CNV was more negative following a reward-prospect cue versus a no-reward-prospect cue, replicating the results of van den Berg et al., (2014) (main effect of reward: $F(1,25) = 9.8$, $p = .004$). Whereas we did not observe a significant interaction between reward and caffeine on the behavioural measures, the effects of reward on CNV amplitude were found to be dependent on caffeine condition. More precisely, the CNV following reward-prospect cue (vs no-reward-prospect) was larger when participants had received caffeine prior to the experiment as compared to placebo, while this CNV difference between reward-prospect and no-reward prospect was not significant in the placebo condition (**Figure 3 A and B**) (reward × caffeine interaction: $F(1,45850) = 8.4$, $p = .004$; caffeine $_{\text{reward-prospect}}$ *minus* $_{\text{no-reward-prospect}}$: $t(36) = 4.04$, $p < .001$; placebo $_{\text{reward-prospect}}$ *minus* $_{\text{no-reward-prospect}}$: $t(36) = 1.7$, $p = 0.11$).

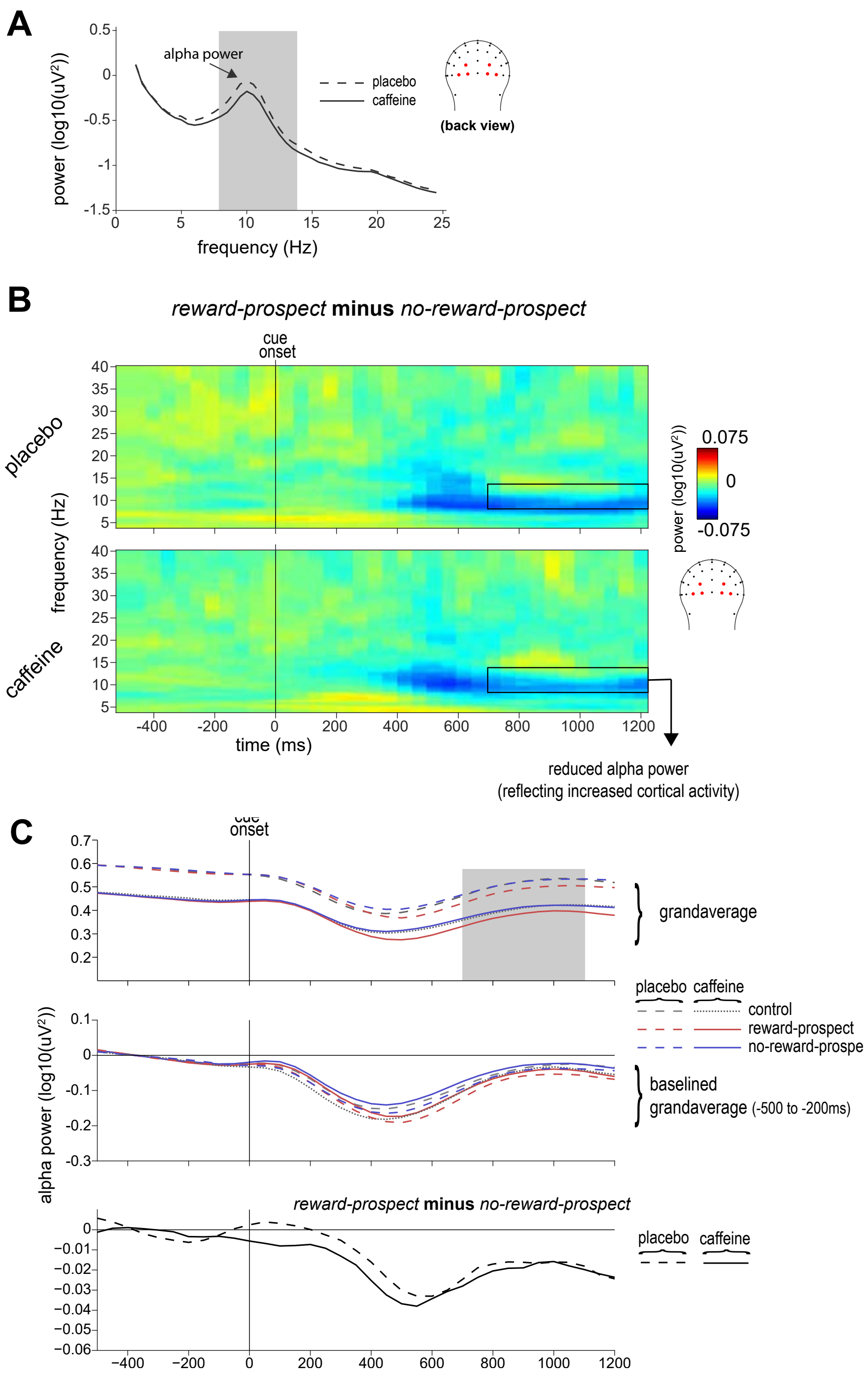
Figure 3: *Cue-evoked ERPs.* **A**. *Grand average cue-evoked responses from fronto-central ROIs for the different conditions.* **B**. *CNV amplitude for the different conditions.* **C**. *Differences in the topographical distributions of the no-reward-prospect versus control and reward-prospect versus no-reward-prospect conditions.*



Cue-evoked brain activation: Influence of reward and caffeine on alpha power

The frequency analyses revealed lower levels of sustained alpha power (8 to 14 Hz) (i.e., sustained across the whole session) over occipital brain regions if participants received coffee with caffeine prior to the experiment as opposed to coffee with placebo (**Figure 4A**) (main effect of caffeine: $F(1,24) = 7.96$, $p = 0.009$), replicating previous work (Kenemans & Lorist, 1995). In addition to these sustained effects of caffeine, we observed cue-evoked changes in alpha power over occipital electrodes. More specifically, after presentation of the cue, occipital alpha power was lower in the caffeine condition as compared to the placebo condition (main effect of caffeine: $F(1, 24) = 14.3$, $p < .001$; **Figure 4B and C**). In addition, alpha power decreased in response to reward-prospect cues compared to no-reward-prospect ones (main effect of reward: $F(1, 25) = 13.3$, $p = .001$; **Figure 4B and C**), replicating previous results for this preparatory-attention effect (van den Berg et al., 2014). In contrast to the CNV effects, no interactions between the effect of caffeine and reward was observed in terms of cue triggered alpha power ($F(1,45847) = 0.01$, n.s.).

Figure 4 (page 100): *Sustained and cue-locked changes in oscillatory power measured over the occipital channels.* **A**. *Sustained EEG power across the entire session revealed lower alpha power when participants received caffeine vs. placebo.* **B**. *Cue-locked spectral power revealed a decrease in alpha power prior to the presentation of the imperative Stroop stimulus, irrespective of caffeine condition.* **C**. *Changes in alpha power over time relative to the onset of the cue-stimulus.*



Differential effects of reward and caffeine on CNV and alpha power

In the above analysis we observed distinctive effects of caffeine and reward on CNV amplitude and cue-evoked alpha power. To statistically test if these relationships were indeed distinctive, we first z-transformed both CNV amplitude and alpha power, ensuring a comparison on the same scale. Next, we added a neural marker (CNV or alpha) as a predictor variable to the model that is described in **Table 1** under cue-interval analysis:

$$z_n = \beta_{0,j} + \beta_1 reward_{j[n]} + \beta_2 caffeine_n + \beta_3 session_{j[n]} + \beta_4 reward \times caffeine_n + \beta_5 reward \times marker_n + \beta_6 caffeine \times marker_n + \beta_7 reward \times caffeine \times marker_n + \epsilon_n$$

This analysis revealed a significant three way interaction (caffeine × reward × marker: $F(1{,}91719) = 4.9$, $p = 0.027$), providing supporting statistical evidence for differential effects of caffeine and reward on CNV amplitude and evoked alpha power.

Influence of reward and caffeine on the processing of the Stroop stimulus

As expected, conflict processing was found to evoke the classical early incongruency related negativity ($F(1{,}24) = 28.4$, $p < .001$), and a larger LPC ($F(1{,}24) = 14.7$, $p < .001$). Consistent with the behavioural effects of congruency, these congruency effects were not significantly modulated by caffeine or reward (**Figure 5C**).

In the target-stimulus interval we also observed a more positive, widespread, and long duration ERP amplitude in the reward-prospect condition compared to the no-reward-prospect trials, starting around 300 ms post-stimulus in both the fronto-central and parietal ROIs (ROI*fc*: 400 – 500 ms; main effect of reward: $F(1{,}25) = 18.6$, $p < .001$; ROI*p*: 700-800 ms; main effect of reward: $F(1{,}25) = 6.7$, $p = 0.016$). In addition, in the caffeine condition, more positive amplitudes were observed in the fronto-central ROI compared to the placebo condition (ROI*fc*: 400 – 500 ms; main effect of caffeine: $F(1,24) = 4.3$, $p = 0.047$[2]). This effect, however, did not reach the level of significance in the parietal ROI (ROI$_p$: 700-800 ms; main effect of caffeine: $F(1,24) = 1.0$, p = n.s.). Visual inspection of **Figure 5 A** shows that the largest positive-polarity activity was observed in the condition in which the participant both received caffeine and was cued with reward-prospect, while the lowest positivity was elicited in the no-reward-prospect condition during which the participant received placebo. This observation was supported by an interaction between caffeine and reward-prospect, which was significant at trend level (at $p < 0.1$) in the fronto-central ROI and while it reached significance in the parietal ROI (caffeine × reward interaction; ROI$_{fc}$: $F(1,40793) = 3.2$, $p = .07$; ROI$_p$: $F(1,40789) = 10.6$, $p = .001$). In the parietal ROI we observed an effect of reward-prospect in the caffeine condition (caffeine *reward-prospect* **minus** *no-reward-prospect*: $t(33) = 3.6$, $p = .001$), while there was no observable effect of reward-prospect in the placebo condition (placebo *reward-prospect* **minus** *no-reward-prospect*: $t(34) = 1.2$, p = n.s.).

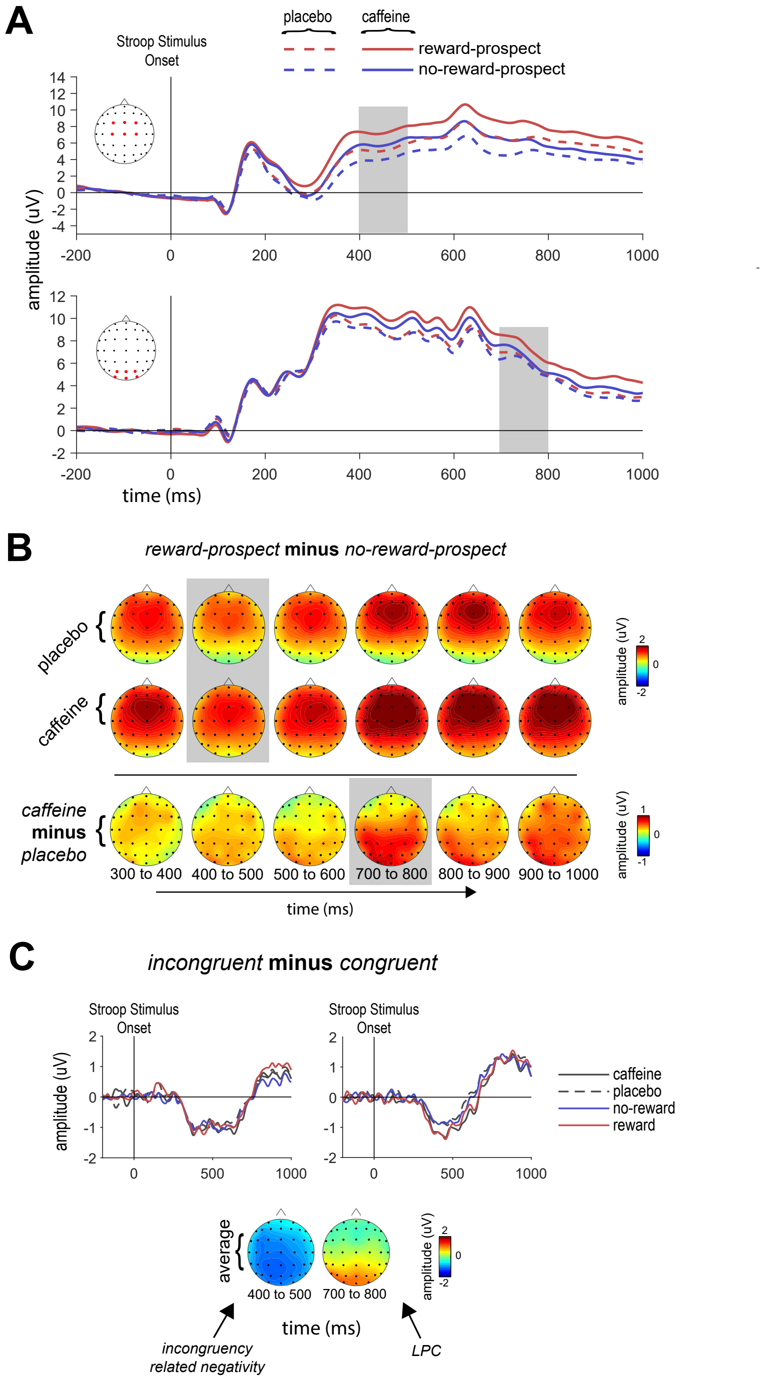


Figure 5: *Stimulus-evoked ERPs (collapsed over congruency).* **A**. *ERP traces from the fronto-central (top panel) and parietal (lower panel) ROIs, showing the large late positive complex (LPC) [also early interval is indicated by a grey bar] wave.* **B**. *Topographical distribution of the difference in ERP amplitude between the reward-prospect and no-reward-prospect trials for the caffeine and the placebo conditions separately, showing the effects of these manipulations on the LPC.* **C**. *ERP difference waves (incongruent minus congruent) for the caffeine and placebo sessions (collapsed across the levels of reward), and for reward-prospect and no-reward-prospect (collapsed across caffeine condition). Topographical scalp plots show the difference in ERP amplitude between incongruent and congruent trials.*

---

[2] *Note that in the GLM-ANOVA (Appendix IV) the effect of caffeine was significant at trend level (F(1,25) = 3.6, p = 0.069).*

Discussion

Two factors that are separately known to influence preparatory attention are caffeine and the prospect of reward. The goal of the present study was to investigate if and how these factors interact to enhance attentional preparatory activity and subsequently improve the processing of and response to task-relevant information that can potentially lead to monetary rewards. To do this, we conducted a two-session experiment in which participants performed a cued-reward Stroop task, while behavioural and neural-activity measures were acquired. Each trial of the task consisted of a cue that indicated whether a colour-word Stroop stimulus would or would not follow, and whether there was a prospect of gaining monetary rewards for good performance on discriminating that colour-word Stroop stimulus. Before the start of each experimental session, participants received either a cup of decaffeinated coffee with caffeine or with lactose (placebo) (both 3 mg/kg bodyweight). The key results showed that caffeine intake indeed resulted in greater enhancements of cue-triggered attention-related neural processes and better behavioural performance especially when the cue indicated that there was a potential for reward for good performance on the subsequent Stroop task. The findings described in this paper advance our understanding of how two previously separately studied factors, both of which impact behavioural performance, interact on a neural level to improve behavioural performance even further.

The level of general neural arousal substantially increased after the intake of caffeine (vs. placebo). Replicating previous studies, we observed reductions of power in the alpha frequency band (8 to 14 Hz), reflecting an increase in cortical brain activity and alertness (Kenemans & Lorist, 1995; Scheeringa et al., 2012). This increased state of alertness when participants had received caffeine prior to the start of the experimental session, was paired with improved behavioural performance (both RTs and accuracy), suggesting that caffeine intake improves the general state of the brain to such an extent that individuals can react more effectively to external stimuli and events.

In line with previous research, the behavioural responses to the colour-word Stroop stimuli were faster and more accurate when preceded by a reward-prospect (Padmala & Pessoa, 2011; Schevernels et al., 2014; van den Berg et al., 2014). The effects of reward-prospect and caffeine, however, appeared to be additive in terms of RTs and accuracy, and thus behavioural performance was most optimal when there was both a prospect of reward and caffeine had been ingested. In contrast to these results, some previous studies have reported that caffeine is especially effective when, due to the subject's state, behavioural performance was *not* optimal. For instance, it has been found that caffeine can have a profound impact on response times and accuracy particularly when participants are fatigued (Lorist & Tops, 2003; Nehlig, 2010). Mental fatigue is a state that has been associated with decreased behavioural performance, usually after continuously performing a taxing task for an extended time period (Lorist & Faber, 2011). Thus, our finding that caffeine actually boosted behavioural performance similarly in both the more optimal reward-prospect condition and the suboptimal no-reward-prospect condition would seem to differ from previous reports that have suggested that the stimulant effects of caffeine on behaviour are most pronounced in situations when attentional control of perceptual functions is reduced.

Increased behavioural performance was preceded by changes in brain activity related to event-related attentional preparation. These effects were modulated by caffeine and reward, resulting in enhancements of both the fronto-central CNV and posterior alpha power starting at ~700 ms following the cue, reflecting the marshalling of neural circuits that have been associated with preparatory attention (Corbetta & Shulman, 2002; Hillyard, 1969) and the enhancement of sensitivity to visual information (i.e., alpha power) (Van Den Berg et al., 2019; Worden et al., 2000), after caffeine consumption and during reward-prospect trials.

In the present study, we used the CNV amplitude and occipital alpha power as markers for cue-triggered preparatory attention. First, we replicated that both caffeine and reward-prospect have an enhancing effect on the slow-wave CNV (Ashton et al., 1974; Schevernels et al., 2014; van den Berg et al., 2014) and on posterior alpha power (Kenemans & Lorist, 1995; van den Berg et al., 2014). Second, and probably even more intriguingly, we observed that the modulation of the CNV by reward-prospect could be increased even further after the intake of caffeine. It is important to note that in previous studies, both the amplitude of the CNV and power in the alpha frequency band have been statistically linked through correlation to improved target processing and more optimal behavioural performance (Grent-'T-Jong et al., 2011; van den Berg et al., 2014, 2016). Hence, one might expect that the CNV and alpha power could be reflective of the same underlying neural mechanism that is involved in attentional preparation. However, this hypothesis did not seem to hold in these studies, since no direct relationship between the CNV amplitude and alpha power could be established.

Here, we observed that the effects of reward and caffeine interacted on the CNV but were additive for alpha power. This dissociation suggests that these markers reflect windows into two different facets of preparatory attention. Although their specific neural and cognitive functions are not fully understood, the CNV has been suggested to reflect a general arousal effect related to anticipation for a task or event. For the CNV this effect is thought to originate from the fronto-parietal attentional control circuits (Grent-'t-Jong & Woldorff, 2007). Decreases in occipital alpha power have been correlated with increased BOLD signal in the visual cortices (Scheeringa et al., 2012), suggesting that the amplitude of this pre-stimulus oscillatory activity is inversely reflective of receptibility to information in the visual sensory cortices, acting like a mechanism that can filter specific incoming sensory information.

In the present study, we found that, on the one hand, caffeine influenced the neural circuitry underlying the CNV when cued with reward-prospect but not when cued with no-reward-prospect. On the other hand, alpha power was influenced by caffeine regardless of cue information. These results suggest that caffeine might work on our neural information processing system in at least two ways. First, by enhancing the general state of the subject, caffeine increases sensory processing capabilities (as reflected by lowered alpha power). Second, the effect of caffeine depends on the context of an event, that is, on its behavioural relevance for the upcoming task. Neurally this can be illustrated by a larger CNV to reward-predicting cues under the influence of caffeine (vs. placebo), while there is little effect of caffeine on the CNV triggered by cues that predict no-reward. Our finding that the prospect of gaining reward and the consumption of caffeine speeded up reactions to subsequent Stroop stimuli in an additive manner confirms, at least partly, the differential effect these factors on our attention system at the underlying neural level.

With regard to these effects of caffeine, previous studies have found a pronounced effect when mental fatigue occurs (and thus when slower RTs would have been observed), indicating that caffeine enhances the state of the subjects in such a way that it helps them overcome (or at least compensate for) low levels of neural arousal that tend to impair behavioural task performance. In line with these observations, we found that caffeine did indeed seem to improve the general arousal state of the subject (as indexed by lower alpha power, thus reflecting higher levels of cortical activation). Here, we manipulated the behavioural importance of events through reward-prospect cues. Under these circumstances we found that caffeine resulted in greater enhancement of preparatory attention (CNV) for the more important rewarding events. Thus, in addition to the more general effect, caffeine can specifically boost attentional preparation for more salient or behaviourally important external events (as signalled by the reward-prospect cue) compared to other events that are less important. These findings suggest that improved behavioural performance due to the enhanced arousal state induced by caffeine depends on the context, and that this context is behavioural relevance.

From a theoretical perspective, the effects of both caffeine and reward-prospect on preparatory attention are expected to influence subsequent Stroop stimulus processing. For example, reward-prospect cues have been argued to enhance the saliency of specific impending events (Schevernels et al., 2014; van den Berg et al., 2014), resulting in the recruitment of the attentional-control circuits to improve the processing of those events. The effect of cueing of the potential for obtaining a monetary reward (vs. no-reward), was followed in time by a larger fronto-central distributed positivity to the Stroop stimulus, which was similar to the effect observed by van den Berg et al. (2014), but was not further modulated by caffeine (additive effects of caffeine and reward). However, the neural interaction effects on the later parietal LPC elicited by the Stroop stimulus paralleled the effects on the preceding cue-triggered CNV, with a larger enhancement of the LPC with reward prospect when caffeine was administered as opposed to a placebo. The observation that modulations of cue-related brain activity (CNV) by caffeine and reward was followed, in time, by a larger Stroop stimulus-related (LPC) neural activity, suggests that recruitment of the attentional-control circuits indeed ramifies in enhanced processing of task-relevant information. However, such an effect did not seem to result into a corresponding behavioural interaction between caffeine and reward in the current paradigm.

Further consideration and future directions

Next, we consider various open standing questions with regard to the findings as described above.

Caffeine did not significantly influence preparatory alpha power differentially as function of reward. Hence, while caffeine can modulate the anticipatory processes reflected by the CNV, caffeine does not seem to influence the sensory specific processes as reflected by posterior alpha power. Perhaps, due to the complex nature of the Stroop task (with 50/50 incongruent vs congruent stimuli), it is difficult to prepare the sensory cortex to the reception for specific task relevant information. For instance, in the classic Worden et al. (2000) study (and numerous replications and extensions since), the cue contained information as to where in space a potential target would occur, and this information resulted in alpha modulation (increase on the irrelevant side and decrease over the relevant one) over sensory cortices. Similar effects can be observed when learning to associate specific stimuli (e.g. faces and houses) with rewards (van den Berg et al., 2019). In this probabilistic learning study we found that when faces were rewarded (i.e., a gain) this was followed by a reduction in alpha power over the face-specific sensory brain regions.

Given the findings, the conclusion that it might be difficult for the participants to prepare given the demands of the word-colour Stroop task seems also illustrated by our behavioural findings that showed that the classic congruency RT effect (MacLeod, 1991) was not significantly modulated by reward-prospect, replicating findings from van den Berg et al (2014). An important point here is that in the present study our main question was not focused on the interaction between reward-prospect and stimulus-conflict processing. Furthermore, we did not find an interaction between caffeine and trial congruency, similar to previous research (Kenemans et al., 1999; Tieges et al., 2007). On the other hand, when congruency was blocked, Kenemans and colleagues did find an influence of caffeine on conflict processing (Kenemans et al., 1999). Accordingly, a point for future research is to understand how and if factors such as reward and caffeine can and do modulate the processing of incoming information to reduce conflict.

An important other point to consider when interpreting these effects of caffeine is the potential role of withdrawal effects. Withdrawal symptoms typically emerge 12–24 hours after stopping the consumption of caffeine (Juliano & Griffiths, 2004; Nehlig et al., 1992), which are manifested in the form of headaches, difficulty of concentrating, and effects on mood. Here, we asked our participants to abstain from drinking coffee for 12 hours prior to the experimental session that started at 9:00 a.m.. The choice of this timeline was specifically to have participants be tested after an overnight abstinence period during which they would normally have not drank coffee, and before withdrawal effects tend to kick in, thereby minimizing the potential for withdrawal effects. Besides, different studies have found that when there are caffeine withdrawal effects they are more pronounced for subjective symptoms compared to performance effects (Mills et al., 2016). In addition, in order to reduce the effect of expectancy on the experience of withdrawal symptoms (Mills et al., 2017), participants were not informed that in one of the sessions the coffee contained placebo instead of caffeine.

To check further for withdrawal effects, behavioural performance in the placebo condition of the present study was compared to behavioural performance in a previous study by Van den Berg and colleagues (2014), which used a similar experimental design with the main difference the absence of the caffeine manipulation and a higher overall monetary reward. Comparison of the behavioural RT effect of reward of the two studies (30ms in van den Berg et al, 2014, vs. 25ms here) suggests that the effect of caffeine observed in the present study is unlikely to be explained by withdrawal effects in the placebo condition. Similarly, Kenemans and Lorist (1995) found that, average response speed in their placebo condition was comparable to the results of a similar experiment without the caffeine manipulation; participants responded faster in caffeine condition than in the other condition/experiment, suggesting that withdrawal effects are unlikely to explain the observed caffeine effects (see also: Kenemans et al., 1999; Kenemans & Verbaten, 1998).

Lastly, it should be emphasized that we did not find an interaction between caffeine and reward on the behavioural responses, as well as some aspects of the neural responses, to the Stroop stimulus. However, it is important to note that the effects of caffeine by means of antagonism of adenosine receptors also include the modulation of other neurotransmitter systems (e.g., acetylcholine, noradrenaline, serotonin) that are known to influence other cognitive processes besides preparatory attention (Fredholm et al., 1999; McLellan et al., 2016). Accordingly, it is difficult to disentangle to what extent the effects on the Stroop stimulus processing were related to changes in preparatory attention prior to the Stroop target stimulus or to modulations of the processing of the Stroop stimulus itself by caffeine. Thus, the interpretation of the observed behavioural and neural effects in response to the Stroop target stimulus should be made with caution. We speculate, for example, that even if there are interactive effects reward and caffeine on preparatory attention, potentially additional caffeine effects (accomplished through other neurotransmitter systems) may swamp out or otherwise affect the observability of these effects in the final behavioural output. As such, this would seem to be an important point for future studies to address.

Summary and conclusions

In summary, we found various neural and cognitive processes were modulated by caffeine, reward, and the interaction between the two. The task consisted of a cue stimulus followed by a Stroop stimulus. Following the cue, preparatory anticipatory attention (as reflected by the fronto-central CNV) was modulated by reward-prospect and caffeine in an interactive way, while preparatory activity in the sensory regions (alpha power) was modulated independently by those factors, with no interaction. We also found that the processing of the subsequently presented Stroop stimulus was modulated by these factors, but again in different patterns. The earlier phase of Stroop stimulus processing (conflict processing - Ninc) showed independent effects of caffeine and reward-prospect. In contrast, the effects on the later phase (LPC) of the Stroop stimulus processing did show interactions, being modulated differentially by reward-prospect as a function of caffeine. Lastly, we observed that the behavioural responses, the final end-product of this cascade of various processes, showed additive effects in terms of caffeine and reward-prospect, but with no interaction. Important to consider here is that the processing of the Stroop stimulus would be expected to be modulated by both preparatory attention, as well as by direct influences of caffeine, although perhaps through other neurotransmitter systems. Accordingly, we cannot rule out that under different conditions we might see dependencies between caffeine and reward at a behavioural level. However, even though no interactive effects were observed in the behavioural responses here, by measuring brain activity during various stages of information processing leading up to that final behaviour, we showed how caffeine and reward can interact in their modulations of cognitive brain activity.

To conclude, we found that reward-prospect and caffeine intake can enhance neural attentional-preparatory activity, showing most optimal preparation and behavioural performance after caffeine intake and after a reward-prospect cue, as reflected by larger CNV and lowest alpha power being triggered by the cue. Additionally, we found that caffeine appears to especially improve preparatory attention (CNV) for a task that could potentially result in a reward. In a broader sense these findings indicate that caffeine can specifically target attentional preparatory neural processes for important salient events relative to events that are less consequential. These improvements in preparatory processes then result in enhanced target processing and more optimal behavioural performance.

Chapter 5

Ethics in design and implementation of technologies for workplace health promotion: a call for discussion

This chapter is based on:

de Jong, M. †
Roossien, C.C. †
Bonvanie, A.M. †
Mackelberghe, E.L.M. † (2021)

† These authors have contributed equally to this work

Ethics in Design and Implementation of Technologies for Workplace Health Promotion: A Call for Discussion. Front. Digit. Health 3:644539. doi: 10.3389/fdgth.2021.644539

Abstract

Aim: This study aims to initiate discussion on the ethical issues surrounding the development and implementation of technologies for workplace health promotion. We believe this is a neglected topic and such a complex field of study that we cannot come up with solutions easily or quickly. Therefore, this study is the starting point of a discussion about the ethics of and the need for policies around technologies for workplace health promotion.

Method: Based on a literature review, the present study outlines current knowledge of ethical issues in research, development, and implementation of technologies in the workplace. Specifically, the focus is on two ethical issues that play an important role in the worker–employer relation: privacy and autonomy.

Application: Two cases indicative for a multidisciplinary project aimed at developing and evaluating sensor and intervention technologies that contribute to keeping ageing workers healthy and effectively employable are explored. A context-specific approach of ethics is used to investigate ethical issues during the development and implementation of sensor and intervention technologies. It is a holistic approach toward the diverse field of participants and stakeholders, and the diversity in perceptions of relevant values, depending on their respective professional languages.

Discussion: The results show how protecting the privacy and autonomy of workers cannot be seen as stand-alone issues, but, rather, there is interplay between these values, the work context, and the responsibilities of workers and employers. Consequently, technologies in this research project are designed to improve worker conscientious autonomy, while concurrently creating balance between privacy and health, and assigning responsibilities to appropriate stakeholders.

Conclusion: Focusing on a contextual conceptualisation of the ethical principles in the design and implementation of digital health technologies helps to avoid compartmentalization, out-of-context generalisation, and neglect of identifying responsibilities. Although it is a long reiterative process in which all stakeholders need to be included in order to assess all ethical issues sufficiently, this process is crucial to achieving the intended goal of a technology. Having laid out the landscape and problems of ethics around technologies for workplace health promotion, we believe policies and standards, and a very overdue discussion about these, are needed.

Keywords: *privacy, autonomy, generalization, responsibility, ethics, responsible research and innovation*

Introduction

A major challenge caused by the aging workforce is to keep workers fit for work (Kenny, Yardley, Martineau, & Jay, 2008) to achieve a sustainable workforce. Technological interventions can assist to maintain individual workability, for instance by addressing the needs of aging workers in an objective manner (Truxillo, Cadiz, & Hammer, 2015) and creating balance between individual capacity and workload through well-designed workplace health interventions (Kenny et al., 2008). Examples of digital health technologies that are applied in the workplace are accelerometers, measuring bending, standing and walking activity (Villumsen, Madeleine, Jørgensen, Holtermann, & Samani, 2017) and wearable sensors for measuring fatigue (Aryal, Ghahramani, & Becerik-Gerber, 2017). Technologies such as these are aimed at automatically measuring and intervening worker behaviour, by giving (automated) feedback through digital means such as smart phones or stand-alone digital applications. These digital health technologies are used in addition to existing workplace health practices.

Research into the design and implementation of digital health technologies is surrounded by ethical issues that require responsible research. It is important to think about what impact this technology might have on individuals who are targeted as potential users or even on society as a whole. Responsible research and innovation (RRI) is a field of science that aims to highlight these socio-ethical issues in research and innovation practices (Grunwald, 2014; Owen, Macnaghten, & Stilgoe, 2012). In the past decade, new knowledge and guidelines have been developed that empower researchers to incorporate the researcher's responsibility throughout the innovation process (Stahl, 2013; Stilgoe, Owen, & Macnaghten, 2013), focussing on anticipation of (un)foreseen ethical qualms, reflexivity on one's own role, inclusion of diverse perspectives, and responsiveness to societal needs. Studies that describe the employed techniques to overcome the socio-ethical issues in development are lacking (Fisher et al., 2015), and publications in the field of responsible research and innovation still struggle with three critical problems: compartmentalization, generalization, and vagueness about responsible use (Efstratiou et al., 2007; Kortuem et al., 2007; Leclercq-Vandelannoitte, 2017; Palm, 2009).

Compartmentalization of focus in the current setting refers to the focus on one part of the development or implementation phase, while not including the tension between the intended and actual use of a technology. Until now, studies have mostly focused on ethical issues in either the design of new technologies (Aryal et al., 2017; Efstratiou et al., 2007; Motti & Caine, 2014; Saurabh, Rao, Amrutur, & Sundarrajan, 2014) or ethical issues in the implementation of existing technologies (Kortuem et al., 2007; Leclercq-Vandelannoitte, 2017; Sole, Musu, Boi, Giusto, & Popescu, 2013). The issues surrounding implementation take technologies as a given and do not question their inherent values in the design. This situation does not do justice to reality: if design and implementation do not acknowledge each other's ethical concerns and intended values, the final use of the technology will not reflect the intentions of both sides. A broader view on the transition between design and implementation is called for (Jakobsen, Fløysand, & Overton, 2019) to facilitate responsiveness between these phases of RRI.

An example of compartmentalization can be found in the field of health care innovation. New innovations are often developed from the viewpoint of a technology-enthusiast designer, whereas many nurses and caretakers are not digitally skilled (Verheijden, Tijink & Heijltjes, 2020). The ethical concerns of designers might be solved by a technical solution, however, due to lack of technical skill, the users do not use the technology properly and bypass these ethical concerns. Take for instance the use of smart glasses in health care. The smart-glass is used to share images of patients in a healthcare institution with colleagues in order to get a second opinion. This is a privacy issue. Therefore, the design forces people to first agree to the terms, and then call the colleague using the tiny screen on the smart glass. This action, however, is difficult and requires training and practice. For digital starters, this is an insurmountable problem. Instead, they use the glass by letting a colleague set it up before they enter the room (thereby violating the right to privacy of the client) or by using other applications to facilitate the sharing of images, such as WhatsApp video calls. This makes the ethical issues and risks of privacy violations even bigger.

In the case of the second problem, generalization, a single issue is identified as a core problem and addressed in a general way without attention to the specific context. For example, privacy is one of the significant issues in the development and application of new technologies that collect large amounts of data of individuals (Al Ameen, Liu, & Kwak, 2012; Conger, Pratt, & Loch, 2013; Nissenbaum, 2010; Zhu, Gao, & Li, 2016). However, most analyses of privacy issues focus on technologies that are used in the public space. These analyses do not necessarily fit other important contexts, such as use of sensor technologies in the work environment designed for health promotion. With regard to new technologies designed for the work environment, specific issues that concern privacy in the worker-employer relationship remain unaddressed. Additionally, discussion lacks about how privacy is embedded in the broader context. For example,, specific features in the design of digital health technologies intended to protect the privacy of the user can actually decrease the autonomy of the user. This could be specifically problematic in the work environment. That is, research suggests that workers experience (Leclercq-Vandelannoitte, 2017) and fear (Damman, van der Beek, & Timmermans, 2015) a loss of privacy ánd autonomy due to the use of technologies and (preventive) health interventions in the workplace. This lack of context-specific knowledge on both privacy and autonomy results in ethical issues that are not appropriately addressed in the development of new technologies. Albeit not an example from the workplace, the recent development of apps to prevent the spread of COVID-19 illustrates this problem of generalization fairly well. During the development of these apps, one single issue, i.e. privacy, was identified as the core problem, while other ethical issues were not addressed as much as they ought to have been (Morley et al., 2020). Based on 349 interviews with participants from nine European countries, Lucivero and colleagues (2021) showed that instead of or besides fear of privacy violations, people were hesitant to use COVID-19 applications due to other issues, such as scepticism of feasibility and fear of reduced autonomy. In most European countries, the application was eventually used by only a small part of the population, which not only vastly reduced its effectiveness, but potentially also reduced trust in and potential use of future applications with similar goals (Morley et al., 2020). This mismatch between values addressed by the developers and the values that are important to the user shows that generalization is a common problem that is not addressed properly in the design of technologies. Even though, as this example illustrates, generalization has the potential to have a large impact on the outcomes and use of a technology.

Finally, the topic of responsible use of digital health technologies remains vague and insufficiently addressed. Providing transparency about responsible use as well as identifying who is responsible are lacking. For example, Leclercq-Vandelannoitte (2017, p. 151) observed that in the use of ubiquitous technologies in the workplace, neither workers nor employers recognize who is responsible for technology, nor do they understand the importance of responsible use of these technologies. Furthermore, designers do not provide insight into the responsible use of their designs. Thus, identifying responsible use is notoriously difficult due to interdependent design-use dynamics (Kiran, 2012). These dynamics entail that design and use continuously impact one another because a particular function is often the reason for the design of a technology. However, the adoption of the design can substantially change the function. An example is the innovation of the short message service (SMS), which was designed to enable mobile owners to receive messages about incoming voicemails as well as bills from their service provider (Taylor & Vincent, 2006). However, SMS developed into a primary function for communication between individuals, thereby posing additional design demands as well as responsibilities that were not relevant to the original function.

Although in principle new sensor technologies are developed to support the user, they can have unforeseen consequences that are unintentionally harmful to the user or to society (Schukat et al., 2016). For instance, health-insurance companies in the Netherlands ask their customers to share their personal activity data, monitored via a pedometer or step counter on their phones. By doing so these individuals could earn back part of their insurance fee. Although these marketing strategies are being framed in a way that they are beneficial to the user, there are other values at stake (e.g. inequality between individuals with different socio-economic status and use of health data by the insurance company). While activity trackers were initially developed in order to help individuals self-manage their health, commercial organisations now make use these simple devices for their own commercial benefits.

Both the example of the SMS and the activity tracker show that the interdependent design-use dynamics of such a technology make it difficult to predict how it will be used in the future and whether or not it will be used as intended. However, this difficulty should not hinder designers from at least outlining the responsibilities inherent in their designs.
This study aims to overcome these issues of generalization and compartmentalization and additionally identify relevant responsibilities in the design and implementation of digital health technologies in the workplace. We want to initiate a discussion on the ethical issues surrounding workplace health promotion and the role of technologies.

We believe this is such a neglected field that we cannot come up with solutions easily or quickly. Therefore, the present study is an invitation to engage in a discussion on the problems we encountered. Ideally, work health considerations and employers' responsibilities would be set in a trajectory of health over the lifetime of work. In this paper, the focus will be limited to the problems of developing and introducing technologies. These technologies, however, are intended to have an effect on health over the lifetime of work. We also want to point out that the problems we signal are not new but are acerbated by the introduction of currently available technologies. The examples we use might seem quite simple, conventional and not new at all. However, they show how slow we are to come up with solutions and how far behind we are in the discussion on ethical considerations on technologies in the workplace.

First, the present study outlines current knowledge on ethical (and legal) issues on the implementation of technologies in the workplace, specifically focusing on the two ethical issues that play an important role in the worker-employer relation: privacy (Spook, Koolhaas, Bültmann, & Brouwer, 2019) and autonomy (Damman et al., 2015; Leclercq-Vandelannoitte, 2017). Secondly, two cases were explored using a context-specific approach of ethics to investigate these ethical issues during the development and implementation of sensor and intervention technologies for health purposes in the workplace. This context-specific approach arose from the diversity of participants and stakeholders and differences in languages (different academic disciplines; fields of application) used.

Privacy of workers

Employers are obligated to guarantee a safe working environment for their workers and should be reluctant to meddle with the workers' private lives and personal data. Interfering with workers' health behaviour, especially as connected to lifestyle, is dubious at best. It targets individuals (at work and in a personal setting) instead of organizational and collective problems, even if the goal is sustainable employability (van Berkel et al., 2014). Therefore, sensor and intervention technologies should comply with several criteria to ensure worker privacy.

Firstly, according to the EU General Data Protection Regulation, article 15, section 1 (GDPR, 2016), the worker should be able to access all personal data and outcomes of sensor and intervention technologies, without the interference of others. Secondly, the employer should not have access to data and outcomes of individual workers or be able to derive these outcomes from group data (GDPR, 2016, sec. 6). Current regulations on data collection and individual privacy limit the possibilities of data sharing (GDPR, 2016). As stated in article 6, section 1, subsection d of the GDPR, data processing is only valid if it is necessary to protect the vital interests of the subject, hence, a life-or-death situation.

Legally, data sharing at a group level is only allowed if the data does not contain identifiable information, such as personal data traceable to individuals (GDPR, 2016, sec. 4). Specifically, when it comes to sensor data that cross the border between work and private life, serious legal concerns arise regarding data and health privacy (Brassart Olsen, 2020). It could be argued, however, that sharing digital health data with relevant actors, such as health and safety workers, is beneficial for workers in specific contexts. In case of workplace improvements, the use of personal data could help to improve working conditions. The GDPR, however, does not provide a legal basis for the exchange of personal data in these specific relationships (Arora, 2019), making it difficult to use digital health data in the work environment, even if it can improve a worker's health.

A needs assessment among workers with physically demanding work identified a demand for sensor and intervention technologies (Spook et al., 2019). However, respondents expressed concerns about what would happen with the personal data retrieved by the sensors, fearing their privacy would be violated, especially if employers had access to the data. These apprehensions confirm the findings of other studies (Choi, Hwang, & Lee, 2017; Jacobs et al., 2019). The GDPR, as described above, offers an extensive legal framework protecting the rights and freedoms of data subjects, ensuring data minimization, informed consent, good practice via the Data Protection Impact Assessment (DPIA), and privacy by design (GDPR, 2016; Lodge & Crabtree, 2019; Mulligan, Koopman, & Doty, 2016). Although this legal framework is intended to protect workers, in some cases workers are not necessarily protected by it, nor do they want to be protected in this manner. That is, workers also declared that they would share their data with their employer to explore possibilities to improve working conditions if they could retain full ownership of the data (Spook et al., 2019).

Absolutizing a legal framework potentially leads to narrowing the fundamental questions of why privacy is an essential moral value. Data protection is significant to ensure privacy, but it does not embrace a comprehensive understanding of the concept. Numerous scholars have warned against a reductionist conceptualisation of privacy as merely about the protection of the personal sphere, raising questions about possible conditions under which this protection can be overruled (Barocas et al., 2013; DeCew, 2015; Dwork, 2006; Mulligan et al., 2016; Nissenbaum, 2010; Solove, 2008). They have argued for a broader understanding of privacy based on a reflection of practice and context. A legal framework for privacy by nature is fixed; however, privacy as a value should be shaped by each situation. Nissenbaum (2010, p. 2) succinctly summarized this concept: *'What people care about is not simply restricting the flow of information, but ensuring it flows appropriately'*.

Privacy as an essentially contested and malleable concept is dependent upon, amongst other things, the context in which it is examined, and the social and technological circumstances that apply to this context. As the theoretical debate about privacy continues, there is a need for a context-specific approach. Mulligan and colleagues (2016, p. 15) have suggested an approach based on four questions: *'While dilemmas between privacy and publicity, or privacy and surveillance, or privacy and security persist, the question we more often face today concerns the plurality available to us amidst contests over privacy: Which privacy? For what purpose? With what reason? As exemplified by what?'*. These questions enable researchers and practitioners to pragmatically define the relevant characteristics of the applicable notion of privacy.

Worker autonomy

A significant challenge for a workforce that will continue working into older age is to keep workers fit for work (Kenny et al., 2008). Van der Klink and colleagues (2016, p. 74) suggest to focus on sustainable employability based on a capabilities approach. Maintaining and supporting the ability of workers to continue working depends on the adaptation of work behaviour to changing circumstances. Worker autonomy in the self-regulation of work behaviour is crucial in this process (Ryan & Deci, 2006). Hence, organizations are introducing an increasing number of digital health devices on the work floor with which workers can regulate their tasks and work behaviour to ensure the autonomy needed for self-regulation.

Technological interventions can assist in maintaining workers' ability to work, for instance by developing technology that addresses the needs of ageing workers objectively, such as interventions that increase physical activity and ergonomically flexible workplaces (Truxillo et al., 2015). Thus, digital workplace health interventions can create a balance between workers' capacity and workload (Kenny et al., 2008), and sensor technologies, such as activity monitors and heart rate monitors, can accurately monitor workload. Additional intervention technologies, such as smart chairs (Goossens, Netten, & Van der Doelen, 2012; Roossien et al., 2017) can support workers in altering behaviour to prevent and solve health problems effectively.

Workers are willing to adopt sensor technologies that are perceived as useful (Choi et al., 2017; Jacobs et al., 2019), but workers' willingness to use these technologies depends on the addressing of concerns about data security and technology misuse (Jacobs et al., 2019). Philosophically, autonomy is complex, and caution is necessary to narrow the notion of autonomy to an idea of self-determination. Autonomy is a normative idea that directs actions governed by a responsible commitment to the norms with which one binds oneself. It can be about one's willed ideals as well as a commitment to the norms and standards people encounter and adopt because of a specific setting, such as the workplace. Thus, autonomy, also referred to as 'conscientious autonomy' (Kukla, 2005), covers the high moral values that direct peoples' lives as well as small practical commitments that shape ordinary happenings. For instance, if someone values being healthy, practical commitments could include walking to work instead of driving and taking the stairs instead of riding in an elevator.

Responsibility in the work environment

The ultimate responsibility for safeguarding the work environment lies with employers. Employers are responsible for the capabilities of their workers, actively preventing harm and accidents (Arbeidsomstandighedenwet, 1999; Palm, 2009). For workers who labour physically, employers must protect workers' safety via periodic occupational health examinations and safety monitoring (Arbeidsomstandighedenwet, 1999). Despite employers' limited access to the outcomes of regular health checks, this examination protects workers because occupational physicians can access health data and warn workers of potential issues while bound to professional confidentiality.

To protect workers while using sensor and intervention technology, all stakeholders must be responsible for the proper use of these technologies (Johnson & Powers, 2005), although employers may have different views on this responsibility than workers (van Berkel et al., 2014). Both workers and employers acknowledge the responsibility to prevent harm in the workplace. However, many employers consider the responsibility to stay healthy and fit for the job to be the worker's responsibility, while workers embrace autonomy in their lifestyle choices (van Berkel et al., 2014). These contrary views see health as either a safety discourse or a lifestyle discourse (Allender, Colquhoun, & Kelly, 2006). Nevertheless, the responsibilities of workers and employers in both discourses must be examined through context-specific ethics to prevent ambivalence in the worker-employer relationship (van Berkel et al., 2014).

Practical examples

Project description

The project SPRINT@Work is an EU-funded interdisciplinary project aimed at developing and evaluating sensor and intervention technologies that contribute to keeping ageing workers healthy and effectively employable (Bonvanie, Broekhuis, Janssen, Maeckelberghe, & Wortmann, 2020; de Jong, Bonvanie, Jolij, & Lorist, 2020; de Jong, Jolij, Pimenta, & Lorist, 2018; Roossien, Baten, van der Waard, Reneman, & Verkerke, 2021; Roossien, Heus, Reneman, & Verkerke, 2020; Roossien, Hodselmans, Heus, Reneman, & Verkerke, 2021; Roossien, Krops, Wempe, Verkerke, & Reneman, 2021; Roossien et al., 2017). These health-related technologies were developed and implemented by researchers and engineers from a variety of disciplines (cognitive neuroscience, information management, biomedical engineering and rehabilitation medicine, community and occupational medicine), in collaboration with companies. The developed sensor and intervention technologies lead toward an automated, digital process of behavioural assessment of employees for health self-management purposes. Cognitive neuroscience and information management were represented by one professor and one PhD candidate, biomedical engineering and rehabilitation medicine were represented by two professors and one PhD candidate, and community and occupational medicine were represented by two professors, one post-doctoral researcher, and one PhD candidate. The four PhD candidates acted as executing researchers.

Procedure: context-specific approach of ethics

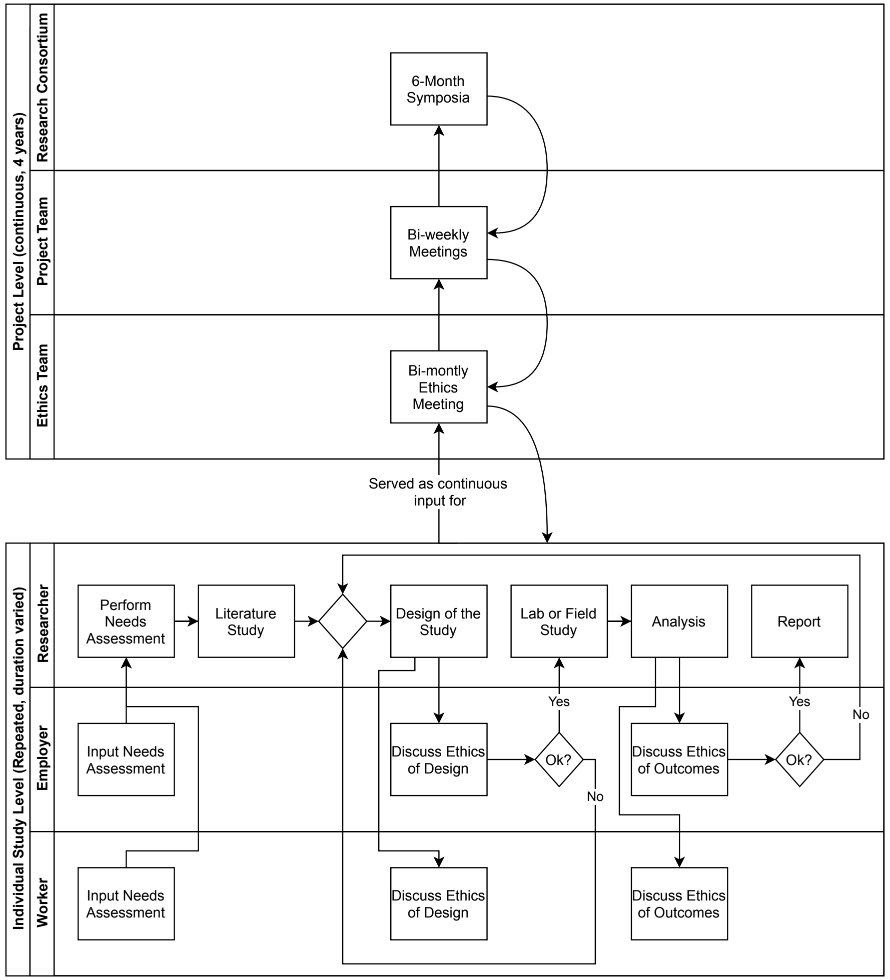
In several intervision sessions between the executing researchers, and later, the entire project team, the following issues were addressed: (a) whether the legal framework of privacy identifies sufficiently what is at stake in the context of the development and implementation of sensor technologies for sustainable employability, and (b) whether self-management devices aimed to promote self-regulation can assist in enabling the autonomy of workers. The team developed a conceptual framework that contextualizes data protection and privacy issues as well as the notion of worker autonomy. This framework of context-specific ethics was helpful for both designing and implementing sensor technologies and it functioned as a benchmark for the researchers. That is, during the project, the researchers continuously checked whether their proposed design was in line with context-specific ethics. Additionally, this normative framework was continuously adapted using insights from the executed studies.

Case studies

The present study highlights two case studies that were performed by the researchers of SPRINT@Work. The first case study was about monitoring the core body temperature as a parameter of heat stress of firefighters. The objective of this study was to validate a wearable non-invasive core thermometer to monitor the core temperature of firefighters during firefighting simulation tasks (Roossien et al., 2020). The second case study was about a research on health self-management applications in the workplace of health care workers. This study aimed at investigating whether use of sensor and intervention technology enhances the autonomy of workers in self-regulating their health-related behaviour (Bonvanie et al., 2020).

In both studies, the employer decided whether the study could be executed within the company. Thereafter, workers could voluntarily participate in the field studies. Employers were not allowed to oblige workers to use the sensor technology, nor can they ask for data if the worker voluntarily uses a sensor technology (Dutch Data Protection Agency, 2016). The intentions were articulated according to the declaration of Helsinki on research involving human subjects (World Medical Association, 2013), stating that participants should voluntarily give informed consent.

Figure 1: *shows how the research process during the project SPRINT@Work took place. The researchers involved in SPRINT@Work executed studies individually, while discussing ethical issues with the employers and workers that participated in their studies. The researchers continuously interacted with fellow executing researchers and an ethicist in the ethics team. This ethics team then shared and discussed findings with the project team, including supervising researchers, and higher-level findings were shared with the consortium. The outcomes of meetings with the consortium, project team, and ethics team were used to improve the studies of individual researchers.*



Case study 1:
The case of firefighters

During their job firefighters are exposed to a high thermal load due to heavy physical activity, external heat exposure from fires and the wear of highly insulated protective clothing (Roossien et al., 2020). This can lead to heat stress and subsequent related health problems, such as exhaustion, dehydration, mental confusion, and loss of consciousness (Roghanchi & Kocsis, 2018). In more extreme cases, heat stress can cause permanent damage and can even be life-threatening (Krishnamurthy et al., 2017; Morgado, Talaia, & Teixeira, 2017), thereby affecting the long-term health of the firefighter, affect productivity and risk perception, and cause safety problems (Roghanchi & Kocsis, 2018). There are large differences between individual firefighters regarding how their body copes with excessive heat. Therefore, general guidelines regarding duration of exposure to heat are not sufficient for the whole population of firefighters. To prevent heat stress among firefighters, Roossien and colleagues (2020) aimed to develop a new technology that would allow for monitoring and intervening in real-time during potentially harmful work situations.

Overcoming compartmentalization

The firefighting department that participated in the design and development of the intervention indicated a desire for a wearable thermometer to measure the real-time body temperature, because they wanted more insight in heat stress during work. This solution was developed in this case study. The thermometer was worn in-ear and registered the real-time core temperature of the firefighter (Roossien et al., 2020). It is dangerous if the firefighters themselves become distracted by immediate feedback about the obtained data, and they neither have time nor opportunity to monitor the feedback and data from their own sensors. Therefore, it is necessary that other colleagues, such as the captain, monitor the current body temperature of their workers on-site. This way they are able to intervene when the monitors show changes in the body temperature which could potentially harm the workers.

During development and testing, the researcher discussed issues regarding data sharing and confidentiality with both workers and captains, as well as with the other researchers, in order to find ways to overcome the potential issues regarding privacy and worker rights (see figure 1). Legally, an employer cannot ask permission to access the personal data of workers (GDPR, 2016, sec. 4), even if it is to the workers' advantage and safety. This issue points to ambiguity in the data protection law regarding the protection of workers' privacy opposed to the responsibility of the employer to safeguard workers' health and safety. Employers cannot, under any circumstance, use personal sensor data for the protection of health and safety of their workers, even though employers have the responsibility to protect workers from harm in the work environment. An ensuing focus for the research team was to explore how privacy could be conceptualized in the specific context of sensor technologies at the workplace, despite such ambiguity.

An agency-based approach to privacy

Following the pragmatic approach of Mulligan, data sharing in the case of the firefighters was analysed to determine what kind of privacy might provide sufficient protection. Control over personal information, such as the core temperature and heart rate of the firefighter, is a critical target for protection. As previously stated, from the perspective of the GDPR, this type of data can only be accessed under stringent circumstances and must be handled by a health professional who is bound by professional confidentiality. Nevertheless, in the case of a fire, no such health professional is available. Thus, the harm that supposedly would be prevented by enforcing data protection might be superseded by the prevention of more prominent harm. This example illustrates how information becomes ethically and normatively significant. Not because it is about specific values such as privacy but because the context allows its use for action. In this case, the possible prevention of overheating. Hence, it is not about what information one has but about what one can do with that information.

Manson and O'Neill (2007) called the above explanation an agency-based model of informing and communicating, where it is necessary to analyse what the agent, in this case, the firefighter captain, can do with the private information obtained. If overheating can be prevented, firefighters might want the option to share sensor information with their captain, although the captain is not bound by confidentiality as a health professional. Hence, the firefighters' permission for the captain to access this information is based on the specific agency of the captain to protect the firefighters from overheating. A different way to protect the privacy of firefighters is making sure firefighter captains are bound by the confidentiality of their own profession.

The answers to Mulligan and colleagues (2016) questions — 'Which privacy? For what purpose? With what reason? As exemplified by what?'— is that in the case of the firefighters, the privacy at stake is the ownership of personal data obtained by sensor technologies. The purpose of privacy is to give the firefighters control over their data, not only to prevent the employer to use of this personal information but also to allow the firefighters to share the data as they deem acceptable. The agency-based model exemplifies this purpose: in an ideal situation, the firefighter can opt to share data for protection from health hazards with the captain, who can act to prevent health hazards but cannot use the data for any other purposes, because the data is formatted in such a way that only the direct hazard of overheating is shown. This could for instance be done by using a traffic light figure, that only shows whether a situation is safe (green), or a reason to be alert (orange) or immediately withdraw the firefighter (red). In cases where direct indication of this risk of overheating is considered too much of a privacy violation, the agency-based approach could also allow including other health and safety indicators, such as an almost empty oxygen tank or another workplace risk. In this way, an orange or red warning light does not solely give the captain information on worker health, but also on health and safety risks in general. This example shows that a narrow interpretation of privacy might result in diminishing safety: if privacy is unidimensional, and the only choice would be to decide to share the data with the employer, either the firefighter would accept more significant risks during the execution of the job because the data would be hidden (as in the GDPR), or the employer would have full access to all data, which could lead to misuse for other purposes.

The case of the firefighters showed a disbalance between what is actually beneficial for the firefighters' health, and the regulations that are meant to protect them. This is a major problem when implementing new technologies in the work environment. Given that the law not yet protects the user in fiduciary relationships in certain professions (GDPR, 2016), it is important to acknowledge these design-use dynamics in the design phase of a new technology, and come up with solutions that could help overcome this gap in the law. Although some researchers already call for changing the law for fiduciary relationships (Arora, 2019), this would be a long and arduous process. Even if the law would change on this matter, it would still be important to define in which situations data sharing is condoned and with whom sharing health-data is necessary. Therefore, the agency-based approach asks for a thorough discussion with all stakeholders involved about what type of data is necessary to share with other actors and with whom in order to protect the firefighters' health (as can be seen in the process described in figure 1: level individual study). For instance, is it necessary to share raw data or would aggregated data suffice? Is it important to collect data for longer periods of time or can the data be removed directly after the fire was put out? But also, who has access to the data and how can it be prevented that other colleagues have access to the data. This can also be an indirect result of the use of a sensor. What happens for instance if one firefighter is called back more often than other firefighters? Agreements on these issues should be strictly documented and revised if necessary.

Responsibilities of stakeholders

In the case of the firefighters, the employer is serious about the responsibility for the health of the workers. The GDPR, however, prevents the employer from using personal data to protect firefighters from overheating in an emergency. In this case, the workers are at an impasse. Distraction from the task could cause immediate risks to themselves and colleagues; thus, it is impossible to self-monitor their current health parameters. This discrepancy between the desired situation and current regulations is frustrating for the fire department because the captains wish to protect their firefighters, but the GDPR makes it impossible for captains to use data for the goal of protection of workers.

Case study 2: The case of the healthcare workers

Healthcare workers are often subject to irregular working hours due to shift work. These work characteristics can make it more difficult to uphold healthy habits, such as daily exercise and a balanced diet (De Jongh & Mcdougal, 2014). An unhealthy lifestyle for a healthcare worker not only impacts their employability in the long term (Hendriksen, Snoijer, de Kok, van Vilsteren, & Hofstetter, 2016), but also impacts the public's view on the healthcare institution, because the healthcare workers are assumed to ‘know best’ about the impact of lifestyle choices on long-term health. Both the issues of long-term health and the exemplary function of their work are well-known to healthcare workers, which is why many of them actively try to keep up good behaviour. In this case study, a healthcare institution asked for an intervention that allows employees to self-manage their health, without having to explain themselves to the employer. An activity tracker supports these workers in their health, because it allows them to monitor their daily behaviour despite the irregular hours and workload, and thereby supports these workers in becoming and staying healthy (Kalantari, 2017).

Overcoming compartmentalization

The healthcare institution where the study took place is eager to improve and tries to incorporate the ideas of workers into their workplace health promotion policies. The activity tracker used is a tracker developed for the consumer market, meaning that the research team did not have any influence on the specifications of the tracker. During implementation however, the researchers decided to use proxy user accounts for all users, thereby enabling the researchers to tailor and alter the information that was given to the workers. These adjustments to the messages were intended to limit the impact on worker autonomy (see iterative process figure 1: individual project - researcher). Apart from the researchers and the participant, nobody had access to the data.

The use of sensor technologies to assist in sustainable employability hinges on offering workers objective feedback and interventions that allow them to self-regulate behaviour. Illustrative for the ideal of autonomy was a participant, self-identified as overweight and unfit, who was eager to experiment with an activity tracker. This activity tracker enabled her to receive automated digital feedback on her daily exercise behaviour. This worker was committed to improving her condition:

> *I value a healthy lifestyle. I have difficulties keeping up with that for all sorts of reasons, and this is an opportunity for me to get some non-intrusive and time-saving support. I also would like to be an example for the patients who visit here. They need people like me as role models, people who struggle but make an effort to improve their health.*

She referred to her value of personal health. Receiving an activity tracker did not provide autonomy. However, due to the activity tracker, she could autonomously commit to her value of becoming healthy. This value had a different application in her work context, a healthcare organization, where she wanted to set an example for others. She wanted to show that increasing daily exercise by walking more and taking the stairs is an essential commitment to improving health. Thus, in the work context, the worker wanted to achieve a healthy lifestyle as well as provide the moral value of being an example. She translated the value of her health and her position at work into a daily practical commitment of taking more steps. Thus, the use of this sensor technology helped her to achieve her ideal.

Nevertheless, the commitment of the worker was not only shaped by a momentous decision to accept the activity tracker. Her commitment was confirmed by making some progress in walking more steps. However, it was disaffirmed when a colleague from higher management saw her waiting for the elevator:

> *And then they are supporting 'the week of taking the stairs' [...], but then, when I am standing in front of the elevator, [colleagues] say "Oh, are you taking the elevator? We are taking the stairs!" That feels terrible. Really terrible.*

This encounter made her question whether the entire experiment was about her improvement in health and realizing her values, or whether it was ultimately about organizational control and cost reduction.

This example, although an individual experience, illustrates how personal autonomy can easily be threatened in a work environment if personal values are not acknowledged. Giving workers a health device does not merely provide a means for self-regulation, because the technology is embedded in a context that can promote or disavow the responsible commitment to the norms to which one is bound. This realization calls for reflection on how the introduction of technology can affect workers' autonomy, and how the context of the implemented technology influences the perceptions regarding workers' autonomy.

Worker autonomy as a prerequisite for health self-regulation was empirically investigated in the study of Bonvanie and colleagues (2020). It examined activity trackers that give feedback information on health-related behaviour to workers. The example of activity trackers is of interest because it is used as a technology that enables workers to self-regulate a healthy lifestyle (Bravata et al., 2007; Mattila et al., 2013). The underlying assumption was that the use of digital health technologies provides workers with autonomy via feedback and the freedom to respond to self-regulate health-related behaviour. Despite adjustments to the messages, intended to limit the impact on worker autonomy, these findings revealed that the use of a sensor technology did not significantly increase perceived autonomy and may have even reduced autonomy under certain conditions, especially for less healthy workers (Bonvanie et al., 2020). Moreover, workers who had used an activity tracker to monitor their behaviour before they received an employer-provided device experienced the same decrease in autonomy as workers who used the activity tracker for the first time. This finding suggests that the activity tracker does not limit the autonomy of workers; instead, perceived autonomy may decrease due to the hierarchical relationship between workers and employers.

A conscientious autonomy enhancing approach

The employer of the health-care institute who participated in this study demonstrated a value for healthy workers. That is, the employer already implemented several other activities and regulations, such as promoting a week of taking the stairs, providing a healthy cafeteria and offering a smoke-free property. Although independent researchers conducted the study, the normative standards of the activity tracker were encouraged by the employer. The goal was to walk 10.000 steps per day and take ten flights of stairs. Some participants agreed with this goal and internalized the normative standard. Others, however, did not and perceived the feedback as pressure to aim for 10.000 steps. Participants who shared the same value of healthy living as the employer but had other ideas to implement it, felt as if the activity tracker forced them to commit to someone else's normative standards.

These findings reflect the idea of conscientious autonomy (Kukla, 2005): autonomy that is committed to one's willed ideals as well as the norms and standards encountered in a particular setting that are adapted as normative. Hence, based on the disbalance between the individual goals and ideals of workers and the norms of their colleagues and employers, one can determine why the autonomy of certain workers declines when using a sensor technology. When implementing technologies or other interventions in the work environment, the employer therefore needs to pay specific attention to how the norms and culture in the work environment influence the worker's autonomy.

Participation in the study and being able to discuss the impact of technologies with different stakeholders within the development process, caused the employer to reconsider the current workplace health promotion policies. The employer altered their strategy into a more conscientious autonomy enhancing approach. This was achieved by including a more diverse group of workers in the decision-making and evaluating process regarding new technologies and interventions, thereby aiming to facilitate a healthy workplace and lifestyle for all workers.

Responsibilities of stakeholders

Similar to the case of the firefighters, the employer was responsible for the health of the health-care workers. This responsibility of the employer is limited to the work context, while the health of workers is also influenced by their private lives. By providing an activity tracker, the employer is walking a thin line between the work and the private context. One can ask the question, where does the responsibility of the employer stop and where does the responsibility of the worker begin, and where do they overlap? Interestingly, participants in the study of Bonvanie and colleagues (2020) stated that the ability to maintain their health is partially the responsibility of the employer, because their work environment has a large impact on this ability, and that their employer took this responsibility quite seriously. Both the employer and the workers experienced the intertwined nature of health, work and the home environment, and aim to improve the collaboration on improving the overall health of the worker (see process figure 1: individual level worker - employer).

Discussion

Previous literature on responsible research and innovation struggled with three major problems: (1) compartmentalization, (2) generalisation, and (3) vagueness about responsibilities. Rather than developing a theoretical approach to these problems, we highlighted two cases of the project SPRINT@Work. We aimed at describing how we explored the critical ethical issues privacy and autonomy in the development and implementation of digital health technologies in the setting of doing research. A context-specific analysis of both values was employed, keeping previous research and the legal context in mind. For the firefighters case study, this analysis resulted in the description of an agency-based concept of privacy, where it is necessary to analyse and regulate what the agent can do with the private information obtained (Manson & O'Neill, 2007). For the case study of the health-care professionals, this resulted in a conscientious autonomy enhancing approach for the design and implementation of digital health technologies in the workplace. When this approach is employed, all stakeholders (with a specific emphasis on the user(s)) have to be actively involved in the design and implementation phase in order to achieve the intended goal of the technology, which is to enhance health-related behaviour (Kukla, 2005).

Decompartmentalization of focus

Responsibilities regarding the assessment of risks of the new technology get indistinguishable when a transition between phases occurs (Jakobsen et al., 2019). More specifically, engineers and researchers might have reflected on the impact of their new technology, however, after the development phase, responsibilities shift towards the user or organizations that implement the technology. They do not necessarily reflect on possible ethical and societal risk, and primarily focus on productivity or increasing product acceptance (Chatfield, Borsella, Mantovani, Porcari, & Stahl, 2017; Leclercq-Vandelannoitte, 2017). Ethical concerns arise as soon as technological innovations are introduced (Martin & Freeman, 2004). Although an ethical script of an innovation shows what the default choices regarding privacy, responsibility and autonomy are, at the same time, the reaction of the environment on this built-in ethical script plays a significant role. The ethical script is mainly developed by the engineers and researchers who develop the technological intervention, but the response of the user and his/her environment on this ethical script largely determines the privacy of the user and their possibility to exercise autonomy. Using a multi-stakeholder approach may help to overcome this problem of compartmentalization by providing a smooth responsible transition from development to implementation.

In the case studies, the reflection on both design and field experiments involving health-related technologies in the workplace, caused both the researchers, employers and workers to reflect on the interpretation and implications of the concepts of privacy and autonomy (see figure 1). This approach of integrating development and use of the digital health technology was necessary to successfully implement techniques from the field of RRI, such as reflexivity and responsiveness. The context-specific approach allowed for a cyclic approach, using outcomes from early implementations of technologies as input for further development. As a result, researchers, employers and workers were able to work together to take unforeseen consequences of the technology into consideration, because they appeared during use by end-users. This then allowed the researchers and engineers to alter the technology or the choices that were made during development and implementation.

Both cases show the benefits of including the tension between the intended and actual use in the development and implementation of a new technology. In the case of the firefighters, the balance between safeguarding privacy and safeguarding health could only be reached because the researchers were able to use input from actual use (during job performance). More specifically, due to the interaction between the researchers, workers and the team captain, the application of the wearable thermometer for use in the workplace could be improved, which consequently benefits the firefighters' health. In the case of the healthcare workers, the researchers closely monitored the impact of the technology on the autonomy of workers at the workplace. By doing so, they were not only able to re-evaluate the benefit of the activity tracker, but also caused the employer to reconsider the current workplace health promotion policies and the manner in which these come to be.

Prevention of out-of-context generalization

A responsible decision to provide workers with sensor technologies to sustain their employability requires careful analysis of the values at stake in the context of the specific workplace and the individual worker (ten Have, van der Heide, Mackenbach, & de Beaufort, 2013). In case of privacy, the GDPR offers a basic framework for the implementation of protection measures, while it also leaves room for interpretation and discussion. The GDPR (2016) obligates and ensures that the decisions regarding data protection taken by the controller, for instance an engineer or a researcher, are taken with great care, especially when "processing of the data could result in high risk to the rights and freedoms of natural persons" (GDPR, 2016, sec. 35 (3)). In order to help the controller making responsible decisions regarding privacy of individuals, the data protection impact assessment (DPIA) (Bieker, Friedewald, Hansen, Obersteller, & Rost, 2016) is developed as a risk assessment method. This includes a multiple stakeholder approach to identify privacy risks. During meetings with stakeholders, a context-specific method of privacy by design is applied to design protection measures that are appropriate for a specific context.

The DPIA's (and the GDPR's) main focus is to protect the privacy of the user without paying much attention to other ethical issues in its analysis. Although it is a step in the right direction, in the development of new digital health technologies, other values, such as health, autonomy and responsibility, and the interplay between these values need to be reflected upon as well. The current study therefore used a context-specific approach of ethics (instead of privacy) to assess privacy and autonomy concerns in the workplace.

For both cases, the context-specific approach of ethics helped to identify the best approach to provide a framework of what is at stake in their specific context. Although from a different perspective, both the agency-based model of privacy (Manson & O'Neill, 2007), and the conscientious autonomy enhancing design (Kukla, 2005), can help identify bottlenecks, implicit norms, and courses of action during the development and implementation of new technologies and policies. These two approaches are a source of moral knowledge, given that the experiences in the field informed the researchers about what users value and the dynamics between the researcher, employer, and user were explored by testing the conceptualization of ethical principles in the work environment and further adjusted as deemed necessary.

Making implied responsibilities explicit

Acting responsibly regarding health in the workplace is considered important (van Berkel et al., 2014) but employers experience difficulties taking their responsibility, and in the case of health promoting technologies in the workplace, other stakeholders find it difficult to share this responsibility. Leclercq-Vandelannoitte (2017), in a study about the use of ubiquitous technologies in the workplace, observed that *"despite their prevalence and the importance of their consequences for workers, neither salespeople nor managers seem to be aware of them, feel responsible for them, or appear able or willing to identify the responsibilities involved in this process"*.

In the case of workplace health promoting technologies, responsibility involves multiple stakeholders with a prominent role for the employers (Palm, 2009), engineers (Doorn, 2012) and researchers (Doorn & van de Poel, 2012; La Fors, Custers, & Keymolen, 2019). To protect the privacy of workers while gathering personal data, all stakeholders need to take their responsibility for the use of the involved technology (Johnson & Powers, 2005).

The engineers and researchers have the responsibility to design the technology in such a way that it guarantees the privacy of the user and supports the user in his/her ability to react autonomously (Robaey, Spruit, & van de Poel, 2018; van de Poel & Verbeek, 2006; Verbeek, 2011). However, engineers and researchers often do not offer sufficient insight into what constitutes a responsible use of their designs (Leclercq-Vandelannoitte, 2017). Technologies are never value neutral (Martin & Freeman, 2004), and it is important that engineers and researchers explore how the development and implementation of their technologies influence and mould the ethical values such as privacy, but also the autonomy of employers and workers and help them reflect on this explorative process (Robaey et al., 2018; van de Poel & Verbeek, 2006; Verbeek, 2011). The responsibility of engineers and/or researchers should focus on perspectives such as value-sensitive design, critical technical practice, reflective design, and values in design (Cummings, 2006; Shilton, 2013).

The reflection on the responsibility of workers and employers is not a one-time action. As stated before, differences in interpretations of responsibilities can cause significant problems between workers and employers (van Berkel et al., 2014), and the use of technology often alters the original function (Kiran, 2012). When using (new) technologies, workers and employers should discuss the responsibilities and intended actions of these technologies with the designers. This discussion should also entail the continuous reflection of the employer, to determine whether the conscientious autonomy of the worker has improved. In the case of the healthcare professionals, sensor technologies enabled workers to take responsibility to target work-related health parameters within the workplace. In general, however, these technologies are most effective when workers feel autonomous to self-regulate health-relevant actions (Bonvanie et al., 2020). Thus, employers should be alert for unintended effects of sensor technologies and ensure an environment that facilitates workers to take their responsibility. When workers and employers share values, such as health, technologies that support the workers' personal goals could increase a sense of conscientious autonomy, thereby improving the self-regulation of healthy behaviour.

Limitations

The context-specific approach of doing ethics has been a valuable method to investigate the core ethical principles of the digital health technologies in the case studies. In specific to obtain a contextual conceptualization of the ethical principles in the design and implementation of these technologies. However, we realize that this approach was not utilized to its full extent.

Multiple useful tools are now available to help researchers perform responsible research and innovation (e.g. rri-tools.eu). However, at the start of the project SPRINT@Work, approaches to incorporate an ethics structure throughout the complete research cycle of multidisciplinary projects were lacking or at least not commonly practiced. Therefore, we started our journey with no clear approach in mind, and we developed our knowledge and approach as we went on.

In the case of the healthcare workers, this resulted in little attention to the pre-defined norms and values of the activity tracker. If we would have identified these norms and values before the start of the field studies, the researchers could have incorporated them in the design of the study. This could have prevented negative experiences of workers with the activity tracker.

In the case of the firefighters, we did not involve a specialist in the field of privacy law to help find potential legal solutions for data sharing while protecting the workers' privacy. The project team would have benefitted from actively involving a privacy specialist from the start of the project. This could have influenced the approach taken by the researcher that designed and implemented the technology, the technology itself and its suitability for future use.

At last, it can be stated that the problem of compartmentalization applies to us researchers as well. In order to be able to manage this extensive project, the individual research projects of the PhD candidates were defined as much as possible. Although this approach was meant to save time, it resulted in delays both in the individual research projects concerning the development as in the projects concerning the implementation of the new technologies.

Based on these limitations, we advise multidisciplinary teams to start exploring potential critical ethical issues right from the start of the project. They could use this paper as a first lead on how such issues could be identified. Even though it might not be their initial field of expertise, we appeal to future multidisciplinary teams to also report their findings and possible adjustments to our proposed approach.

Conclusion: Call for an overdue discussion

Based on a substantial literature review, we aimed to discuss the importance of context-specific ethics in design and implementation of digital health technologies. Focusing on a contextual conceptualization of the core ethical principles in the design and implementation of digital health technologies helps to avoid compartmentalization, out-of-context generalization, and neglect of identifying responsibilities. Although it is a long reiterative process in which all stakeholders need to be included in order to assess all critical ethical issues sufficiently, this process is crucial for achieving the intended goal of a technology. We call for multidisciplinary teams, including relevant stakeholders, involved in innovation practices in workplace health promotion to publish their way of doing ethics. Future research teams can learn from these experiences and use and improve their approaches.

Finally, having laid out the landscape and problems of ethics surrounding technologies for workplace health promotion, we believe that an evaluation of policies and standards and a very overdue discussion guided by the signalled ethical problems are needed. Laws and regulations aim to offer protection to users of new technologies, but tend to focus on data access and privacy. Thereby they leave out other values, such as responsibility and autonomy, that are in close interplay with privacy. It is therefore important that engineers and researchers of workplace health promotion themselves enter this debate. They should consider how the design and implementation of their technologies influence and mould the values of the users and adapt their technologies, to protect the user from harm, and increase the acceptance. However, it does not stop there. They should also enter the debate about how policies and standards hinder or promote workplace health promotion.

Chapter 6

Summary and General discussion

Aim and summary

The main objective of the present thesis was to investigate ways to monitor and improve cognitive performance during prolonged office work, taking into account the increasing number of older employees in our society. Previously, research mainly focused on different aspects of this overarching aim, in isolation. Research, for example, focused on the effects of mental fatigue on specific cognitive processes in a lab setting (e.g., Boksem et al., 2006; Hopstaken et al., 2016; van der Linden et al., 2003), the feasibility of sensor technology to detect mental fatigue in a working environment (Pimenta et al., 2013), or performing responsible research and innovation in order to make sure new sensor technologies are developed and implemented in a responsible manner (Flipse et al., 2014; Grunwald, 2014; Stilgoe et al., 2013). The aim of the present thesis was to implement a more holistic way of investigating cognitive performance in the working environment. That is, in the present thesis different stages of the research process, dedicated to the development cycle of sensor technology applicable to monitor and counteract performance decline during prolonged performance in a working environment, are described, including a discussion of collaborative research efforts.

In the present thesis, we investigated whether typewriting indices could be used as a method to monitor mental fatigue in the working environment. This research was conducted both in the lab, using markers in the EEG signal related to mental fatigue to validate the typewriting markers as indicators of mental fatigue (**Chapter 2**), and in the working environment to investigate whether the selected markers were sensitive to changes with time-on-task under less controlled task conditions. In addition, we also examined whether the effects of time-on-task on different time-scales (time-of-day and day-of-week) also influence task performance (**Chapter 3**). Due to our inability to keep up with task demands for an infinitive amount of time, caffeine is frequently used to boost our mental capacity.

In Chapter 4, we reported on a study aimed to examine the effect of caffeine on our ability to focus attention to relevant stimuli in our (work-) environment. In this study, we examined the effects of caffeine on both performance and underlying brain activity (**Chapter 4**). In addition to these experimental studies, we also focused on conducting responsible research and innovation. **Chapter 5** describes an illustration of the ethics approach of the multidisciplinary research project SPRINT@Work that aimed to develop sensor and intervention technologies for a sustainable workforce. The present thesis is part of this project. The study described the analysis of two ethical issues that accompany the development and implementation of new sensor technology (i.e. privacy and autonomy of workers), in a real-life research setting, using a contexts-specific approach of doing ethics that helps to avoid compartmentalization, generalization, and neglect of responsibility, three major problems in *responsible research and innovation*.

Part I:

Monitoring and counteracting mental fatigue

During a full-time work week, employees in the Netherlands work around 8 hours per day, for 5 days a week. Research in various settings shows that over the day, and with prolonged task performance, employees slowly start to disengage from their task, while performance declines (i.e. becoming slower, making more mistakes, and being less able to correct for mistakes; Boksem et al., 2006; Lorist et al., 2005). This state, called mental fatigue emerges during or after prolonged task performance and is generally characterised by feelings of aversion against continuing the current task). In order to maintain optimal performance, employees are therefore allowed to take either scheduled or self-paced breaks. Given that the manifestation of mental fatigue depends on several factors, both internal (motivation; Boksem et al., 2006) and external (environmental noise), and people are not very good at estimating when performance declines (e.g., Zhang et al., 2011) it is challenging to effectively prevent mental fatigue in the working environment without the help of external measures or regulations.

The key to mental fatigue

Developments in information technology now made it possible to monitor human-computer interactions. Given the increasing use of computers in the working environment, it might provide a non-intrusive way to monitor the effects of mental fatigue on job performance. One could, for example, monitor human-computer interactions by analysing changes in typing performance.

It is technically possible to continuously monitor typing speed, for instance, by measuring the time between two subsequent keystrokes (interkey interval) or the duration of a keypress (Pimenta et al., 2013, 2014). Although the interkey interval varies within and between key and word type (e.g., Kalfaoğlu & Stafford, 2014), in Chapter 2, we established a clear link between interkey interval and the time someone takes to finish a sentence, reflecting performance speed. Monitoring typing accuracy, however, is considered to be more advanced. First of all, there are different types of typing errors (Logan, 1999). For instance, someone could type the wrong letter or make a grammatical error. Typos and grammatical errors are easily recognizable by algorithms that have been built into different word processors, such as MS Word. These typing errors, however, just partly reflect task accuracy. Someone could also produce words that do not fit with the sentence or produce sentences of which the content is wrong. Due to the complexity of tasks in the working environment, especially in case of unstructured tasks, it is complicated to monitor changes in accuracy during regular office work.

Pimenta et al. (2013, 2014) investigated changes in typewriting dynamics during regular working activities. They found that employees type slower, press keys longer, and press the backspace key more often over the course of the day. These findings suggest a relation between typing performance and mental fatigue. However, because of possible interactions with confounding factors (e.g., task type), it is not possible to determine whether these changes in typewriting are actually related to mental fatigue specifically. In the second chapter of this thesis, we therefore investigated the relation between typewriting dynamics during prolonged task performance and brain activation patterns that were previously related to mental fatigue. In this study, twenty-four young (18-30 y) and twenty-four middle-aged (50-67 y) participants performed a standardized typewriting task for 2 hours, during which typewriting measures were recorded using a keylogger and brain activity was recorded using electroencephalography (EEG).

In this study, participants performed an adapted version of the task of Hoeks et al. (2004). During this task, participants were presented with a sentence of which the last word is either congruent (i.e. good semantic fit) or incongruent (i.e. bad semantic fit) with the rest of the sentence. The task was adapted to a typing task (see Chapter 2 for full task description). Typing behaviour was recorded using software that was developed by Pimenta et al. (2013, 2014). By analysing the output of the software, we could further look into the typing dynamics during prolonged task performance. Besides measures of typing speed (i.e., typing time, general interkey interval, and stimulus–key interval) and typing accuracy (incorrect words, corrected words, backspace use, and backspace sequence length), we also investigated the letter–letter interval, an index for speed during correct typewriting, the letter–backspace interval, which is an index for the initiation of error-correction, and the backspace–backspace interval, which is an index of speed during error-correction.

In order to investigate the relation between typewriting dynamics and neural activation patterns related to mental fatigue, event-related potentials (ERPs) were extracted from the EEG signal. ERPs describe the timing and organization of cognitive processes before, during, or after external events. The P3 is a positive wave between 300 and 500 ms after stimulus presentation that has been associated with attention and memory operations (Polich, 2007, 2012). Previous research showed that the amplitude of the P3 decreased when participants became mentally fatigued (Hopstaken et al., 2015; Lorist & Jolij, 2012; Wascher & Getzmann, 2014). By correlating the effects of prolonged task performance on typewriting behaviour and changes in the P3 over time, we were able to connect changes in typing performance to changes in neural activation related to mental fatigue. Our results revealed that participants who showed a larger decline in performance also showed a larger decrease in P3 amplitude over time, which indicates that typewriting markers can monitor changes in mental fatigue, at least in a controlled environment.

Consistently, previous studies have shown that performance speed declines with prolonged task performance (Langner, Steinborn, et al., 2010; Langner, Willmes, et al., 2010). In line with these findings, we found that performance speed decreased over time, reflected in the time that participants took to complete one trial. When we looked further into the effects of prolonged task performance on typing dynamics, we observed that this general effect of time-on-task on performance speed could be explained by three main effects: participants typed more slowly in general (1), there was an increase in the number of error-corrections with time-on-task (2), and participants typed slower during error-corrections compared to correct typewriting (3).

The first main effect regarding decreased typing speed during task performance can be related to deteriorations in cognitive functioning due to mental fatigue. Among others, fatigued participants have difficulties focusing their attention and preparing for their responses, resulting in decreased performance speed and lower performance accuracy (Boksem et al., 2005). In Chapter 2, the effects of mental fatigue on typing speed were correlated with a decreased P300 amplitude, suggesting that this decline in typing speed was related to lowered attention and memory processes.

Besides this general decrease in typing speed over time, there is an indirect effect of error-corrections on the time someone needs to complete a sentence correctly. That is, if people notice a mistake, they first have to correct this mistake by using the backspace key and then retype the correct words, increasing the number of keystrokes that is needed to type the correct(ed) sentence. This increase in the number of error-corrections during typewriting is not surprising. Monitoring performance requires higher-order mental functions, which are prone to the effects of mental fatigue (Lorist et al., 2000; van der Linden et al., 2003). That is, erroneous responses are usually followed by specific brain activation patterns, called the error-related negativity (ERN; Falkenstein et al., 1991; Gehring et al., 1993), and result in decreased response speed on the next trial (e.g., Rabbitt, 1979). Lorist and colleagues (2005) investigated these behavioural and brain activity patterns during prolonged performance on an Eriksen flanker task. They found that performance monitoring declined over time, which was reflected in a significant decrease of brain activity patterns related to error processing (ERN), and was accompanied by a decrease of post-error slowing. In Chapter 2, this decreased ability to monitor behaviour and concurrently adapt performance might have resulted in the increase in error-corrections, but specifically in the increase in errors that were not corrected.

Besides the increase in error-corrections resulting in a decline in typing speed, there was an additional, indirect effect of the increase in error-corrections over time that was related to the decrease in typing speed.

In line with Kalfaoğlu and Stafford (2014), we found that there was a difference between typing speed during correct typewriting and during correction typewriting. More specifically, participants took longer to initiate the first backspace keystroke and to initiate typewriting after a correction compared to correct typewriting. Thus, because typing speed during error-corrections was slower in general, and due to the increase in error-corrections, the time to type the text increased with time-on-task, as well.

As these findings illustrate, speed and accuracy influence each other. Error-corrections improve overall accuracy, but also clearly decrease typing speed. This pattern, where improving performance in one dimension (i.e. accuracy) results in a decline in the other dimension (i.e. speed) is called the speed-accuracy trade-off (Wickens & Hollands, 2000). Whether prolonged task performance will influence either one of these dimensions depends on context and the strategy of an individual. In some situations, one might prefer to work fast, while in other situations it is more important to perform accurately. Additionally, specific conditions and personal traits can also influence someone's tendency to perform a strategy that either favours speed or accuracy.

Starns & Ratcliff (2010), for example, studied age-related differences in the speed-accuracy trade-off over several tasks. Consistently, they found that younger adults are adjusting their strategy in order to balance the trade-off between speed and accuracy, while older adults generally tend to minimize errors, even if that means that they have to perform slowly. In line with these findings, we found that age indeed influenced the manifestation of mental fatigue in our study. Older adults tend to slow down with increasing time-on-task, while typing accuracy, measured by typing errors that were corrected and typing errors that were not corrected, remained stable over time. In contrast, younger adults slowed down, and showed an increase in both the corrected and uncorrected errors over time, reflecting a more balanced decline in performance with prolonged task performance.

These differences in strategy suggest that, in general, younger and older adults cope differently with the increasing demands of their environment due to prolonged task performance. That is, younger adults tend to adopt a strategy that results in a decline in accuracy, reflected in an increase in error-corrections and uncorrected errors, while older adults show a decline in speed. Previously, we mentioned that an increase in backspace use indirectly results in a decline in overall typing speed. The time that younger adults save by not slowing down over time, seems to disappear due to the increase in the number of letters that have to be typed, and the additional correction slowing. Given that the number of uncorrected mistakes also increased over time, we could argue that younger adults show a larger decline in performance over time than older adults. It is important to acknowledge that, dependent on the requirements of the task that is being performed, a decline in accuracy might even be more concerning than a decline in speed, because these mistakes eventually have a larger impact on the quality of the work.

Chapter 2 specifically focused on the relation between typing behaviour and mental fatigue in a controlled setting, establishing a link between typewriting markers and markers in EEG related to mental fatigue. In the third chapter of this thesis, we investigated whether the typewriting markers that were found to be susceptible to the effects of mental fatigue in the lab, showed similar patterns during prolonged task performance in the working environment. By analysing typewriting data of 22 office workers (*M*=48.1 year, *SD*=13.4) that regularly performed typewriting activities over six consecutive weeks, we found that, similar to our findings in the lab, both the percentage of backspaces and the interkey interval were sensitive to the effects of prolonged task performance. That is, office workers use the backspace key more often over time and leave more time between two subsequent keystrokes (interkey interval), suggesting that with time-on-task people become slower and perform less accurately.

Before we interpret these results in the context of this thesis, it is important to acknowledge that prolonged task performance in the office study (de Jong et al., 2020) is different from the prolonged task performance manipulation used in the lab setting (de Jong et al., 2018). That is, in the lab, we were able to monitor the participants' performance and enforce them to continuously perform a task. Hence, we knew that participants did not take breaks or switched tasks. In a real-life office setting however, prolonged task performance meant that workers were performing office work during which they also had to type. It is possible and even probable, that participants switched tasks or even took small breaks in between these typing sessions.

In general, the effects of prolonged task performance on typing changed throughout the day. In the morning, a decline in typing accuracy, but not in typing speed, was observed, while in the afternoon, both typing speed and typing accuracy showed a decline with time-on-task. These findings suggest that office workers are able to perform at stable speed during prolonged task performance in the morning, however, perhaps due to little recovery during the day, task demands might pile up, resulting in additional decline in typing speed in the afternoon. Similar results have been found in the lab, where the effects of mental fatigue were studied in isolation (Lorist et al., 2000). More specifically, when participants were instructed to perform fast, the researchers found a steady decline in accuracy during the first hour. After an hour participants adjusted their strategy and showed a decline in speed, as well.

In Chapter 3, we found no evidence for a general decline in typewriting performance over the days of the working week, providing additional proof that the effects of mental fatigue remained isolated to one working day, and suggesting that employees were able to recover from work-related demands overnight. Nonetheless, typewriting dynamics were subject to daily variations. That is, on Fridays, and especially on Mondays, typing speed increased and backspace use was higher than on the other days of the week. Several studies found that employees experienced their performance and work-related demands differently on Mondays. Some researchers argued that on Mondays employees need to recover from pleasurable activities that took place during the weekend (Binnewies et al., 2010; Fritz & Sonnentag, 2005), and other studies revealed that Mondays serve as a transition day between the work-free weekend and the start of a new workweek, during which a more negative mood and higher stress levels are being experienced (Croft & Walker, 2001; Devereux et al., 2011; Rook & Zijlstra, 2006). This transition might be reflected in the typewriting data, as well. On Mondays compared to the other days of the week, employees were faster and made more corrections by using the backspace. These changes in the speed-accuracy patterns show similarities with previous research on the speed-accuracy trade-off: participants who experience increased stress levels, generally perform faster, but make more mistakes (Wickens & Hollands, 2000).

Using data to shape a more optimal working environment

In the first two chapters, we showed that typing indices could be used to describe subtle changes in speed and accuracy, both in the lab and in the working environment. These findings indicate that a keylogging tool could be used to monitor performance without interfering with regular working activities, and could in theory be used to provide feedback directly to the user. In order to be able to provide informative feedback, first, a baseline state to which alterations in behaviour can be compared, needs to be established. It is only possible to detect changes in a person's state, if typewriting patterns during this baseline state are known.

Given that tasks and subjective state of the employee, independent of prolonged task performance, can vary on a day to day, or even hour to hour basis, it is challenging to find a baseline state to compare the current state of the employee with. In Chapter 2, for example, we showed that typing speed during words that were incongruent with the sentence was significantly lower compared to words that were a good fit with the sentence. Although these effects did not interact with time-on-task, changes in task difficulty could influence typewriting patterns. In order to provide an employee with informative feedback, the tool would need information about the task someone is working on or the algorithm would need to be trained on large amounts of data to determine this by itself.

When a baseline state is established, and changes in performance can be detected reliably, there are several ways to provide feedback to the user. One possibility is to use the tool to provide feedback on the employees' individual cognitive performance, for instance by suggesting to take a break or switch tasks in order to recover from the work-related demands (Helton & Russell, 2017). However, due to inter-personal, environmental, and task variations, it might be difficult to provide reliable feedback in real-time. On top of that, previous experimental studies showed that feedback uncertainty decreases feedback use and slows down behavioural adaptation (Wei & Körding, 2010). In case of the keylogging tool, this would mean that if the user would clearly experience the subjective feelings of mental fatigue or notice the behavioural effects of mental fatigue (e.g., increase in errors), there is a larger chance that feedback will be incorporated than if there are no signs that fatigue is increasing. Given that subjective feelings of fatigue are usually not a good indicator of mental fatigue (Zhang et al., 2011), and the ability to monitor one's performance decreases over time (Lorist et al., 2005), fatigued individuals might perceive the feedback as less reliable and are less likely to use the feedback they received from such a tool. In this case, the tool would not fully serve its purpose and could even become a stressor, negatively influencing job performance.

Feedback does not necessarily have to be provided in real-time in order to be informative to the user. Another option is to provide feedback on an individual level to help employees realize a more optimal work-break schedule that is complementary with their individual state and specific work-related demands. An example of such an implementation is the study described in Chapter 3 of this thesis, but on a more individual level. Here, we compared typing dynamics in the morning to the afternoon, and over the different days of the week. By comparing behaviour dynamics over several weeks, the tool could for instance help decide when a specific employee should work on tasks that need high accuracy or when it is better to work on easy tasks. By averaging keylogging activity over several weeks under specific conditions, noise decreases, resulting in feedback that is less prone to errors than real-time feedback. This way, the tool could enable employees to make informed decisions about their work-rest schedule and planning specific tasks.

Besides implementations focused on the individual, the tool could also be used to test effectiveness of specific interventions in the working environment. An example of such an implementation would be evaluating the effectiveness of a 6-hour working day compared to a working day of 8 hours. Previously, Akerstedt and colleagues (2001) investigated the effects of a shortened work week in nurses over the course of a year using a questionnaire before and a year after the implementation of the 6-hour working week. They showed that shortened schedules improved subjective sleep quality, mental fatigue, and health. Although this study shows long-term positive effects on health and wellbeing, short-term, daily variations in performance have not been taken into account. When addressing the effectiveness of similar schedules in office workers, a major advantage of implementing the keylogging tool would be that it measures behavioural dynamics without demanding the effort of employees and without influencing employees by reminding them of the intervention by asking them to keep a diary or fill out daily questionnaires.

Take a break or switch tasks

In the first two chapters of this thesis, we confirmed that using a keylogging tool to monitor behaviour in a working environment could provide information about behaviour dynamics which could eventually help shape a more optimal working environment. From previous studies, we know that mental fatigue is more likely to develop if the homogeneity of a task is high, specifically homogeneity regarding the cognitive functions involved (Robinson & Bills, 1926). Besides taking breaks, switching to different tasks depending on other cognitive functions could therefore relieve the effects of mental fatigue.

This can be illustrated by a study of Helton and Russell (2017). In their experiments, participants either had to continuously perform a spatial vigilance task, perform a spatial vigilance task with the interruption of a different task (verbal 1-back or 3-back, spatial 1-back or 3-back), or perform a spatial vigilance task interrupted by a break. They found that performance on the vigilance task improved after taking a break and declined during continuous task performance. Performing a different task in the break improved post-interruption performance on the vigilance task compared to continuous performance. However, post-interruption performance differed based on the tasks that were performed during the break. If the break-task was more similar to the vigilance task, there was a larger transfer of mental fatigue to the spatial vigilance task, resulting in larger decrements in post-interruption performance. To fully recover, it is therefore important to take breaks during a working day.

Taking breaks in the working environment is often accompanied by the consumption of coffee or tea. Besides the relief of taking a break from work, the pharmacological effects of caffeine also improve mood, alertness and productivity (Lieberman et al., 1987). Caffeine is the most used psychoactive stimulant worldwide. Around eighty percent of the world's population is consuming caffeine-containing substances on a daily basis, where it is mostly ingested as coffee or tea, although other products, such as chocolate and soft drinks, also contain caffeine. In habitual coffee drinkers, caffeine doses as low as 32 mg have been found to decrease reaction times (RTs) accompanied by improved accuracy (Lieberman et al., 1987; for a review see Lorist and Tops, 2003). The effects of caffeine have been found to be most pronounced when attentional control is reduced, for instance during or after prolonged task performance (Lorist et al., 1994). Caffeine consumption therefore seems a specifically useful method to maintain performance during prolonged office work.

In Chapter 4, we investigated whether the previously described effects of caffeine can be ascribed to improved processing of behaviourally relevant information or whether these effects are general, improving processing of all incoming stimuli. This distinction is especially relevant for the working environment, given that there is lots of incoming information in the working environment, and focusing attention towards particular stimuli or tasks is important. For instance, finishing a task because the deadline is approaching, while resisting the urge to join colleagues in the coffee corner. Although one could argue what is more rewarding.

In order to investigate the working of caffeine on our attention system, we conducted a two-session lab-experiment in which thirty-one participants (*M*=22.0 year, *SD*=3.6) performed a cued-reward Stroop task (see Chapter 4 for full task description). During each trial, participants were presented with a cue that indicated whether a colour-word Stroop stimulus would be presented, and whether the participants could earn a reward based on their performance (manipulation of behavioural relevance). Before each session, the participants either received decaffeinated coffee with caffeine or with a placebo (3 mg/kg body weight).

During the experiment, we measured RTs and correct responses as indices of performance, and recorded electrical brain activity using EEG to assess changes in specific attentional processes. We used several brain activity markers reflecting these processes. First, sustained alpha power was used as a marker of neural arousal. Modulations in alpha power have previously been related to cortical activation and excitability (Scheeringa et al., 2012). Second, the contingent negative variation (CNV) and event-related alpha power were used as markers of more specific attentional-related processes. Where the CNV reflects the level of arousal in anticipation of a stimulus or event (Brunia et al., 2012; Hillyard, 1969), and event-related alpha power can be used as a marker of information sensitivity of the visual cortex (Jensen & Mazaheri, 2010).

Using these markers of behavioural performance and brain activity, we were able to find differences in attentional processes during cued-reward Stroop-task performance as a result of ingesting caffeine compared to placebo. In the placebo-condition, we found that both neural arousal while preparing for a stimulus (↑CNV) and information sensitivity of the visual areas of the brain before presentation of the stimulus (↓alpha power) increased, specifically if performance on the experimental task could potentially lead to a reward. These neural enhancements were followed by improved processing of the colour-word Stroop stimulus and improved behavioural performance on the task (↑correct responses; ↓ RTs). This result is already interesting in itself, given that it indicates that people focus their attention more to tasks that are behaviourally relevant (replication of Van den Berg et al. (2014). However, we found that caffeine modulates this process. If participants received caffeine, on top of a general arousing effect of caffeine, which was visible in an average decrease in alpha power, widely spread over the brain, event-related attentional processes were only enhanced for stimuli that could potentially lead to reward, indicating that caffeine can enhance attention towards specific information.

Thus, our findings in Chapter 4 suggest that caffeine not only increases general arousal, making perception more susceptible to stimuli in our environment, but also specifically enhances the processing of stimuli that are behaviourally relevant (van den Berg et al., 2020). Increasing our knowledge on the influence of caffeine on these fundamental processes is important for our understanding of its effects in daily life. Specifically the manner in which caffeine works on our attention system could influence when and in what environment to use caffeine as a stimulant. Our dynamic working environment provides lots of incoming information of which a large part is not necessarily important for us to be able to interact with our environment (e.g., colleagues talking just outside your office). If a stimulant would result in a general increase in arousal, attention towards all stimuli, including the distracting ones, would be enhanced. This would not cause any problems in a controlled environment, such as the lab, where there is little distracting information. However, in a dynamic working environment with all kinds of distractors, using a general stimulant would not provide an advantage. Our study now showed that caffeine would enhance attention towards specific relevant stimuli, which indicates that caffeine intake would provide an advantage in a turbulent working environment.

Part II:

Responsible research and innovation

In the first part of this thesis, we discussed the technical usability of monitoring typing indices to detect mental fatigue. Besides practical issues that need to be considered when implementing such a tool (See Using data to shape a more optimal working environment), developing a new tool to monitor behaviour with or without the aim to intervene in people’s life also has major ethical implications. It is therefore important to think about what impact this technology might have on individuals who are targeted as potential users or even on society as a whole (Grunwald, 2014; Stilgoe et al., 2013). Responsible research and innovation is a field of science that aims to highlight these socio-ethical issues in research and innovation practices. In the past decade, new knowledge and guidelines have been developed that empower researchers to incorporate the researcher's responsibility throughout the innovation process (Stahl, 2013; Stilgoe et al., 2013), focusing on anticipation of (un)foreseen ethical qualms, reflexivity on one's own role, inclusion of a diversity of perspectives, and responsiveness to societal needs.

These developments in responsible research and innovation helped researchers to become more aware of their role in the implementation of new technologies, and reflect on the impact of their innovations. However, still, three major problems remain: compartmentalization of focus, generalization over different contexts, and vagueness of responsibilities of the stakeholders involved in the design and implementation process (Efstratiou et al., 2007; Kortuem et al., 2007; Leclercq-Vandelannoitte, 2017; Palm, 2009). In order to overcome these problems in the projects of SPRINT@Work (See Appendix I for more information about SPRINT@Work), the researchers conducted a context-specific approach of doing ethics during the development and small-scale implementation of technologies in the workplace (Chapter 5).

Chapter 5 described how this context-specific approach of ethics was conducted for two case studies. One of these case studies focused on a wearable non-invasive core thermometer to monitor the core temperature of firefighters. Because firefighters do not have time nor the opportunity to monitor the data from the wearable core thermometer when they enter a fire, they would benefit from sharing these data with their captains. However, legally, employers or supervisors, such as firefighter captains, are not allowed to access personal data of their subordinates (European Union, 2016; GDPR, 2016; Arora, 2019).

This disbalance between what is actually beneficial for the firefighters health, and the regulations that are meant to protect them is a major problem when implementing new technologies in the working environment. Given that the law not yet protects the user in these fiduciary relationships (GDPR, 2016), it is important to already acknowledge these design-use dynamics in the design phase of a new technology, and come up with solutions that could help overcome this gap in the law. Although some researchers already stand for changing the law for fiduciary relationships (Arora, 2019), this would be a long and arduous process. Even if the law would change on this matter, it would still be important to define in which situations data sharing is condoned and with whom sharing data is really necessary. In the firefighter case, an agency-based approach of privacy was proposed (Manson & O'Neill, 2007), in which the user is enabled to share their health-data with data controllers (e.g., their firefighter captain). The outcome of this agency-based approach would depend on the context (where and with whom) in which the sensor technology is implemented (Chapter 5). The agency-based approach would ask for a discussion with all the stakeholders involved and for documentation of the agreements.

This practical example endorses previous findings which show that, although the existing legal framework is intended to protect workers, in some cases, users are not necessarily protected by it (Arora, 2019) or do they want to be protected in this manner (Spook et al., 2019). This is not only the case for sensor technologies that are developed for use in the working environment, where there is a clear dependent relationship between the worker and his or her employer. The rapid development of health-related technologies that are being used to support us during our daily life activities resulted in a change of players that are involved in the exchange of confidential data. Although the players have changed, the law still only protects specific confidential relationships, such as general practitioners, while other important fiduciary relationships, such as data controllers that deal with personal (health) data, have no specific basis in the law (Arora, 2019). This not only makes it difficult to share data with these new data controllers, but also makes it difficult to really protect the individual who wants to share or even benefits from sharing his or her personal data.

A call for practical and adaptive ethics guidelines

Responsibilities regarding the reflection on risks of new technologies get indistinguishable during the transition between the design and implementation phase (Jakobsen et al., 2019). More specifically, although the researchers might have reflected on the impact of a new technology, after the development phase, responsibilities shift towards the user or organizations that implement the technology, who do not necessarily reflect on possible ethical and societal risk, and primarily focus on productivity or increasing product acceptance (Chatfield et al., 2017; Leclercq-Vandelannoitte, 2017). Providing a smooth responsible transition from development to implementation, using a multi-stakeholder approach may help to overcome this problem of compartmentalization. The cases described in Chapter 5 illustrate how a context-specific approach of ethics could protect the user from harm, but could also help increase acceptance of a new technology. Until now very little research on ethics practices in the assessment of the interplay between values inherent in the design and implementation of new technologies has been done (Fisher et al., 2015). In order to be able keep tools and guidelines aligned with technological advancements, more multidisciplinary teams involved in innovation practices should publish their solutions of doing ethics in a responsible way. Given the independent and objective nature of science compared to business, where the primary focus lies on productivity and product acceptance, researchers should take on this responsibility.

Chapter 7

Nederlandse samenvatting
Dutch summary

Als mensen langere tijd achter elkaar bepaalde werkzaamheden uitvoeren, kunnen zij vermoeid raken. Wanneer mensen mentaal vermoeid raken hebben zij vaak geen zin meer om door te gaan met de taak waar ze mee bezig zijn, kunnen zij hun aandacht minder goed bij de taak houden en voeren zij de taak minder goed uit. Ze worden bijvoorbeeld langzamer, maken meer fouten en detecteren hun fouten minder snel en minder vaak. Dit is voornamelijk een probleem als mensen onder tijdsdruk werken, zoals bijvoorbeeld in een werkomgeving. Fulltime werknemers in Nederland werken 8 uur per dag, 5 dagen per week. Er zijn dus voldoende mogelijkheden vermoeid te raken. Door af en toe een pauze te nemen, kunnen werknemers de effecten van vermoeidheid voorkomen of tegengaan. Eerder onderzoek heeft echter laten zien dat mensen over het algemeen niet goed kunnen inschatten wanneer zij vermoeid raken en dus ook niet wanneer ze moeten pauzeren. Tijdens mijn promotieonderzoek heb ik daarom onderzoek gedaan naar methoden om vermoeidheid te detecteren en naar methoden om vermoeidheid tegen te gaan. Op deze manier kunnen we werknemers helpen om meer optimale werkomstandigheden te creëren.

Ongeveer 40% van de werknemers in Nederland werkt meer dan 6 uur per dag met een computer. Dit biedt mogelijkheden om op een objectieve manier het effect van mentale vermoeidheid op prestatie te detecteren. Zowel snelheid als accuraatheid van gedrag kunnen namelijk op basis van patronen in typgedrag worden bepaald. Typsnelheid kan bijvoorbeeld worden bepaald op basis van de tijd tussen twee toetsaanslagen of door de duur van een toetsaanslag te meten. Door de complexiteit van taken waar men in een werkomgeving mee te maken heeft is het bepalen van de mate van accuraatheid iets lastiger. Grammaticale fouten en tikfouten zijn te detecteren met behulp van algoritmes die ook zijn ingebouwd in verschillende woordverwerkers, zoals Microsoft Word. Dit type fouten reflecteert echter maar gedeeltelijk de accuraatheid van een tekst. Werknemers kunnen daarnaast woorden typen die niet goed in de zin passen of een tekst produceren waarvan de inhoud niet klopt. Deze vorm van accuraatheid is moeilijker te bepalen, om nog maar te zwijgen over de inbreuk op de privacy van de werknemers als een programma continu de inhoud van hun typgedrag controleert. Accuraatheid kan echter ook indirect worden bepaald op basis van het gedrag dat mensen vertonen als zij fouten detecteren en corrigeren. Correctiegedrag kan tijdens het typen worden gemeten op basis van backspaceaanslagen.

In Hoofdstuk 2 beschrijven we een studie waarin we in een experimentele setting hebben onderzocht of patronen in typgedrag gevoelig zijn voor de effecten van vermoeidheid. Om dit te kunnen onderzoeken, voerden vierentwintig deelnemers tussen 18 en 30 jaar en vierentwintig deelnemers tussen 50 en 67 jaar 2 uur lang een gestandaardiseerde typ-taak uit. Tijdens deze taak kregen deelnemers een zin te zien. Het laatste woord van de zin verscheen op het beeldscherm nadat het eerste deel van de zin was getypt, en kon goed of niet goed bij de rest van de zin passen (Bijvoorbeeld: “Het brood werd door de bakker gebakken/bedreigd.”). De deelnemers werden geïnstrueerd om de zin en het woord zo snel mogelijk en foutloos te typen en eventuele fouten te corrigeren met de backspacetoets. Tijdens het experiment werden toetsaanslagen geregistreerd en werd hersenactiviteit gemeten met behulp van elektro-encefalografie (EEG). Op deze manier konden we achteraf bepalen hoe snel en accuraat deelnemers waren en hoe patronen in typgedrag samenhangen met hersenactiviteit gerelateerd aan mentale vermoeidheid. We waren met name geïnteresseerd in patronen in hersenactiviteit die aandachtsprocessen reflecteren, omdat eerder onderzoek herhaaldelijk heeft aangetoond dat vermoeidheid en aandacht sterk aan elkaar gerelateerd zijn.

In lijn met onze verwachtingen vonden we, door te kijken naar patronen in het EEG signaal, dat de deelnemers hun aandacht minder goed bij de taak hielden naarmate zij langer aan de taak werkten. Deze achteruitgang in aandacht was gerelateerd aan veranderingen in typgedrag. Als de deelnemers vermoeid raakten, hielden zij hun aandacht minder goed bij de tekst en typten zij langzamer. Typsnelheid lijkt dus gerelateerd aan vermoeidheid.

Er zijn verschillende oorzaken die ten grondslag liggen aan dit effect van vermoeidheid op typsnelheid. Allereerst gingen de deelnemers over het algemeen langzamer typen wanneer zij vermoeid raakten. Daarnaast zagen we dat jonge deelnemers naarmate het experiment vorderde meer fouten maakten en deze fouten in ieder geval gedeeltelijk corrigeerden. Dit zorgde er indirect voor dat zij meer tijd nodig hadden om de zin te typen. Het corrigeren van een tekst kost nou eenmaal tijd: mensen produceren eerst de incorrecte tekst, vervolgens moeten zij deze tekst weghalen met de backspacetoets en daarna moeten zij de correcte tekst produceren. Kortom, er zijn veel meer typbewegingen. We zagen echter ook dat deelnemers tijdens het corrigeren van hun fouten langzamer typten dan tijdens de tekst die zij in één keer goed typten. Het maken en vervolgens corrigeren van fouten heeft dus op twee manieren een negatieve impact op typsnelheid.

Op welke manier vermoeidheid snelheid, accuraatheid of allebei beïnvloedt verschilt per persoon en situatie. Uit eerder onderzoek weten we dat ouderen het vaak belangrijker vinden om accuraat te presteren dan om een taak snel uit te voeren, terwijl jongeren zowel snel als accuraat willen presteren. Eenzelfde patroon zagen we in ons onderzoek. De oudere groep leek vaker voor een strategie te kiezen waarbij ze gedurende het experiment langzamer werden, maar wel even accuraat bleven presteren. De jongeren lieten daarentegen een achteruitgang in zowel accuraatheid als snelheid zien.

In Hoofdstuk 2 hebben we dus laten zien dat we op basis van typgedrag iets kunnen zeggen over het verloop van vermoeidheid over tijd. Echter, deze studie vond plaats in een gecontroleerde laboratoriumsituatie, waarin we ervoor kunnen zorgen dat alle deelnemers dezelfde taken uitvoeren, tussendoor geen pauze nemen en niet worden gestoord. In een alledaagse werkomgeving moeten we rekening houden met allerhande factoren die de mate van vermoeidheid en daarnaast ook het typgedrag kunnen beïnvloeden. Zo voeren werknemers tijdens hun normale werkzaamheden meestal verschillende taken uit. Doordat verschillende taken een andere cognitieve belasting vereisen, kan vermoeidheid zich binnen deze taken anders manifesteren. Daarnaast is het goed mogelijk en zelfs waarschijnlijk dat werknemers op verschillende momenten pauze nemen. Ook dit beïnvloedt de manier waarop vermoeidheid zich uit.

Met deze informatie in het achterhoofd hebben we in **Hoofdstuk 3** onderzocht of de effecten van vermoeidheid op typgedrag die we vonden in de laboratoriumsetting ook zichtbaar zijn in een werkomgeving. Om dit te kunnen onderzoeken hebben we gedurende zes weken het typgedrag van kantoormedewerkers gemonitord. Ondanks dat we geen controle hadden over de werkzaamheden en de omgeving van de deelnemers, vonden we in deze studie patronen in typegedrag die overeenkwamen met de effecten van vermoeidheid bij de jongere deelnemers in de laboratoriumstudie. Werknemers werden over het algemeen langzamer en gebruikten de backspacetoets vaker als zij gedurende langere tijd typwerkzaamheden uitvoerden.

Deze patronen in typgedrag als gevolg van vermoeidheid hangen af van het moment van de dag. In de ochtend gebruikten de werknemers steeds vaker de backspacetoets terwijl hun typsnelheid stabiel bleef. In de middag zagen we dat de deelnemers naast het vaker corrigeren van hun tekst ook langzamer werden. Het lijkt er dus op dat vermoeidheid in de middag sterker tot uiting kwam in het typegedrag. Dit komt mogelijk doordat werknemers in de middag onvoldoende zijn hersteld van de vermoeidheid die in de ochtend al is opgebouwd. Het is opvallend dat de effecten van vermoeidheid niet sterker werden naarmate de werkweek vorderde. Hieruit maken we op dat werknemers in ieder geval goed kunnen herstellen van de mentale belasting die zij hebben ervaren gedurende de vorige dag.

Typgedrag is dus gevoelig voor de effecten van vermoeidheid, zowel in een gecontroleerde laboratoriumsituatie als in een turbulente werkomgeving. Een methode waarmee we typgedrag kunnen monitoren in de werkomgeving zou gebruikt kunnen worden om werknemers te ondersteunen bij het nemen van beslissingen over hun werkomstandigheden. Er zijn verschillende manieren waarop een dergelijke methode de werknemer kan ondersteunen.

Directe feedback aan de werknemer. Real-time typgedrag kan informatie geven over de mate waarin een werknemer op dat moment vermoeid is. Op basis van deze informatie kan direct advies worden gegeven over het nemen van een pauze of het wisselen van taak. Doordat werknemers in een werkomgeving worden beïnvloed door allerlei factoren die naast vermoeidheid het typgedrag beïnvloeden, is het waarschijnlijk dat deze feedback niet altijd volledig betrouwbaar is. Het kan bijvoorbeeld voorkomen dat een werknemer net van taak wisselt, waardoor typsnelheid omlaag gaat en het aantal backspaceaanslagen omhoog gaat. Op basis van deze gedragspatronen kan het advies worden gegeven om een pauze te nemen of om van taak te wisselen, terwijl de werknemer eigenlijk helemaal niet vermoeid is.

Feedback op basis van grote hoeveelheden data. Dit probleem kan worden opgelost door feedback te geven op basis van typgedrag dat gedurende een langere periode is gemeten. Door werknemers over langere tijd, bijvoorbeeld weken of maanden, te monitoren, kan er worden onderzocht op welke momenten werknemers vaker patronen in typgedrag vertonen die overeenkomen met vermoeidheid. Dus in plaats van directe feedback te geven op het gedrag van een werknemer op dat moment, kan er een gepersonaliseerd advies worden gegeven over de taakindeling en de momenten waarop iemand het beste pauze kan nemen.

Onderzoek op de werkvloer. Als laatste kan het monitoren van typgedrag ook gebruikt worden bij onderzoek naar vermoeidheid in een werkomgeving. Het meeste onderzoek naar vermoeidheid dat in de werkomgeving wordt uitgevoerd is gebaseerd op dagboekverslagen of vragenlijsten. Door gebruik te maken van een methode die op basis van patronen in typgedrag een indicatie geeft over de mate van vermoeidheid, kan er onderzoek worden gedaan zonder dat de normale werkzaamheden worden verstoord. Op deze manier kan bijvoorbeeld de effectiviteit van een 6-urige werkweek worden onderzocht.

Deze methoden kunnen worden gebruikt om meer optimale werkomstandigheden te creëren voor werknemers door er bijvoorbeeld voor te zorgen dat werknemers op de juiste momenten rust nemen. Het nemen van pauzes gaat vaak gepaard met het drinken van koffie of thee. De cafeïne die men hierdoor binnen krijgt verbetert de stemming, alertheid en productiviteit van werknemers bovenop de positieve effecten van het nemen van de pauze. Eerder onderzoek heeft laten zien dat cafeïne extra goed werkt als iemand vermoeid is. Voordat we ons na een ochtend hard werken direct naar de koffiecorner begeven om een bakje koffie te drinken met onze collega's, is het echter wel belangrijk om meer te weten te komen over de manier waarop cafeïne stimulerend werkt. Heeft cafeïne een positief effect op de aandacht voor informatie die relevant is voor de taak of wordt álle informatie die binnenkomt beter verwerkt? Aangezien werknemers in een werkomgeving worden blootgesteld aan veel informatie kan het laatste juist een negatief effect hebben op het uitvoeren van werkzaamheden. Dit zou namelijk betekenen dat zij juist méér worden afgeleid als zij een kopje koffie drinken.

In **Hoofdstuk 4** hebben we onderzoek gedaan naar de effecten van cafeïne op aandacht. Daarbij hebben we specifiek onderzocht of de positieve effecten van cafeïne kunnen worden verklaard door een verbetering in de aandacht voor relevante informatie of doordat alle informatie beter binnenkomt. Om dit te kunnen onderzoeken, hebben eenendertig proefpersonen op twee verschillende momenten een taak uitgevoerd, waarbij zij voorafgaand aan de taak een kop koffie kregen, één keer met cafeïne en één keer zonder cafeïne. Tijdens het experiment hebben we hersenactiviteit gemeten door middel van EEG.

Tijdens het experiment werd 1200 keer het woord ROOD, BLAUW, GROEN of GEEL gepresenteerd. Dit woord had ofwel dezelfde kleur als het geschreven woord (congruent), ofwel een andere kleur dan het geschreven woord (incongruent). Nadat het woord in beeld verscheen, moesten de deelnemers zo snel mogelijk aangeven in welke kleur het woord was geschreven. Voordat het woord op het scherm werd gepresenteerd, kregen de deelnemers een symbool op het scherm te zien. Dit symbool vertelde hen of zij een beloning konden krijgen als zij op tijd de juiste kleur van het geschreven woord zouden aangeven. Hierdoor konden we onderzoeken op welke manier cafeïne aandacht richting relevante versus minder relevante informatie beïnvloedt. De taak wordt namelijk een stuk relevanter als er een beloning in het vooruitzicht ligt.

In lijn met eerder onderzoek zagen we dat deelnemers hun aandacht beter focusten op de taak en beter presteerden als ze geld konden verdienen. Cafeïne beïnvloedt dit proces. Wanneer deelnemers voorafgaand aan het experiment koffie met cafeïne hadden gedronken, richtten respondenten hun aandacht alleen beter op de taak als zij geld konden verdienen, maar niet wanneer zij geen konden verdienen. Buiten de algemeen stimulerende effecten van cafeïne die ervoor zorgen dat mensen zich beter en actiever voelen, zorgt cafeïne er dus ook nog eens voor dat mensen hun aandacht beter focussen op belangrijke informatie.

Dit onderzoek naar fundamentele processen in het brein lijkt op het eerste gezicht misschien niet heel relevant voor ons eigen leven. De manier waarop cafeïne werkt geeft ons echter belangrijke informatie over wanneer en op welke manier we cafeïne het best kunnen gebruiken. Een stimulerend middel dat ervoor zorgt dat alle informatie beter wordt verwerkt, zou misschien geen problemen opleveren in een gecontroleerde omgeving, zoals het laboratorium, maar in een dynamische werkomgeving met allerlei afleiders zou het eerder een negatieve impact kunnen hebben op onze productiviteit. Ons onderzoek toont nu aan dat cafeïne de aandacht richting relevante informatie kan vergroten, terwijl aandacht richting informatie die niet relevant is voor de taak niet wordt vergroot. Een kopje koffie met cafeïne levert in een werkomgeving dus juist een extra voordeel op.

De focus van de eerste drie inhoudelijke hoofdstukken van mijn dissertatie (Hoofdstuk 2 t/m 4) ligt vooral op het monitoren en verbeteren van het gedrag van werknemers. Buiten de praktische problemen die bij het implementeren van een dergelijke methode op de werkvloer komen kijken, heeft het monitoren en beïnvloeden van gedrag ook ethische implicaties, al helemaal in een werkomgeving waar men te maken heeft met een afhankelijke werkgever-werknemer relatie. In de afgelopen decennia zijn er al grote stappen gezet in de ontwikkeling van kennis en richtlijnen die onderzoekers en ingenieurs ondersteunen bij het verantwoord ontwikkelen van nieuwe producten. In **Hoofdstuk 5** laat ik samen met andere onderzoekers van het project SPRINT@Work[3] aan de hand van twee voorbeelden zien hoe we, naast de ontwikkeling en implementatie van nieuwe technologieën in de werkomgeving, reflecteerden op de ethische en praktische problemen die bij de ontwikkeling van nieuwe technologieën komen kijken. Tijdens deze

---

[3] *Het huidige onderzoek is een onderdeel van het onderzoeksproject SPRINT@Work. Appendix IV geeft een volledige beschrijving van het project.*

reflectie gebruikten we een context-sensitieve benadering van ethiek die gedurende het gehele project, de ontwikkeling en het implementeren van technologieën in de werkomgeving, werd toegepast. Door het hele proces heen werd tijdens intervisiesessies met de onderzoekers en een ethicus geëvalueerd of het huidige juridische kader de implementatie van nieuwe technologieën voldoende dekt. Daarnaast onderzochten we op welke manier de technologieën de autonomie van de werknemers beïnvloeden.

Eén van de voorbeelden ging over de ontwikkeling van een draagbare sensor die gebruikt kan worden om de kerntemperatuur van brandweermannen te monitoren terwijl zij een brand blussen. Brandweermannen hebben niet de tijd of de mogelijkheid om tijdens het blussen van een brand te letten op de data die uit de sensor komt en daarom zou het voordelig zijn als de zij de data met de brandweercaptain zouden kunnen delen. Volgens de algemene verordening gegevensbescherming (AVG) mogen werkgevers en supervisors, zoals brandweercaptains, geen persoonlijke data van hun werknemers inzien. Dit bemoeilijkt de implementatie van deze sensor, die positieve effecten kan hebben op de gezondheid van de brandweermannen, in hun dagelijks werkleven.

Deze disbalans tussen wat beter is voor de gezondheid van de werknemer en de wetgeving die de werknemer dient te beschermen tegen onrecht, is een probleem bij het implementeren van nieuwe technologieën in de werkomgeving. Het feit dat de wetgeving nog niet voldoende rekening houdt met de rol van meerdere betrokken partijen tijdens de toepassing van nieuwe technologieën maakt het belangrijk om in een vroeg stadium van de ontwikkeling al na te denken over de manier waarop een technologie uiteindelijk gebruikt gaat worden. Op deze manier kan er op tijd worden begonnen met het onderzoeken van mogelijk alternatieve opties, waarmee problemen met de wetgeving kunnen worden vermeden. In het voorbeeld van de brandweermannen is uiteindelijke gekozen voor een methode waar de gebruiker expliciet de keuze krijgt om gezondheidsdata te delen

met data-controllers. De uitkomst van deze benadering hangt af van de context (waar en met wie) waarin de technologie wordt geïmplementeerd. Dit betekent niet dat de gebruiker zomaar data met iedereen kan delen. De methode is er juist op gericht dat er discussie is met alle mensen die met het proces te maken hebben en vraagt daarnaast om duidelijke documentatie van de afspraken.

Dit voorbeeld laat duidelijk zien dat de wetgeving die bedoeld is om werknemers te beschermen, in bepaalde omstandigheden niet beschermt. Dit is niet alleen het geval voor technologieën die ontwikkeld zijn voor het gebruik in de werkomgeving, waar er een duidelijke afhankelijke relatie is tussen de werkgever en werknemer. De ontwikkeling van gezondheid-gerelateerde apps die ontwikkeld worden om de werknemer te ondersteunen in het dagelijks leven hebben ervoor gezorgd dat er veel meer mensen betrokken zijn bij de uitwisseling van vertrouwelijke informatie. Deze data-controllers hebben volgens de wetgeving, in tegenstelling tot artsen en andere zorgverleners, geen rechten en geen plichten. Dit maakt het niet alleen moeilijk om op een verantwoorde manier data te delen met deze data-controllers, zoals in het voorbeeld van de brandweerman, maar maakt het ook moeilijker om de mensen te beschermen die op een onverantwoorde manier hun data delen met data-controllers, zoals zorgverzekeraars.

Conclusies

Binnen mijn promotieonderzoek heb ik op verschillende manieren onderzocht hoe we werknemers kunnen helpen bij het creëren van meer optimale werkomstandigheden. We hebben laten zien dat patronen in typgedrag gevoelig zijn voor de effecten van vermoeidheid, zowel in een experimentele setting als in een werkomgeving. Veranderingen in typgedrag kunnen gebruikt worden om feedback te geven aan de werknemer, maar bieden ook waardevolle informatie voor onderzoek in

de werkomgeving. Gedrag kan op deze manier continu worden gemonitord zonder dat reguliere werkzaamheden onderbroken hoeven te worden.
Naast het bekende positieve effect van pauzes, is cafeïne een goede manier om vermoeidheidseffecten tegen te gaan. Men kan aandacht beter focussen op belangrijke informatie in de omgeving na een kop koffie. Dit is een voordeel in een dynamische werkomgeving, waar het belangrijk is dat informatie die op dat moment belangrijk is voor de werkzaamheden, goed wordt verwerkt.

In het onderzoek naar technologieën die gebruikt kunnen worden om optimale werkomstandigheden te creëren, is het belangrijk dat er al tijdens de ontwikkelingsfase rekening gehouden wordt met de impact van deze technologieën op de mensen die ermee te maken krijgen. Het is van belang dat onderzoekers in multidisciplinaire teams hun verantwoordelijkheid nemen en een prominentere rol gaan spelen door ethiek een geïntegreerd onderdeel van het volledige onderzoeks- en innovatieproces te laten zijn.

References

Aarts, E., van Holstein, M., & Cools, R. (2011). Striatal dopamine and the interface between motivation and cognition. *Frontiers in Psychology, 2*(JUL), 1–11. https://doi.org/10.3389/fpsyg.2011.00163

Ackerman, P. L. (2011). Cognitive fatigue: Multidisciplinary perspectives on current research and future applications. (P. L. Ackerman (ed.)). *American Psychological Association*. https://doi.org/10.1037/12343-000

Akerstedt, T., Olsson, B., Ingre, M., & Holmgren, M. (2001). *J. Human Ergol.*, 30:197-202, 2001. 197–202.

Al Ameen, M., Liu, J., & Kwak, K. (2012). Security and privacy issues in wireless sensor networks for healthcare applications. *Journal of Medical Systems, 36*(1), 93–101. https://doi.org/10.1007/s10916-010-9449-4

Allender, S., Colquhoun, D., & Kelly, P. (2006). Competing discourses of workplace health. *Health, 10*(1), 75–93. https://doi.org/10.1177/1363459306058989

Arora, C. (2019). Digital health fiduciaries: protecting user privacy when sharing health data. *Ethics and Information Technology, 21*(3), 181–196. https://doi.org/10.1007/s10676-019-09499-x

Aryal, A., Ghahramani, A., & Becerik-Gerber, B. (2017). Monitoring fatigue in construction workers using physiological measurements. *Automation in Construction, 82*, 154–165. https://doi.org/10.1016/j.autcon.2017.03.003

Ashton, H., Millman, J. E., Telford, R., & Thompson, J. W. (1974). The effect of caffeine, nitrazepam and cigarette smoking on the contingent negative variation in man. *Electroencephalography and Clinical Neurophysiology, 37*(1), 59–71. https://doi.org/10.1016/0013-4694(74)90245-4

Baker, K., Olson, J., & Morisseau, D. (1994). Work practices, fatigue, and nuclear power plant safety performance. Human Factors: *The Journal of the Human Factors and Ergonomics Society, 36*(2), 244–257.

Barocas, S., & Nissenbaum, H. (2013). Big data's end run around anonymity and consent. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (pp. 44-75). https://doi.org/10.1017/CBO9781107590205.004

Bartlett, F. C. (1941). Fatigue following highly skilled work. *Nature, 147*(3736), 717–718. https://doi.org/10.1038/147717a0

Bartley, S., & Chute, E. (1949). Fatigue and Impairment in Man. *The Quarterly Review of Biology, 24*(1), 68. https://doi.org/10.1086/396874

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. *2000*. http://arxiv.org/abs/1506.04967

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bertelson, P., & Joffe, R. (1963). Blockings in prolonged serial responding. *Ergonomics, 6*(2), 109–116. https://doi.org/10.1080/00140136308930682

Bieker, F., Friedewald, M., Schiffner, S., Serna, J., Ikonomou, D., Rannenberg, K., Hansen, M., Obersteller, H., & Rost, M. (2016). A Process for Data Protection Impact Assessment Under the European General Data Protection Regulation. *Privacy Technologies and Policy, 1*, 21–37. https://doi.org/10.1007/978-3-319-44760-5

Bills, A. G. (1931). Blocking: a new principle of mental fatigue. *The American Journal of Psychology, 43*, 230–245. https://doi.org/10.2307/1414771

Binnewies, C., Sonnentag, S., & Mojza, E. J. (2010). Recovery during the weekend and fluctuations in weekly job performance: A week-level study examining intra-individual relationships. *Journal of Occupational and Organizational Psychology, 83*(2), 419–441. https://doi.org/10.1348/096317909X418049

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: an ERP study. Brain Research. *Cognitive Brain Research, 25*(1), 107–116. https://doi.org/10.1016/j.cogbrainres.2005.04.011

Boksem, M. A. S., Meijman, T., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology, 72*(2), 123–132. https://doi.org/10.1016/j.biopsycho.2005.08.007

Boksem, M. A. S., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews, 59*, 125–139. https://doi.org/10.1016/j.brainresrev.2008.07.001

Bonnefond, A., Doignon-Camus, N., Hoeft, A., & Dufour, A. (2011). Impact of motivation on cognitive control in the context of vigilance lowering: An ERP study. *Brain and Cognition, 77*(3), 464–471. https://doi.org/10.1016/j.bandc.2011.08.010

Bonvanie, A., Broekhuis, M., Janssen, O., Maeckelberghe, E., & Wortmann, J. (Hans) C. (2020). Health Self-Management Applications in the Work Environment: The Effects on Employee Autonomy. *Frontiers in Digital Health, 2*(July), 1–17. https://doi.org/10.3389/fdgth.2020.00009

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*(3), 345–370. https://doi.org/10.1007/BF02294361

Brainard, D. H. (1996). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Bravata, D. M., Smith-Spangler, C., Sundaram, V., Gienger, A. L., Lin, N., Lewis, R., Stave, C. D., Olkin, I., & Sirard, J. R. (2007). Using Pedometers to Increase Physical Activity and Improve Health A Systematic Review. *JAMA, 298*(19), 2296–2304. https://doi.org/10.1001/jama.298.19.2296

Brown, J. P., Martin, D., Nagaria, Z., Verceles, A. C., Jobe, S. L., & Wickwire, E. M. (2020). Mental Health Consequences of Shift Work: An Updated Review. *Current Psychiatry Reports, 22*(2), 1–7. https://doi.org/10.1007/s11920-020-1131-z

Brunia, C. H. M., van Boxtel, G. J. M., & Böcker, K. B. E. (2012). Negative Slow Waves as Indices of Anticipation: The Bereitschaftspotential, the Contingent Negative Variation, and the Stimulus-Preceding Negativity. *The Oxford Handbook of Event-Related Potential Components* (January 2021). https://doi.org/10.1093/oxfordhb/9780195374148.013.0108

Cappell, K. A., Gmeindl., L., & Reuter-Lorenz, P. A. (2010). Age Differences in Prefrontal Recruitment During Verbal WM Maintenance Depend on Memory Load. *Cortex, 46*(4), 462–473. https://doi.org/10.1016/j.cortex.2009.11.009.Age

Chan, M. (2011). Fatigue: The most critical accident risk in oil and gas construction. *Construction Management and Economics, 29*(4), 341–353. https://doi.org/10.1080/01446193.2010.545993

Chatfield, K., Borsella, E., Mantovani, E., Porcari, A., & Stahl, B. C. (2017). An investigation into risk perception in the ICT industry as a core component of responsible research and innovation. *Sustainability (Switzerland), 9*(8). https://doi.org/10.3390/su9081424

Choi, B., Hwang, S., & Lee, S. H. (2017). What drives construction workers' acceptance of wearable technologies in the workplace?: Indoor localization and wearable health devices for occupational safety and health. *Automation in Construction, 84*(May), 31–41. https://doi.org/10.1016/j.autcon.2017.08.005

Conger, S., Pratt, J. H., & Loch, K. D. (2013). Personal information privacy and emerging technologies. *Information Systems Journal, 23*(5), 401–417. https://doi.org/10.1111/j.1365-2575.2012.00402.x

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215. https://doi.org/10.1038/nrn755

Croft, G. P., & Walker, A. E. (2001). Are the Monday blues all in the mind? The role of expectancy in the subjective experience of mood. *Journal of Applied Social Psychology, 31*(6), 1133–1145. https://doi.org/10.1111/j.1559-1816.2001.tb02666.x

Csathó, Á., Van Der Linden, D., Hernádi, I., Buzás, P., & Kalmár, G. (2012). Effects of mental fatigue on the capacity limits of visual attention. *Journal of Cognitive Psychology, 24*(5), 511–524. https://doi.org/10.1080/20445911.2012.658039

Cummings, M. L. (2006). Integrating ethics in design through the value-sensitive design approach. *Science and Engineering Ethics, 12*(4), 701–715. https://doi.org/10.1007/s11948-006-0065-0

Damman, O. C., Van der beek, A. J., & Timmermans, D. R. M. (2015). Employees are ambivalent about health checks in the occupational setting. *Occupational Medicine, 65*(6), 451–458. https://doi.org/10.1093/occmed/kqv048

Davis, D. R. (1946). The disorganization of behaviour in fatigue. Journal of Neurology, *Neurosurgery, and Psychiatry, 9*(1), 23–29. https://doi.org/10.1136/jnnp.9.1.23

Davis, J. M., Zhao, Z., Stock, H. S., Mehl, K. A., Buggy, J., & Hand, G. A. (2003). Central nervous system effects of caffeine and adenosine on fatigue. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology, 284*(2), R399—404. https://doi.org/10.1152/ajpregu.00386.2002

de Jong, M., Bonvanie, A. M., Jolij, J., & Lorist, M. M. (2020). Dynamics in typewriting performance reflect mental fatigue during real-life office work. *PLOS ONE, 15*(10), e0239984. https://doi.org/10.1371/journal.pone.0239984

de Jong, M., Jolij, J., Pimenta, A., & Lorist, M. M. (2018). Age Modulates the Effects of Mental Fatigue on Typewriting. *Frontiers in Psychology, 9*(July), 1–15. https://doi.org/10.3389/fpsyg.2018.01113

De Jongh, J., & Mcdougal, F. (2014). Gezonde Leefstijl en Vitaliteit op de Werkvloer: een kwantitatief onderzoek onder werknemers in opdracht van Convenant Gezond Gewicht. Available online at: https://www.stvda.nl/~/media/Files/Stvda/Thema/Arbeidsomstandigheden/vitaliteit/20140925-ruigrok-Netpanel.ashx

DeCew, J. (2018). Privacy. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (2015)*. Metaphysics Research Lab, Stanford University.

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences, 95*(24), 14529–14534. https://doi.org/10.1073/pnas.95.24.14529

Delorme, A., & Makeig, S. (2004). EEGLAB: an open sorce toolbox for analysis of single-trail EEG dynamics including independent component anlaysis. *Journal of Neuroscience Methods, 134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Devereux, J., Rydstedt, L. W., & Cropley, M. (2011). An exploratory study to assess the impact of work demands and the anticipation of work on awakening saliva cortisol. *Psychological Reports, 108*(1), 274–280. https://doi.org/10.2466/09.14.17.PR0.108.1.274-280

Doorn, N. (2012). Exploring Responsibility Rationales in Research and Development (R&D). *Science Technology and Human Values, 37*(3), 180–209. https://doi.org/10.1177/0162243911405344

Doorn, N., & van de Poel, I. (2012). Editors' Overview: Moral Responsibility in Technology and Engineering. *Science and Engineering Ethics, 18*(1), 1–11. https://doi.org/10.1007/s11948-011-9285-z

Dunwiddie, T. V, & Masino, S. a. (2001). T He R Ole and R Egulation of a Denosine. Annual *Review of Pharmacology and Toxicology, 41*(1), 145–174.

Dwork, C. (2006). Differential Privacy. In M. Bugliesi, B. Preneel, V. Sassone, & I. Wegener (Eds.), *Automata, Languages and Programming* (pp. 1–12). Springer, Berlin, Heidelberg.

Efstratiou, C., Davies, N., Kortuem, G., Finney, J., Hooper, R., & Lowton, M. (2007). Experiences of designing and deploying intelligent sensor nodes to monitor hand-arm vibrations in the field. In Proceedings of the 5th international conference on mobile Systems, applications and services (MobiSys '07). *Association for Computing Machinery*, 127–138. https://doi.org/10.1145/1247660.1247677

Engelmann, J. B., & Pessoa, L. (2007). Motivation Sharpens Exogenous Spatial Attention. *Emotion, 7*(3), 668–674. https://doi.org/10.1037/1528-3542.7.3.668

Ennis, G. E., Hess, T. M., & Smith, B. T. (2014). Implications for Cognitive Engagement in Older Adulthood. *Psychology and Aging, 28*(2), 495–504. https://doi.org/10.1037/a0031255.

Eurofound. (2015). Workplace practices – Patterns, performance and well-being (3rd European Company Survey). https://doi.org/10.2806/85175

European Union. (2016). General data protection regulation. *Official Journal of the European Union, 119*(2).

Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental Fatigue Affects Visual Selective Attention. *PLoS ONE, 7*(10), 1–10. https://doi.org/10.1371/journal.pone.0048073

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology, 78*(6), 447–455. https://doi.org/https://doi.org/10.1016/0013-4694(91)90062-9

Ferré, S. (2008). An update on the mechanisms of the psychostimulant effects of caffeine. *Journal of Neurochemistry, 105*(4), 1067–1079. https://doi.org/10.1111/j.1471-4159.2007.05196.x

Ferré, S. (2016). Mechanisms of the psychostimulant effects of caffeine: implications for substance use disorders. *Psychopharmacology, 233*(10), 1963–1979. https://doi.org/10.1007/s00213-016-4212-2

Ferré, S., Bonaventura, J., Zhu, W., Hatcher-Solis, C., Taura, J., Quiroz, C., Cai, N. S., Moreno, E., Casadó-Anguera, V., Kravitz, A. V., Thompson, K. R., Tomasi, D. G., Navarro, G., Cordomí, A., Pardo, L., Lluís, C., Dessauer, C. W., Volkow, N. D., Casadó, V., … Zwilling, D. (2018). Essential control of the function of the striatopallidal neuron by pre-coupled complexes of adenosine A2A-dopamine D2 receptor heterotetramers and adenylyl cyclase. *Frontiers in Pharmacology, 9*(APR), 1–18. https://doi.org/10.3389/fphar.2018.00243

Ferré, S., Popoli, P., Giménez-Llort, L., Rimondini, R., Müller, C. E., Strömberg, I., Ögren, S. O., & Fuxe, K. (2001). Adenosine/dopamine interaction: Implications for the treatment of Parkinson's disease. *Parkinsonism and Related Disorders, 7*(3), 235–241. https://doi.org/10.1016/S1353-8020(00)00063-8

Fisher, E., O'Rourke, M., Evans, R., Kennedy, E. B., Gorman, M. E., & Seager, T. P. (2015). Mapping the integrative field: taking stock of socio-technical collaborations. *Journal of Responsible Innovation, 2*(1), 39–61. https://doi.org/10.1080/23299460.2014.1001671

Flipse, S. M., van der Sanden, M. C. A., Radstake, M., de Winde, J. H., & Osseweijer, P. (2014). The DNA of socially responsible innovation. *EMBO Reports, 15*(2), 134–137.

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology, 2*(JUL), 1–13. https://doi.org/10.3389/fpsyg.2011.00154

Fredholm, B. B., Bättig, K., Holmén, J., Nehlig, A., & Zvartau, E. E. (1999). Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews, 51*(1), 83–133.

Fredholm, B. B., Irenius, E., Kull, B., & Schulte, G. (2001). Comparison of the potency of adenosine as an agonist at human adenosine receptors expressed in Chinese hamster ovary cells. *Biochemical Pharmacology, 61*(4), 443–448. https://doi.org/10.1016/S0006-2952(00)00570-0

Fritz, C., & Sonnentag, S. (2005). Recovery, Health, and Job Performance: Effects of Weekend Experiences. *Journal of Occupational Health Psychology, 10*(3), 187–199. https://doi.org/10.1037/1076-8998.10.3.187

Galván, V. V, Vessal, R. S., & Golley, M. T. (2013). The Effects of Cell Phone Conversations on the Attention and Memory of Bystanders. *PLoS ONE, 8*(3), 1–10. https://doi.org/10.1371/journal.pone.0058579

Gander, P., Purnell, H., Garden, A., & Woodward, A. (2007). Work patterns and fatigue-related risk among junior doctors. *Occupational and Environmental Medicine, 64*(11), 733–738. https://doi.org/10.1136/oem.2006.030916

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science, 4*(6), 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gevaerd, M. S., Takahashi, R. N., Silveira, R., & Da Cunha, C. (2001). Caffeine reverses the memory disruption induced by intra-nigral MPTP-injection in rats. *Brain Research Bulletin, 55*(1), 101–106. https://doi.org/10.1016/S0361-9230(01)00501-9

Goossens, R. H. M., Netten, M. P., & Van Der Doelen, B. (2012). An office chair to influence the sitting behavior of office workers. *Work, 41*(SUPPL.1), 2086–2088. https://doi.org/10.3233/WOR-2012-0435-2086

Grandjean, E. (1979). Fatigue in industry. *British Journal of Industrial Medicine, 36*(3), 175–186. https://doi.org/10.1136/oem.36.3.175

Grent-'T-Jong, T., Boehler, C. N., Kenemans, J. L., & Woldorff, M. G. (2011). Differential functional roles of slow-wave and oscillatory-alpha activity in visual sensory cortex during anticipatory visual-spatial attention. *Cerebral Cortex, 21*(10), 2204–2216. https://doi.org/10.1093/cercor/bhq279

Grent-'t-Jong, T., & Woldorff, M. G. (2007). Timing and sequence of brain activity in top-down control of visual-spatial attention. *PLoS Biology, 5*(1), 0114–0126. https://doi.org/10.1371/journal.pbio.0050012

Grunwald, A. (2014). The hermeneutic side of responsible research and innovation. *Journal of Responsible Innovation, 1*(3), 274–291. https://doi.org/10.1080/23299460.2014.968437

Hansen, J. C., & Hillyard, S. A. (1980). Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology, 49*(3–4), 277–290. https://doi.org/10.1016/0013-4694(80)90222-9

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage, 30*(4), 1383–1400. https://doi.org/10.1016/j.neuroimage.2005.11.048

Helton, W. S., & Russell, P. N. (2017). Rest Is Still Best: The Role of the Qualitative and Quantitative Load of Interruptions on Vigilance. *Human Factors, 59*(1), 91–100. https://doi.org/10.1177/0018720816683509

Hendriksen, I. J. M., Snoijer, M., De Kok, B. P. H., Van Vilsteren, J., & Hofstetter, H. (2016). Effectiveness of a Multilevel Workplace Health Promotion Program on Vitality, Health, and Work-Related Outcomes. *Journal of Occupational and Environmental Medicine, 58*(6), 575–583. https://doi.org/10.1097/JOM.0000000000000747

Hertzog, C., Vernon, M. C., & Rypma, B. (1993). Age differences in mental rotation task performance: the influence of speed/accuracy tradeoffs. *Journal of Gerontology, 48*(3), 150-156. https://doi.org/10.1093/geronj/48.3.P150

Hess, T. M., & Ennis, G. E. (2011). Age differences in the effort and costs associated with cognitive activity. *The Journals of Gereontology, Series B: Psychological Sciences and Social Sciences, 67*, 447–455. https://doi.org/10.1093/geronb/gbr129.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 30*(33), 11096—11103. https://doi.org/10.1523/jneurosci.1026-10.2010

Hickey, C., & van Zoest, W. (2012). Reward creates oculomotor salience. *Current Biology, 22*(7), R219–R220. https://doi.org/https://doi.org/10.1016/j.cub.2012.02.007

Hillyard, S. A. (1969). Relationships between the contingent negative variation (CNV) and reaction time. *Physiology & Behavior, 4*(3), 351–357. https://doi.org/https://doi.org/10.1016/0031-9384(69)90188-7

Hockey, G. R. J. (1993). Cognitive-energetical control mechanisms in the management of work demands and psychological health. *In Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 328–345). Clarendon Press/Oxford University Press.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology, 45*(1–3), 73–93. https://doi.org/10.1016/S0301-0511(96)05223-4

Hockey, G. R. J. (2011). A motivational control theory of cognitive fatigue. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications.* (pp. 167–187). American Psychological Association. https://doi.org/10.1037/12343-008

Hockey, G. R. J. (2013). The psychology of fatigue: Work, effort and control. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139015394

Hockey, G. R. J., & Earle, F. (2006). Control over the scheduling of simulated office work reduces the impact of workload on mental fatigue and task performance. *Journal of Experimental Psychology: Applied, 12*(1), 50–65. https://doi.org/10.1037/1076-898X.12.1.50

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19*(1), 59–73. https://doi.org/10.1016/j.cogbrainres.2003.10.022

Hooff, M. L. M., Geurts, S. A. E., Kompier, M. A. J., & Taris, T. W. (2007). Workdays, in-between workdays and the weekend: A diary study on effort and recovery. *International Archives of Occupational and Environmental Health, 80*(7), 599–613. https://doi.org/10.1007/s00420-007-0172-5

Hooftman, W. ., Mars, G. M. ., Knops, J. C. M., Janssen, B. M. J., & Pleijers, A. J. S. F. (2019). Nationale Enquête Arbeidsomstandigheden 2018: Methodologie en globale resultaten. Leiden / Heerlen: TNO / CBS; 2019.

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology, 52*(3), 305–315. https://doi.org/10.1111/psyp.12339

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A. J., & Leung, Y. K. (2016). Shifts in attention during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology: Human Perception and Performance, 42*(6), 878–889. https://doi.org/10.1037/xhp0000189

Jacobs, J. V., Hettinger, L. J., Huang, Y. H., Jeffries, S., Lesch, M. F., Simmons, L. A., Verma, S. K., & Willetts, J. L. (2019). Employee acceptance of wearable technology in the workplace. *Applied Ergonomics, 78*(March), 148–156. https://doi.org/10.1016/j.apergo.2019.03.003

Jakobsen, S. E., Fløysand, A., & Overton, J. (2019). Expanding the field of Responsible Research and Innovation (RRI)–from responsible research to responsible innovation. *European Planning Studies, 27*(12), 2329–2343. https://doi.org/10.1080/09654313.2019.1667617

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience, 4*(November), 1–8. https://doi.org/10.3389/fnhum.2010.00186

Jing, S., & Doorn, N. (2020). Engineers' Moral Responsibility: A Confucian Perspective. *Science and Engineering Ethics, 26*(1), 233–253. https://doi.org/10.1007/s11948-019-00093-4

Johnson, D. G., & Powers, T. M. (2005). Computer systems and responsibility: A normative look at technological complexity. *Ethics and Information Technology, 7*(2), 99–107. https://doi.org/10.1007/s10676-005-4585-0

Juliano, L. M., & Griffiths, R. R. (2004). A critical review of caffeine withdrawal: Empirical validation of symptoms and signs, incidence, severity, and associated features. *Psychopharmacology, 176*(1), 1–29. https://doi.org/10.1007/s00213-004-2000-x

Kalantari, M. (2017). Consumers adoption of wearable technologies: literature review, synthesis, and future research agenda. *International Journal of Technology Marketing, 12*(1), 1. https://doi.org/10.1504/ijtmkt.2017.10008634

Kalfaoğlu, Ç., & Stafford, T. (2014). Performance breakdown effects dissociate from error detection effects in typing. *The Quarterly Journal of Experimental Psychology, 67*(3), 508–524. https://doi.org/10.1080/17470218.2013.820762

Kanfer, R., & Ackerman, P. L. (1989). Motivation and Cognitive Abilities: An Integrative/Aptitude-Treatment Interaction Approach to Skill Acquisition. *Journal of Applied Psychology, 74*(4), 657-690. https://doi.org/10.1037/0021-9010.74.4.657

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology, 47*(5), 888–904. https://doi.org/10.1111/j.1469-8986.2010.01009.x

Kecklund, G., & Axelsson, J. (2016). Health consequences of shift work and insufficient sleep. BMJ (Clinical research ed.), 355, i5210. https://doi.org/10.1136/bmj.i5210

Kenemans, J. L., & Lorist, M. M. (1995). Caffeine and Selective Visual Processing. *Pharmacology Biochemistry Behavior, 52*(3), 461–471. doi: 10.1016/0091-3057(95)00159-t.

Kenemans, J. L., & Verbaten, M. N. (1998). Caffeine and visuo-spatial attention. *Psychopharmacology, 135*(4), 353–360. https://doi.org/10.1007/s002130050522

Kenemans, J. L., Wieleman, J. S. T., Zeegers, M., & Verbaten, M. N. (1999). Caffeine and stroop interference. *Pharmacology Biochemistry and Behavior, 63*(4), 589–598. https://doi.org/10.1016/S0091-3057(99)00022-2

Kenny, G. P., Yardley, J. E., Martineau, L., & Jay, O. (2008). Physical work capacity in older adults: Implications for the aging worker. *American Journal of Industrial Medicine, 51*(8), 610–625. https://doi.org/10.1002/ajim.20600

Kimura, M., Sato, H., Ochi, M., Hosoya, S., & Sadoyama, T. (2007). Electromyogram and perceived fatigue changes in the trapezius muscle during typewriting and recovery. *European Journal of Applied Physiology, 100*(1), 89–96. https://doi.org/10.1007/s00421-007-0410-2

Kiran, A. H. (2012). Responsible design. A conceptual look at interdependent design-use dynamics. *Philosophy and Technology, 25*(2), 179–198. https://doi.org/10.1007/s13347-011-0052-5

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). *What ' s new in Psychtoolbox-3. Perception, 36*, 1–16.

Koelega, H. S. (1993). Stimulant drugs and vigilance performance: a review. *Psychopharmacology, 111*(1), 1–16. https://doi.org/10.1007/BF02257400

Kortuem, G., Alford, D., Ball, L., Busby, J., Davies, N., Efstratiou, C., Finney, J., White, M. I., & Kinder, K. (2007). Sensor Networks or Smart Artifacts? An Exploration of Organizational Issues of an Industrial Health and Safety Monitoring System. In: Krumm J., Abowd G.D., Seneviratne A., Strang T. *(eds) UbiComp 2007: Ubiquitous Computing. UbiComp 2007*. Lecture Notes in Computer Science, vol. 4717. Springer, Berlin, Heidelberg. ttps://doi.org/10.1007/978-3-540-74853-3_27

Krishnamurthy, M., Ramalingam, P., Perumal, K., Kamalakannan, L. P., Chinnadurai, J., Shanmugam, R., Srinivasan, K., & Venugopal, V. (2017). Occupational Heat Stress Impacts on Health and Productivity in a Steel Industry in Southern India. *Safety and Health at Work, 8*(1), 99–104. https://doi.org/10.1016/j.shaw.2016.08.005

Kühnel, J., Zacher, H., de Bloom, J., & Bledow, R. (2017). Take a break! Benefits of sleep and short breaks for daily work engagement. *European Journal of Work and Organizational Psychology, 26*(4), 481–491. https://doi.org/10.1080/1359432X.2016.1269750

Kukla, R. (2005). Conscientious autonomy: displacing decisions in health care. *Hastings Center Report, 35*, 34–44. doi: 10.1353/hcr.2005.0025

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161–163. https://doi.org/10.1038/307161a0

Kutas, M., & Petten, C. van (1988). Event-related brain potential studies of language. *Advances in Psychophysiology, 3*, 139–187. http://kutaslab.ucsd.edu/people/kutas/pdfs/1988.AP.139.pdf

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(2), 1–26. https://doi.org/10.18637/jss.v082.i13

La Fors, K., Custers, B., & Keymolen, E. (2019). Reassessing values for emerging big data technologies: integrating design-based and application-based approaches. *Ethics and Information Technology, 21*(3), 209–226. https://doi.org/10.1007/s10676-019-09503-4

Langner, R., Steinborn, M. B., Chatterjee, A., Sturm, W., & Willmes, K. (2010). Mental fatigue and temporal preparation in simple reaction-time performance. *Acta Psychologica, 133*(1), 64–72. https://doi.org/10.1016/j.actpsy.2009.10.001

Langner, R., Willmes, K., Chatterjee, A., Eickhoff, S. B., & Sturm, W. (2010). Energetic effects of stimulus intensity on prolonged simple reaction-time performance. *Psychological Research, 74*(5), 499–512. https://doi.org/10.1007/s00426-010-0275-6

Leclercq-Vandelannoitte, A. (2017). An Ethical Perspective on Emerging Forms of Ubiquitous IT-Based Control. *Journal of Business Ethics, 142*(1), 139–154. https://doi.org/10.1007/s10551-015-2708-z

Lieberman, H. R., Wurtman, R. J., Emde, G. G., Roberts, C., & Coviella, I. L. G. (1987). The effects of low doses of caffeine on human performance and mood. *Psychopharmacology, 92*(3), 308–312. https://doi.org/10.1007/BF00210835

Linder, J. A., Doctor, J. N., Friedberg, M. W., Reyes Nieva, H., Birks, C., Meeker, D., & Fox, C. R. (2014). Time of Day and the Decision to Prescribe Antibiotics. *JAMA Internal Medicine, 174*(12), 2029–2031. https://doi.org/10.1001/jamainternmed.2014.5225

Linssen, A. M. W., Vuurman, E. F. P. M., Sambeth, A., Nave, S., & Spooren, W. (2011). Contingent negative variation as a dopaminergic biomarker : evidence from dose-related effects of methylphenidate. *Psychopharmacology, 218*, 533–542. https://doi.org/10.1007/s00213-011-2345-x

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia, 38*(5), 701–711. https://doi.org/10.1016/S0028-3932(99)00106-2

Lodge, T., & Crabtree, A. (2019). Privacy engineering for domestic IoT: Enabling due diligence. *Sensors (Switzerland), 19*(20). https://doi.org/10.3390/s19204380

Logan, F. A. (1999). Errors in copy typewriting. *Journal of Experimental Psychology-Human Perception and Performance, 25*(6), 1760–1773. https://doi.org/10.1037/0096-1523.25.6.1760

Lorist, M. M., Boksem, M. A. S., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. Brain Research. *Cognitive Brain Research, 24*(2), 199–205. https://doi.org/10.1016/j.cogbrainres.2005.01.018

Lorist, M. M., & Faber, L. G. (2011). Considerations of the influence of mental fatigue on controlled and automatic cognitive processing and related neuromodulatory effects. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 105–126). American Psychological Association.

Lorist, M. M., & Jolij, J. (2012). Trial history effects in Stroop task performance are independent of top-down control. *PLoS ONE, 7*(6), 1–13. https://doi.org/10.1371/journal.pone.0039802

Lorist, M. M., Kernell, D., Meijman, T. F., & Zijdewind, I. (2002). Motor fatigue and cognitive task performance in humans. *The Journal of Physiology, 545*(1), 313–319. https://doi.org/10.1113/jphysiol.2002.027938

Lorist, M. M., Klein, M., Nieuwenhuis, S., De Jong, R., Mulder, G., & Meijman, T. F. (2000). Mental fatigue and task control: planning and preparation. *Psychophysiology, 37*, 614–625. https://doi.org/10.1111/1469-8986.3750614

Lorist, M. M., Snel, J., & Kok, A. (1994). Influence of caffeine on information processing stages in well rested and fatigued subjects. *Psychopharmacology, 113*(3), 411–421. https://doi.org/10.1007/BF02245217

Lorist, M. M., Snel, J., Mulder, G., & Kok, A. (1995). Aging, caffeine, and information processing: an event-related potential analysis. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials, 96*(5), 453–467. https://doi.org/10.1016/0168-5597(95)00069-5

Lorist, M. M., & Tops, M. (2003). Caffeine, fatigue, and cognition. *Brain and Cognition, 53*(1), 82–94. https://doi.org/10.1016/S0278-2626(03)00206-9

Lucci, G., Berchicci, M., Spinelli, D., Taddei, F., & Di Russo, F. (2013). The Effects of Aging on Conflict Detection. *PLoS ONE, 8*(2). https://doi.org/10.1371/journal.pone.0056566

Lucivero, F., Marelli, L., Hangel, N., Zimmermann, B. M., Prainsack, B., Galasso, I., Horn, R., Kieslich, K., Lievevrouw, E., Ongolly, F., Samuel, G., Siffels, L., Stendahl, E., Hoyweghen, I. Van, Lucivero, F., Marelli, L., Hangel, N., Zimmermann, B. M., Prainsack, B., Stendahl, E. & van Hoyweghen, I. (2021). Normative positions towards COVID-19 contact-tracing apps: Findings from a large-scale qualitative study in nine European countries. *Critical Public Health*, 1–14. https://doi.org/10.1080/09581596.2021.1925634

Luck, S. J. (2005). An introduction to the event-related potential technique. Cambridge Mass: MIT Press.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

MacLeod, C. M. (1991). Half a century of reseach on the stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Manson, N. C., & O'Neill, O. (2007). Rethinking Informed Consent in Bioethics. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511814600

Martin, B. J. (1996). Keyboard Reaction Force and Finger Flexor Electromyograms during Computer Keyboard Work. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 38*(4), 654–664. https://doi.org/10.1518/001872096778827288

Martin, K. E., & Freeman, R. E. (2004). The separation of technology and ethics in business ethics. *Journal of Business Ethics, 53*(4), 353–364. https://doi.org/10.1023/B:BUSI.0000043492.42150.b6

Mattila, E., Orsama, A. L., Ahtinen, A., Hopsu, L., Leino, T., & Korhonen, I. (2013). Personal health technologies in employee health promotion: Usage activity, usefulness, and health-related outcomes in a 1-year randomized controlled trial. *Journal of Medical Internet Research, 15*(7). https://doi.org/10.2196/mhealth.2557

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McCormick, F., Kadzielski, J., Landrigan, C. P., Evans, B., Herndon, J. H., & Rubash, H. E. (2012). Surgeon Fatigue. *Archives of Surgery, 147*(5), 430–435. https://doi.org/10.1001/archsurg.2012.84

McLellan, T. M., Caldwell, J. A., & Lieberman, H. R. (2016). A review of caffeine's effects on cognitive, physical and occupational performance. *Neuroscience and Biobehavioral Reviews, 71*, 294–312. https://doi.org/10.1016/j.neubiorev.2016.09.001

Mills, L., Boakes, R. A., & Colagiuri, B. (2016). Placebo caffeine reduces withdrawal in abstinent coffee drinkers. *Journal of Psychopharmacology, 30*(4), 388–394. https://doi.org/10.1177/0269881116632374

Mills, L., Dar-Nimrod, I., & Colagiuri, B. (2017). Effect of Genetic Information and Information About Caffeine Content on Caffeine Withdrawal Symptoms. *Scientific Reports, 7*(1), 8407. https://doi.org/10.1038/s41598-017-08678-4

Miozzo, M., Pulvermüller, F., & Hauk, O. (2015). Early parallel activation of semantics and phonology in picture naming: Evidence from a multiple linear regression MEG study. *Cerebral Cortex, 25*(10), 3343–3355. https://doi.org/10.1093/cercor/bhu137

Morgado, M., Talaia, M., & Teixeira, L. (2017). A new simplified model for evaluating thermal environment and thermal sensation: An approach to avoid occupational disorders. *International Journal of Industrial Ergonomics, 60*, 3–13. https://doi.org/10.1016/j.ergon.2015.11.001

Morley, J., Cowls, J., Taddeo, M., & Floridi, L. (2020). Ethical guidelines for COVID-19. *Nature, 582*, 29–31. 10.1038/d41586-020-01578-0.

Motti, V. G., & Caine, K. (2014). Human factors considerations in the design of wearable devices. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*(1), 1820–1824. https://doi.org/10.1177/1541931214581381

Mulligan, D. K., Koopman, C. & Doty, N. (2016). Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy. *Philosophical Transactions of the Royal Society a Mathematical, Physical and Engineering Sciences*. 374:20160118. doi: 10.1098/rsta.2016.0118

Muscio, B. (1921). Feeling-Tone in Industry. *British Journal of Psychology, 12,* 150–162.

Nehlig, A. (2010). Is caffeine a cognitive enhancer? *Journal of Alzheimer's Disease, 20*(S1), S85-S94. https://doi.org/10.3233/JAD-2010-091315

Nehlig, A., Daval, J. L., & Debry, G. (1992). Caffeine and the central nervous system: mechanisms of action. *Brain Research Reviews, 17*(2), 139–170.

Nissenbaum, H. (2010). *Privacy in Context : Technology, Policy, and the Integrity of Social Life.* Redwood City: Stanford University Press. https://doi.org/10.1515/9780804772891

Norbert, S. (1999). Self-Reports: How the Questions Shape the Answers. *American Psychologist, 54*(2), 93–105.

Okada, M., Mizuno, K., & Kaneko, S. (1996). Adenosine A1 and A2 receptors modulate extracellular dopamine levels in rat striatum. *Neuroscience Letters, 212*(1), 53–56. https://doi.org/10.1016/0304-3940(96)12780-4

Olsen, C. B. (2020). To track or not to track? Employees' data privacy in the age of corporate wellness, mobile health, and GDPR. *International Data Privacy Law, 10*(3), 236–252. https://doi.org/10.1093/idpl/ipaa004

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011. https://doi.org/10.1155/2011/156869

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy, 39*(6), 751–760. https://doi.org/10.1093/scipol/scs093

Padmala, S., & Pessoa, L. (2011). Reward Reduces Conflict by Enhancing Attentional Control and Biasing Visual Cortical Processing. *Journal of Cognitive Neuroscience, 23*(11), 3419–3432. https://doi.org/10.1162/jocn_a_00011

Palm, E. (2009). Securing privacy at work: The importance of contextualized consent. *Ethics and Information Technology, 11*(4), 233–241. https://doi.org/10.1007/s10676-009-9208-8

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10* (4), 437–442. https://doi.org/10.1163/156856897X00366

Persson, R., Helene Garde, A., Schibye, B., & Ørbæk, P. (2006). Building-site camps and extended work hours: A two-week monitoring of self-reported physical exertion, fatigue, and daytime sleepiness. *Chronobiology International, 23*(6), 1329–1345. https://doi.org/10.1080/07420520601058021

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience, 35,* 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Pimenta, A., Carneiro, D., Novais, P., & Neves, J. (2013). Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns. *Lecture Notes in Computer Science*, 8073 LNAI, 222–231. https://doi.org/10.1007/978-3-642-40846-5_23

Pimenta, A., Carneiro, D., Novais, P., & Neves, J. (2014). Analysis of Human Performance as a Measure of Mental Fatigue. *Hybrid Artificial Intelligence Systems*, 389–401. https://doi.org/10.1007/978-3-319-07617-1_35

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J. (2012). Neuropsychology of P300. In The Oxford Handbook of Event-Related Potential Components (Issue June 2016). https://doi.org/10.1093/oxfordhb/9780195374148.013.0089

Quarta, D., Ferré, S., Solinas, M., You, Z. B., Hockemeyer, J., Popoli, P., & Goldberg, S. R. (2004). Opposite modulatory roles for adenosine A1 and A2A receptors on glutamate and dopamine release in the shell of the nucleus accumbens. Effects of chronic caffeine exposure. *Journal of Neurochemistry, 88*(5), 1151–1158. https://doi.org/10.1046/j.1471-4159.2003.02245.x

R.Studio Team. (2016). RStudio: Integrated Development for R. https://doi.org/10.1145/3132847.3132886

R Development Core Team. (2017). R: A language and environment for statistical computing. In Vienna, Austria. Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rabbitt, P. M. (1979). How old and young subjects monitor and control responses for accuracy and speed. *British Journal of Psychology*, 305–311. https://doi.org/10.1111/j.2044-8295.1979.tb01687.x

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin, 26*(4), 419–435. https://doi.org/10.1177/0146167200266002

Ricci, J. A., Chee, E., Lorandeau, A. L., & Berger, J. (2007). Fatigue in the U.S. workforce: Prevalence and implications for lost productive work time. *Journal of Occupational and Environmental Medicine, 49*(1), 1–10. https://doi.org/10.1097/01.jom.0000249782.60321.2a

Robaey, Z., Spruit, S. L., & van de Poel, I. (2018). The Food Warden: An Exploration of Issues in Distributing Responsibilities for Safe-by-Design Synthetic Biology Applications. *Science and Engineering Ethics, 24*(6), 1673–1696. https://doi.org/10.1007/s11948-017-9969-0

Robinson, E. S., & Bills, A. G. (1926). Two factors in the work decrement. *Journal of Experimental Psychology, 9*(6), 415–443. https://doi.org/10.1037/h0072598

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences, 14*(2), 64–71. https://doi.org/10.1016/j.tics.2009.11.005

Roghanchi, P., & Kocsis, K. C. (2018). Challenges in Selecting an Appropriate Heat Stress Index to Protect Workers in Hot and Humid Underground Mines. *Safety and Health at Work, 9* (1), 10–16. https://doi.org/10.1016/j.shaw.2017.04.002

Rook, J., & Zijlstra, F. (2006). The contribution of various types of activities to recovery. *European Journal of Work and Organizational Psychology, 15*, 218–240. https://doi.org/10.1080/13594320500513962

Roossien, C.C., Baten, C.T.M., van der Waard, M.W.P., Reneman, M.F. & Verkerke, G.J. Automatically Determining Lumbar Load during Physically Demanding Work: A Validation Study. *Sensors 2021, 21*, 2476. https://doi.org/10.3390/s21072476

Roossien, C. C., Heus, R., Reneman, M. F., & Verkerke, G. J. (2020). Monitoring core temperature of firefighters to validate a wearable non-invasive core thermometer in different types of protective clothing: Concurrent in-vivo validation. *Applied Ergonomics, 83,* 103001. https://doi.org/10.1016/j.apergo.2019.103001

Roossien, C.C., Hodselmans, A.P., Heus, R., Reneman, M.F. & Verkerke, G.J. (2021) Evaluation of a Wearable Non-Invasive Thermometer for Monitoring Ear Canal Temperature during Physically Demanding (Outdoor) Work. *International Journal of Environmental Research and Public Health, 18*(9), 4896. https://doi.org/10.3390/ijerph18094896

Roossien, C. C., Krops, L. A., Wempe, J. B., Verkerke, G. J., & Reneman, M. F. (2021). Can breathing gases be analyzed without a mouth mask? Proof-of-concept and concurrent validity of a newly developed design with a mask-less headset. *Applied Ergonomics, 90*. https://doi.org/10.1016/j.apergo.2020.103266

Roossien, C. C., Stegenga, J., Hodselmans, A. P., Spook, S. M., Koolhaas, W., Brouwer, S., Verkerke, G. J., & Reneman, M. F. (2017). Can a smart chair improve the sitting behavior of office workers? *Applied Ergonomics, 65,* 355–361. https://doi.org/10.1016/j.apergo.2017.07.012

Ruijter, J., De Ruiter, M. B., Snel, J., & Lorist, M. M. (2000). The influence of caffeine on spatial-selective attention: An event-related potential study. *Clinical Neurophysiology, 111*(12), 2223–2233. https://doi.org/10.1016/S1388-2457(00)00462-4

Ruijter, J., Lorist, M. M., & Snel, J. (1999). The influence of different doses of caffeine on visual task performance. *Journal of Psychophysiology, 13*(1), 37–48. https://doi.org/10.1027//0269-8803.13.1.37

Ruijter, J., Lorist, M. M., Snel, J., & De Ruiter, M. B. (2000). The influence of caffeine on sustained attention: An ERP study. *Pharmacology Biochemistry and Behavior, 66*(1), 29–37. https://doi.org/10.1016/S0091-3057(00)00229-X

Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality, 74*(6), 1557–1586. https://doi.org/10.1111/j.1467-6494.2006.00420.x

Saliasi, E., Geerligs, L., Lorist, M. M., & Maurits, N. M. (2013). The Relationship between P3 Amplitude and Working Memory Performance Differs in Young and Older Adults. *PLoS ONE, 8*(5). https://doi.org/10.1371/journal.pone.0063701

Salthouse, T. A. (1984). Effects of age and skill in typing. Journal of Experimental Psychology. *General, 113*(3), 345–371. https://doi.org/10.1037/0096-3445.113.3.345

Saurabh, K., Rao, H., Amrutur, B., & Sundarrajan, A. (2014). Continuous core body temperature estimation via SURFACE temperature measurements using wearable sensors: Is it feasible? In *Proceedings of International Conference on Biomedical electronics and devices* - BIOSTEC 2014, 181–186. https://doi.org/10.5220/0004889901810186

Saville, C. W. N., de Morree, H. M., Dundon, N. M., Marcora, S. M., & Klein, C. (2018). Effects of caffeine on reaction time are mediated by attentional rather than motor processes. *Psychopharmacology, 235*(3), 749–759. https://doi.org/10.1007/s00213-017-4790-7

Scheeringa, R., Koopmans, P. J., van Mourik, T., Jensen, O., & Norris, D. G. (2016). The relationship between oscillatory EEG activity and the laminar-specific BOLD signal. *Proceedings of the National Academy of Sciences, 113*(24), 6761–6766. https://doi.org/10.1073/pnas.1522577113

Scheeringa, R., Petersson, K. M., Kleinschmidt, A., Jensen, O., & Bastiaansen, M. C. m. (2012). EEG Alpha Power Modulation of fMRI Resting-State Connectivity. *Brain Connectivity, 2*(5), 254–264. https://doi.org/10.1089/brain.2012.0088

Schevernels, H., Krebs, R. M., Santens, P., Woldorff, M. G., & Boehler, C. N. (2014). Task preparation processes related to reward prediction precede those related to task-difficulty expectation. *NeuroImage, 84*, 639–647. https://doi.org/10.1016/j.neuroimage.2013.09.039

Schukat, M., McCaldin, D., Wang, K., Schreier, G., Lovell, N. H., Marschollek, M., & Redmond, S. J. (2016). Unintended Consequences of Wearable Sensor Use in Healthcare. Contribution of the IMIA Wearable Sensors in Healthcare WG. *Yearbook of Medical Informatics, 2*(1), 73–86. https://doi.org/10.15265/iy-2016-025

Schultz W. (2000). Multiple reward signals in the brain. Nature reviews. *Neuroscience, 1*(3), 199–207. https://doi.org/10.1038/35044563

Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews, 95*(3), 853–951. https://doi.org/10.1152/physrev.00023.2014

Shilton, K. (2013). Values Levers: Building Ethics into Design. *Science Technology and Human Values, 38*(3), 374–397. https://doi.org/10.1177/0162243912436985

Sluiter, J. K., de Croon, E. M., Meijman, T. F., & Frings-Dresen, M. H. W. (2003). Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. *Occupational and Environmental Medicine, 60*, i62-i70. https://doi.org/10.1136/oem.60.suppl_1.i62

Smith-Jackson, T. L., & Klein, K. W. (2009). Open-plan offices: Task performance and mental workload. *Journal of Environmental Psychology, 29*(2), 279–289. https://doi.org/10.1016/j.jenvp.2008.09.002

Smolders, K. C. H. J., De Kort, Y. A. W., Tenner, A. D., & Kaiser, F. G. (2012). Need for recovery in offices: Behavior-based assessment. *Journal of Environmental Psychology, 32*(2), 126–134. https://doi.org/10.1016/j.jenvp.2011.12.003

Sole, M., Musu, C., Boi, F., Giusto, D., & Popescu, V. (2013). RFID sensor network for workplace safety management. *2013 IEEE 18th Conference on Emerging Technologies Factory Automation (ETFA)*, 1–4. https://doi.org/10.1109/ETFA.2013.6648157

Solove, D. J. (2008). Privacy: A Concept in Disarray Privacy. *University of Pennsylvania Law Review, 154*(3), 477-564. https://doi.org/10.2307/40041279

Sparks, K., Cooper, C., Fried, Y., & Shirom, A. (1997). The effects of hours of work on health: A meta-analytic review. *Journal of Occupational and Organizational Psychology, 70*(1991), 391–408. https://doi.org/10.1111/j.2044-8325.1997.tb00656.x

Spook, S. M., Koolhaas, W., Bültmann, U., & Brouwer, S. (2019). Implementing sensor technology applications for workplace health promotion: a needs assessment among workers with physically demanding work. *BMC Public Health, 19*(1), 1100. https://doi.org/10.1186/s12889-019-7364-2

Stahl, B. C. (2013). Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy, 40*(6), 708–716. https://doi.org/10.1093/scipol/sct067

Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed–accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging, 25*(2), 377–390. https://doi.org/10.1037/a0018022

Steege, L. M., Drake, D. A., Olivas, M., & Mazza, G. (2015). Evaluation of physically and mentally fatiguing tasks and sources of fatigue as reported by registered nurses. *Journal of Nursing Management, 23*(2), 179–189. https://doi.org/10.1111/jonm.12112

Steinborn, M. B., & Huestegge, L. (2016). A Walk Down the Lane Gives Wings to Your Brain. Restorative Benefits of Rest Breaks on Cognition and Self-Control. *Applied Cognitive Psychology, 30*(5), 795–805. https://doi.org/10.1002/acp.3255

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy, 42*(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Swaen, G. M. H., Van Amelsvoort, L. G. P. M., Bültmann, U., & Kant, I. J. (2003). Fatigue as a risk factor for being injured in an occupational accident: Results from the Maastricht Cohort Study. *Occupational and Environmental Medicine, 60,* 88–92. https://doi.org/10.1136/oem.60.suppl_1.i88

Taylor, A. S., & Vincent, J. (2005). An SMS History In: Hamill L., Lasen A., Diaper D. (eds) *Mobile World. Computer Supported Cooperative Work*. Springer, London. https://doi.org/10.1007/1-84628-204-7_5

Ten Have, M., Van Der Heide, A., MacKenbach, J. P., & De Beaufort, I. D. (2013). An ethical framework for the prevention of overweight and obesity: A tool for thinking through a programme's ethical aspects. *European Journal of Public Health, 23*(2), 299–305. https://doi.org/10.1093/eurpub/cks052

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford University Press.

Thorndike, E. (1900). Mental fatigue. I. *Psychological Review, 7*(6), 547–579. https://doi.org/10.1037/h0069511

Tieges, Z., Snel, J., Kok, A., Plat, N., & Ridderinkhof, R. (2007). Effects of caffeine on anticipatory control processes: Evidence from a cued task-switch paradigm. *Psychophysiology, 44*(4), 561–578. https://doi.org/10.1111/j.1469-8986.2007.00534.x

Truxillo, D. M., Cadiz, D. M., & Hammer, L. B. (2015). Supporting the Aging Workforce: A Review and Recommendations for Workplace Intervention Research. *Annual Review of Organizational Psychology and Organizational Behavior, 2*, 351-381. https://doi.org/10.1146/annurev-orgpsych-032414-111435

Van Berkel, J., Meershoek, A., Janssens, R. M., Boot, C. R., Proper, K. I., & Van Der Beek, A. J. (2014). Ethical considerations of worksite health promotion: An exploration of stakeholders' views. *BMC Public Health, 14*(1), 1–10. https://doi.org/10.1186/1471-2458-14-458

Van De Poel, I., & Verbeek, P. P. (2006). Ethics and engineering design. *Science Technology and Human Values, 31*(3), 223–236. https://doi.org/10.1177/0162243905285838

van den Berg, B., Appelbaum, L. G., Clark, K., Lorist, M. M., & Woldorff, M. G. (2016). Visual search performance is predicted by both prestimulus and poststimulus electrical brain activity. *Scientific Reports, 6*(1), 37718. https://doi.org/10.1038/srep37718

van den Berg, B., de Jong, M., Woldorff, M. G., & Lorist, M. M. (2020). Caffeine Boosts Preparatory Attention for Reward-related Stimulus Information. *Journal of Cognitive Neuroscience, 33*(1),104–118. https://doi.org/10.1162/jocn_a_01630

Van Den Berg, B., Geib, B. R., Martín, R. S., & Woldorff, M. G. (2019). A key role for stimulus-specific updating of the sensory cortices in the learning of stimulus–reward associations. *Social Cognitive and Affective Neuroscience, 14*(2), 173–187. https://doi.org/10.1093/scan/nsy116

van den Berg, B., Krebs, R. M., Lorist, M. M., & Woldorff, M. G. (2014). Utilization of reward-prospect enhances preparatory attention and reduces stimulus conflict. *Cognitive, Affective & Behavioral Neuroscience, 14*(2), 561–577. https://doi.org/10.3758/s13415-014-0281-z

van der Klink, J. J. L., Bültmann, U., Burdorf, A., Schaufeli, W. B., Zijlstra, F. R. H., Abma, F. I., Brouwer, S., & Van der wilt, G. J. (2016). Sustainable employability – definition, conceptualization, and implications: A perspective based on the capability approach. *Scandinavian Journal of Work, Environment and Health, 42*(1), 71–79. https://doi.org/10.5271/sjweh.3531 van der Linden, D. (2010). The urge to stop: The cognitive and biological nature of acute mental fatigue. In P.L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 149–164). American Psychological Association. https://doi.org/10.1037/12343-007

van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: effects on perseveration and planning. *Acta Psychologica, 113*(1), 45–65. https://doi.org/10.1016/S0001-6918(02)00150-6

van Dijk, H., Schoffelen, J.-M., Oostenveld, R., & Jensen, O. (2008). Prestimulus Oscillatory Activity in the Alpha Band Predicts Visual Discrimination Ability. *The Journal of Neuroscience, 28*(8), 1816 LP – 1823. https://doi.org/10.1523/JNEUROSCI.1853-07.2008

Verbeek, P. (2011). *Moralizing Technology : Understanding and Designing the Morality of Things*. University of Chicago Press.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122*(3), 231–249. https://doi.org/10.1037/0033-2909.122.3.231

Villumsen, M., Madeleine, P., Jørgensen, M. B., Holtermann, A., & Samani, A. (2017). The variability of the trunk forward bending in standing activities during work vs. leisure time. *Applied Ergonomics, 58*, 273–280. https://doi.org/10.1016/j.apergo.2016.06.017

Walter, W., Cooper, R., Aldridge, V. J., Mccallum, W. M., & Winter, A. (1964). Contingent Negative Variation : An Electric Sign of Sensori-Motor Association and Expectancy in the Human Brain. *Nature, 203*, 380–384.

Wascher, E., & Getzmann, S. (2014). Rapid Mental Fatigue Amplifies Age-Related Attentional Deficits. *Psychophysiology, 28*(3), 215–224. https://doi.org/10.1027/0269-8803/a000127

Wei, K., & Körding, K. (2010). Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Frontiers in Computational Neuroscience, 4*(May), 1–9. https://doi.org/10.3389/fncom.2010.00011

Weiss, B., & Laties, V. G. (1962). *Enhancement of human performance by caffeine and the amphetamines. Pharmacological Reviews*, 14(April), 1–36.

West, R., & Alain, C. (2000). Effects of task context and fluctuations of attention on neural activity supporting performance of the Stroop task. *Brain Research, 873*(1), 102–111. https://doi.org/10.1016/S0006-8993(00)02530-0

West, R., & Bell, M. A. (1997). Stroop color-word interference and electroencephalogram activation: Evidence for age-related decline of the anterior attention system. *Neuropsychology, 11*(3), 421–427. https://doi.org/10.1037/0894-4105.11.3.421

Wickens, C. D., & Hollands, J. G. (2000). Engineering Psychology and Human Performance. *SA Journal of Industrial Psychology, 13* (1). https://doi.org/10.4102/sajip.v13i1.457

Wilhelmus, M. M. M., Hay, J. L., Zuiker, R. G. J. A., Okkerse, P., Perdrieu, C., Sauser, J., Beaumont, M., Schmitt, J., Van Gerven, J. M. A., & Silber, B. Y. (2017). Effects of a single, oral 60 mg caffeine dose on attention in healthy adult subjects. *Journal of Psychopharmacology, 31*(2), 222–232. https://doi.org/10.1177/0269881116668593

Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 20*(6), 1–6. https://doi.org/10.1523/jneurosci.20-06-j0002.2000

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA, 310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

Zhang, Y., Gong, J., Miao, D., Zhu, X., & Yang, Y. (2011). Subjective evaluation of mental fatigue and characteristics of attention during a driving simulation. *Social Behavior and Personality: An International Journal, 39*(1), 15–20. https://doi.org/10.2224/sbp.2011.39.1.15

Zhu, H., Gao, L., & Li, H. (2016). Secure and privacy-preserving body sensor data collection and query scheme. *Sensors (Switzerland), 16*(2). https://doi.org/10.3390/s16020179

Acknowledgements

Eindelijk mag ik het leukste stuk van dit boekwerk schrijven. Vorige week heb ik te horen gekregen dat de leescommissie mijn proefschrift heeft goedgekeurd en dat ik mijn proefschrift op donderdag 25 november om 11 uur mag verdedigen. De weg naar dit moment was lang en vol met mooie, maar ook moeilijke momenten en perioden. Het promotietraject heeft me gemaakt tot de onderzoeker, maar ook tot de mens die ik nu ben en ik had het voor geen goud willen missen. Dit heb ik niet alleen gedaan. Zonder de hulp en support van mijn omgeving – collega's, vrienden en familie – weet ik zeker dat ik dit resultaat niet neer had kunnen neerzetten.

Monicque, bedankt voor de afgelopen jaren. Met jouw hulp heb ik een breed proefschrift met veel verschillende hoofdstukken kunnen schrijven. Zo heb ik me kunnen ontwikkelen op een manier die bij mij past. Ik heb van jou geleerd om (iets meer) orde in mijn chaos te scheppen. Ook al was ik er soms een beetje klaar mee, ik wist dat je feedback de stukken scherper maakten en dit heeft mij als onderzoeker ook kritischer gemaakt in mijn werk. Dank daarvoor!

Jacob, sinds mijn bachelorthese hebben we samengewerkt. Door jou ben ik enthousiast geworden over het doen van onderzoek. Terwijl ik op dat moment zelf nog niet doorhad wat een promotietraject eigenlijk inhield. Bedankt!

Berry, collega, kamergenootje, lieve vriend en paranimf, in de periode dat ik extra ondersteuning nodig had bij mijn promotie, was jij daar. Jij hebt voor mij op een andere manier leren nadenken over data. Ik pluk hier nog elke dag de vruchten van. Daarnaast ben je ook nog eens een fantastische vriend waar ik altijd terecht kan. Ik ben zo blij dat jij mijn paranimf bent!

Lieve Rob, vandaag precies 7 jaar geleden – ik schrijf dit op 30 augustus 2021 - zijn we tegelijk aan ons promotietraject begonnen en beiden staan we nu op het punt om onze promotie af te ronden. Het voelt goed dat jij mijn paranimf bent. We waren niet direct heel close, maar ik weet nog heel goed dat we buiten voor de USVA stonden en beseften dat we eigenlijk heel erg op elkaar lijken. Vanaf dat moment begon er een tijd waarin we samen een reis door ons visuele bewustzijn hebben afgelegd. Deze meer filosofische ontwikkeling die parallel liep aan mijn promotietraject heeft mij veel gebracht en ik ben blij dat ik dat samen met jou heb mogen doen.

Jeffry, lieve vriend. Ik kan een waslijst met dingen opnoemen waar ik je allemaal voor wil bedanken, maar daar word je vast ongemakkelijk van, dus het gros zal ik even achterwege laten. Ik ben zo blij dat we een paar jaar geleden tijdens een lunch bij Soof's & Co hebben besloten om te gaan samenwonen. Jij bent dé stabiele factor in mijn leven en door jou heb ik de rust gehad om naast mijn baan bij Gronings Perspectief, mijn promotie af te kunnen ronden. Een groter cadeau had je me niet kunnen geven.

Lieve Lisette, ik wil je bedanken voor je warme vriendschap, je steun, je support in alles wat ik doe. Je bent het gekke, warme zonnetje in mijn leven dat ik nooit meer wil missen. Bedankt voor alle fijne gesprekken en momenten, maar ook alvast voor die momenten waarvan ik zeker weet dat ze nog gaan komen[4].

Lieve Pap, lieve Lara, bedankt dat jullie me hebben gesteund tijdens dit langdurige traject. Ik vond het heerlijk om de drukte van werk en proefschrift te kunnen ontvluchten voor een heerlijke fietstocht door het mooie Drenthe en/of een dagje knuffelen met mijn lieve nichtje.

---

[4] *Ik wil je ook nog graag bedanken voor je hulp tijdens de voorbereiding van mijn promotie en bij de lay-out. Zonder jou was deze periode echt een stuk stressvoller geweest. Bedankt voor alles!*

Lieve Mam, lieve Gert, bedankt dat jullie er de afgelopen jaren altijd voor me zijn geweest. Jullie hebben me de trainingen Essence en Source gegeven. Dit heeft me geholpen om meer de rust in mijzelf te vinden. Hierdoor heb ik mezelf de ruimte kunnen geven om mijn proefschrift rustig in mijn eigen tempo af te maken. Ook ben ik jullie erg dankbaar voor de weekenden dat ik bij jullie heb mogen schrijven aan mijn proefschrift. En Mam, natuurlijk bedankt voor het checken van de Nederlandse samenvatting!

Sanne, mijn lieve kamergenootje, huisgenootje, vriendin. Wat hebben we een fijne tijd gehad op kantoor en in ons appartementje in Helpman. Ik wil je bedanken voor je harde reflectie op mijn levensstijl in de tijd dat ik dat nodig had. Je gaf me het duwtje in de rug dat ik nodig had om mijn problemen aan te pakken.

Atser, bedankt voor de fijne gesprekken, je lieve persoonlijkheid, het zijn van mijn fotobuddy (oa. Vlieland op de duin en de bowlingfoto) en het checken van mijn Nederlandse samenvatting!

Josh, fijne vriend, wat een mooie, leuke, gekke uitjes hebben we gehad. Van verkleedpartijtjes tot een mooie wandelingen in het bos. Bedankt dat ik soms even bij je mocht zeuren.

Wouter, je intelligentie inspireert me, ook al vind ik het soms moeilijk om je bij te benen. Bedankt voor alle gezellige avonden, klimsessies en gesprekken (met ook af en toe wat geklaag tussendoor).

Minou, bedankt voor alle hulp en steun bij mijn proefschrift en het traject. Heerlijk om met jou allerlei lunchrestaurantjes te ontdekken door Groningen. Dat was en is nog steeds een welkome afwisseling van het werk.

Martin, Nadine (and also Atser): you creative masterminds. I admire how you found your own way in science. Thank you for inspiring me.

Bedankt Charissa en Anne voor samenwerking. Ik vond het fijn dat ik het project SPRINT@Work met jullie uiteindelijk tot een einde heb kunnen te brengen. Ik heb onze samenwerking binnen het project als een lichtpuntje ervaren tussen alle frustraties rondom het project.

Collega’s van Experimentele Psychologie, door de wekelijkse meetings, borrels en uitjes heb ik een leuke en leerzame tijd bij jullie gehad. Onderzoekers binnen experimentele psychologie hebben een creatieve, bedachtzame en kritische kijk op de wereld en ik voelde me er qua persoonlijkheid erg op mijn plek.

Lieve collega’s van Research Support, jullie hebben mij door de jaren heen op verschillende manieren geholpen en ondersteund. Eerst de technische kant binnen mijn onderzoeken en later als collega’s. Fijn dat ik af en toe nog een kopje koffie met jullie mag komen drinken (al kwam dat er niet echt vaak van tijdens COVID-19).

Collega’s van Gronings Perspectief, bedankt dat ik de afgelopen jaren met jullie heb mogen samenwerken. Deze toegepaste manier van onderzoek doen in een groot maar hecht team heeft me veel geleerd en ook weer nieuwe inzichten gegeven die ik nog heb kunnen verwerken in mijn proefschrift.

Als laatste wil ik mijn lieve vrienden bedanken. Kirsten, Lydia, Simone, Elise, Amarins, Chantal, Madnes 2016 – Astrid, Riekje en Sebastiaan, Random rijsje – Jelmer, Kim en Stijn, Wine-a-Longs – Wendelien, Lisette, Sigrid en nog maar een keer Jeffry, jullie hebben me altijd gesteund als uitlaatklep maar ook als afleiding met leuke uitjes.

Appendix I:
SPRINT@Work

Project description

SPRINT@Work is a project that focuses on sustainable employability and specifically on investigating how to keep the ageing population healthy and employable until and even beyond their expected retirement. To realise sustainable employability, workers were made aware of their condition by (1) objectively monitoring their cognitive and physical workload and capacity and (2) providing interventions to alter their behaviour to improve their work capacity or lower the workload. Moreover, sensor technologies were developed to enable monitoring, and interventions were created. All of these innovative technologies were validated in controlled laboratory studies, as well as in real-life working situations. Multiple aspects related to workload were investigated, such as cognitive and physical demands, individual responses to these exposures and feedback responses. SPRINT@Work comprised a broad consortium that included five knowledge institutes, 13 companies involved in the development of sensor technologies and seven pilot companies with workers and employers wishing to maintain a healthy working situation and willing to test the developed sensor and intervention technologies.

Outcomes

According to the needs assessment for workplace health promotion, several workers pointed out that priority should be given to monitoring fatigue, occupational heat stress and exposure to physically demanding jobs using sensor technologies (Spook et al., 2019). Mental fatigue negatively influences productivity during regular working activities. One way to detect productivity deteriorations in the office environment is by monitoring computer usage, for instance, by monitoring typing behaviour. Therefore, a study was performed to investigate whether typing indices can monitor deteriorations in attentional and memory processes by monitoring changes in neural activation (de Jong et al., 2018). This study was performed in a lab setting. The results showed that both younger and older participants became slower over time, which was reflected in the interkey interval. Moreover, younger adults became less accurate with prolonged task performance. However, they partly corrected their mistakes using the backspace key. Such changes in the typing indices were correlated with changes in neural activation; that is, those who showed larger deteriorations in attentional and memory processes also showed larger deteriorations in typing performance. The next question was whether the markers that were found to be susceptible to the effects of mental fatigue in a lab setting could also describe the behavioural dynamics in the work environment. To answer this question, typing performance data from a real-life office environment were analysed (de Jong et al., 2020). The results showed that the workers' typing speed decreased over time, which was reflected in a larger interkey interval. In addition, the workers used the backspace key more often. Interestingly, these effects of prolonged task performance interacted with the effects of time of day. That is, in the morning, workers were able to perform at a constant speed, with an increase in backspace keystrokes, whereas in the afternoon, both the typing speed, measured by the interkey interval, and accuracy, measured by the percentage of backspace keystrokes, decreased. These results suggest that even though these workers take precautions to counteract the effects of mental fatigue during the day (e.g., drinking coffee or taking breaks), the effects of prolonged task performance accumulate over the day. A different study investigated how consuming caffeinated beverages may help counteract the effects of mental fatigue (van den Berg et al., 2020). The results showed that, besides its general arousing effects, caffeine can enhance attention towards relevant information, which is specifically helpful in the work environment, where it is important to pay attention to specific tasks.

To monitor the energetic workload of physically active workers as a parameter of physical fatigue, a portable breathing gas analyser was developed and validated (patent pending; Roossien et al., 2021). The proof of concept of this analyser was found to be more valid than heart rate monitoring and more practical than indirect calorimetry with a mouth mask. Its users reported that the headset is more comfortable and more usable than mouth-mask systems. This proof-of-concept version is not yet as good as mouth masks; however, it has potential and provides opportunities for further professionalisation. This headset will be further developed and validated in a follow-up study together with a company specialised in breathing analysis, with the aim of making this system available not only for a large target group of workers, but also for rehabilitation and sports applications. To monitor occupational heat stress, a wearable core thermometer was developed and validated against a commercially available wearable thermometer. Despite the good usability of these thermometers, they are not yet suitable for measuring the core temperature while performing physically demanding jobs (Roossien et al., 2020). In a follow-up project, a new technology will be developed to fulfil the need for such a device. A suit equipped with sensors was used to investigate the exposure of physically demanding jobs. This suit monitored work postures and related back muscle activity and automatically calculated the net moment of the lower back with a specially developed artificial neural network based method. This technology was validated on different types of workers, and its function was also demonstrated. However, both the sensor system and software require further development before validating the function of the system in an operational work environment. A smart chair equipped with sensors was used to measure the physical load of office workers. Although the feedback signal did not improve the sedentary behaviour, this smart chair was a useful non-obstructive tool for monitoring the sitting behaviour of office workers (Roossien et al., 2020). Indeed, these systems and technologies will be further developed and validated in follow-up studies and will be made available for workplace health promotion.

To allow workers to benefit from such sensor and intervention technologies in the workplace, the effectiveness and effects of such technologies on employee autonomy were studied in two experimental field studies. The first study investigated the effects of real-time actionable feedback on workers' sitting and typing behaviour, in which the typing behaviour is considered a measure of fatigue. If a worker receives feedback messages on fatigue, they alter their typing behaviour almost immediately. However, if they receive feedback on their sitting behaviour, they alter their sitting behaviour only in the long term. This difference is explained by the fact that workers are considerably able to estimate their sitting bouts but hardly able to assess their level of fatigue. These findings show that workers are willing to alter their behaviour if they receive new information, as in the case of the typing behaviour. However, if they can self-monitor their behaviour, as in the case of the sitting behaviour, they show a learning effect over a longer period of time (Bonvanie, 2020).

The second study examined the effects of workers' use of health self-management applications in the work environment on their perceived autonomy in self-regulating their health-related behaviour. The results showed that workers experience a decline in their perceived autonomy either at home or at work or even both, depending on the type of feedback that they receive, and that this effect is strongest for employees with a high body mass index (BMI). Employees with a high BMI experience more negative emotions when they receive feedback pertaining to not reaching the given norm for physical exercise, and they become more aware of their work environment limitations that prevent them from altering their daily behaviour.

During SPRINT@Work, a context-sensitive perspective was used to contextualise ethical issues in both the development and implementation of sensor and intervention technologies for the work environment. The results of this context-sensitive analysis of ethics showed that the current legal framework for the privacy of workers limits the employers' opportunities to take full responsibility for the workers' health. This can, however, be solved using an agency-based approach, in which specific employees with clear roles (agents) have the power to use the personal data of other workers for specific reasons. Additionally, the autonomy of workers using sensor and intervention technologies is affected when the workers are not by default enabled to uphold their own norms and values but rather perceive the norms inherent to the design of these sensor and intervention technologies as pressing. These insights show that applying a context-sensitive approach of ethics may enhance the position of both workers and employers and provide valuable input for future research regarding technologies aimed at health improvement in the workplace.

Appendix II: Questionnaire with demographic and work-related questions

1. *What is your year of birth?*

2. *What is your sex?*
a. *Man*
b. *Woman*

3. *How many days do you work per week (on average)?*
a. *1*
b. *2*
c. *3*
d. *4*
e. *5*

4. *How do you go to work? Indicate the number of days per week per means of transport.*
.. days by car
.. days by public transport
.. days by bike
.. walking for days
Otherwise, namely:

5. *How many kilometers do you travel to work? (one way)*

6. *Do you use applications or products to monitor your health or exercise behaviour, or have you done this in the past? Think for example of a Fitbit, Nike Fuelband, Runkeeper or Strava*
a. *No, I have never used apps or products to improve my health or monitor exercise behaviour*
b. *Yes, I now use an app or product, namely:*
c. *Yes, I have used an app or product in the past, namely:*

7. *Do you exercise regularly?*
a. *Yes, I exercise several times a week*
b. *Yes, I exercise once a week*
c. *No, I don't exercise regularly*

8. *How many hours a week do you exercise on average?*

9. *Do you smoke?*
a. *Yes*
b. *No*

10. *What is your job category within the Faculty of Economics and Business?*
a. *Scientific staff (WP)*
b. *Support and management staff (OBP)*

11. *How would you describe your workplace distribution?*
a. *I usually work in the office (75-100% office time)*
b. *I regularly work in the office, but also sometimes in other locations (50-75% office time)*
c. *I sometimes work in the office, but also regularly in other locations (25-50% office time)*
d. *I usually work in other locations (0-25% office time)*

12. *How would you describe your task orientation or focus during your work?*
a. *If I have to finish a task, I focus entirely on that task and I do not want to be interrupted*
b. *When I have to finish a task, I mainly focus on that task, but I can be disturbed for important matters*
c. *If I have to finish a task, I do it while doing other important things and can be disturbed*
d. *If I have to finish a task, I do it while I also other important and do minor things, and can I be disturbed for anything*

13. *What health-related aids or programs at the RUG do you use or have you used in the past? (multiple answers possible)*

- *a.* *None*
- *b.* *BALANCE-activities*
- *c.* *the Health Check*
- *d.* *A high-low desk*
- *e.* *A custom office chair*
- *f.* *A bicycle office chair*
- *g.* *Otherwise, namely:*

14. *Have you been bothered or troubled by work-related physical complaints? (think: RSI, low back pain)*

- *a.* *Yes, I now suffer from work-related physical complaints*
- *b.* *Yes, I have had work-related physical complaints in the past*
- *c.* *No, I do not (have) suffered from work-related physical complaints*

15. *Have you been bothered or troubled by work-related mental complaints? (think: burnout, depression, stress)*

- *a.* *Yes, I now suffer from work-related mental complaints*
- *b.* *Yes, I have had work-related mental complaints in the past*
- *c.* *No, I do not (have) suffered from work-related mental complaints*

16. *How many hours per week (on average) do you teach in block IIB?*

17. *How many hours per week (on average) do you teach in block IA next year (2017-2018)?*

18. *Do you have plans to be away for a period (3 or more working days) during block IIB or IA, for example due to holidays or conference attendance? If so, when?*

19. *Is the computer at your workplace managed by the CIT or by yourself?*

- *a.* *Managed by the CIT*
- *b.* *Managed by myself*
- *c.* *I do not know*

20. *What is your mobile phone number? This number is only used for sending part of the feedback messages.*

Appendix III:
Weekly questionnaire

1. *We are curious if you have benefited from the feedback you have received by email and SMS in the past week. Please indicate to what extent you agree with the statements below.*

a. *I have read the feedback in the past week - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

b. *I find the feedback from the past week reliable - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

c. *The observations regarding my behaviour and environment corresponded to my own assessment - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

2. *The second set of questions is about the degree to which you plan or consider how to change your behaviour. So: "When it comes to my daily health and exercise behaviour ..."*

a. *I proceed systematically with the data from the sensors - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

b. *I set goals with regard to the sensor data - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

c. *I pay attention to planning activities based on the sensor data - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

3. *The third set of questions is about the extent to which you are making your planned change. So: "When it comes to my daily health and exercise behaviour ..."*

a. *it often happens that I do things differently than I have agreed with myself by using the sensors - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

b. *It often happens that I do things differently than I agreed with myself using the sensors - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

c. *what I think of using the sensors and what I do matches - I completely agree, I partly agree, I don't agree / I don't disagree, I partly disagree, I completely disagree.*

4. *How many hours does your employer expect you to work in a normal 7-day work week? (If this varies, please estimate the average)*

5. *In total, how many hours did you work in the past 7 days?*

6. *Please indicate below what percentage of these working hours of the past 7 days you have worked in your own office, at home or elsewhere. The total must be 100.*

At my own office: _______ (1)
At home: _______ (2)
Elsewhere: _______ (3)
Total: ________

7. *On a scale from 0 to 10, where the score 0 corresponds to the worst possible performance and the score 10 the best possible performance in your work, how do you rate the work performance of most employees who do similar work?*

8. *How would you rate your overall job performance on the same scale from 0 to 10 in the past 7 days?*

Appendix IV:
Alternative GLM-ANOVA approach

Table 1: *Standard statistical GLM-ANOVA analysis. The right column indicates the adjusted statistics when the between subject factor of order was included in the model (results of this analysis are reported when the p-value is reduced in comparison to the standard statistical GLM-ANOVA). P-values are reported as when the corresponding F-value is larger than 1*

RT = reward × caffeine × congruency

Effect	*DFn*	*DFd*	*F*	*P*	*(statistics when including the between subject factor order)*
reward	1	25	41.5	< .001	
caffeine	1	25	4.8	0.037	F(1,24) = 6.4, p = 0.02
congruency	1	25	195.4	< .001	F(1,24) = 313, p < .001
reward × caffeine	1	25	<1	n.s.	
reward × congruency	1	25	<1	n.s.	
caffeine × congruency	1	25	1.2	0.28	
reward × caffeine × congruency	1	25	<1	n.s.	
CNV = reward × caffeine					
reward	1	25	10.0	0.004	
caffeine	1	25	1.6	0.22	
reward × caffeine	1	25	6.1	0.021	F(1,24) = 6.2, p = 0.02
Alpha$_{sustained}$ = caffeine					
caffeine	1	25	7.1	0.012	F(1,24) = 7.9, p = 0.01
Alpha$_{evoked}$ = reward × caffeine					
reward	1	25	13.1	0.001	
caffeine	1	25	11.0	0.003	F(1,24) = 13.7, p = 0.001
reward × caffeine	1	25	< 1	n.s.	
z = reward × caffeine × marker (CNV or Alpha)					
reward	1	25	14.9	< .001	
caffeine	1	25	15.4	< .001	F(1,24) = 17.5, p < .001
marker	1	25	0		
reward × caffeine	1	25	2.8	0.11	
reward × marker	1	25	1.5	0.23	
caffeine × marker	1	25	5.4	0.029	F(1,24) = 6.7, p = 0.016
reward × caffeine × marker	1	25	4.2	0.051	F(1,24) = 4.4, p = 0.048
ROI$_{fc[400-500]}$ = reward × caffeine × congruency					
reward	1	25	18.2	< .001	
caffeine	1	25	3.6	0.069*	
congruency	1	25	26.4	< .001	
reward × caffeine	1	25	1.2	0.28	
reward × congruency	1	25	<1	n.s.	
caffeine × congruency	1	25	<1	n.s.	
reward × caffeine × congruency	1	25	3.7	0.064	
ROI$_{p[700-800ms]}$ ***= reward × caffeine × congruency***					
reward	1	25	6.4	0.018	F(1,24) = 6.9, p = 0.015
caffeine	1	25	1.8	0.20	
congruency	1	25	14.1	< .001	
reward × caffeine	1	25	5.0	0.034	F(1,24) = 6.0, p = 0.022
reward × congruency	1	25	<1	n.s.	
caffeine × congruency	1	25	1.7	0.21	
reward × caffeine × congruency	1	25	<1	n.s.	

** effect reached significance at p < 0.05 in the mixed model approach*

Table 2: *Summarizing the effects of order (i.e. the addition of caf × session interaction to the model). AIC (Akaike information criterion) and BIC (Bayesian information criterion) values (lower values reflect a better model in terms of explaining variance while weighing the number of model parameters) show that was no clear indication of order effects in any of the measurements. Both AIC and BIC attempt to solve overfitting by penalizing the number of model parameters.*

Behavioral models	df	ΔAIC	ΔBIC
RT ~ reward × congruency × caf + session	20	-	-
RT ~ reward × congruency × caf + caf × session	21	+1	+10
Alpha sustained models			
Alpha Sustained ~ caf + session	8	-	-
Alpha Sustained ~ caf × session	9	+2	+10
Cue models			
CNV ~ reward × caf + session	13	-	-
CNV ~ reward × caf + caf × session	14	0	+8
Alpha Cue ~ reward × caf + session	13	-	-
Alpha Cue ~ reward × caf + caf × session	14	+2	+11
Target models			
ROI$_{fc[400-500]}$ ~ reward × congruency × caf + session	20	-	-
ROI$_{fc[400-500]}$ ~ reward × congruency × caf + caf × session	21	+10	+19
ROI$_{p[700-800ms]}$ ~ reward × congruency × caf + session	20	-	-
ROI$_{p[700-800ms]}$ ~ reward × congruency × caf + caf × session	21	+22	+30