Exploring persuasive technology in the context of health and wellbeing at work

de Korte, E.M.

DOI
10.4233/uuid:cbec4bb0-a54c-424a-b1ba-d32c5567b366

Publication date
2021

Document Version
Final published version

Citation (APA)
de Korte, E. M. (2021). *Exploring persuasive technology in the context of health and wellbeing at work*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:cbec4bb0-a54c-424a-b1ba-d32c5567b366

Exploring persuasive technology in the context of health and wellbeing at work

Elsbeth de Korte

Exploring persuasive technology in the context of health and wellbeing at work

Cover design: Esther Scheide, www.proefschriftomslag.nl

Lay-out: Henny Knijnenburg

Printing by Ipskamp Printing

ISBN: 978-94-6421-304-1

Exploring
persuasive technology
in the context of
health and wellbeing
at work

Dissertation

For the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen
chair of the Board for Doctorates
to be defended publicly on
Thursday 22 April 2021 at 15:00 o'clock

by

Elsbeth Marieke DE KORTE
Master of Science in Human Movement Sciences, Vrije Universiteit,
Amsterdam, the Netherlands
born in Wageningen, the Netherlands

This dissertation has been approved by the promotors:

Composition of doctoral committee:

Rector Magnificus	chairperson
Prof. dr. P. Vink	Delft University of Technology, promotor
Prof. dr. ir. W. Kraaij	Leiden University, promotor
Dr. N.M. Wiezer	TNO, copromotor

Independent members:

Prof. dr. A. Windel	BAUA: Federal Institute for Occupational Safety and Health, Germany
Prof. dr. M.M.R. Vollenbroek-Hutten	University of Twente/ Ziekenhuis Groep Twente
Prof. dr. ir. R.H.M. Goossens	Delft University of Technology
Dr. ir. M. Melles	Delft University of Technology
Prof. dr. ir. G.W. Kortuem	Delft University of Technology, reserve member

Table of contents

1

General introduction

'Code red! Quit working! Your stress levels are too high and you did not have a break yet!'

'Congratulations! You have reached your physical activity goal for today!'

'Because of your night shift, it is recommended to go to sleep in about one hour.'

'Beep! You are entering a high risk zone with hazardous substances, please wear protective clothing!'

We are not able to imagine life without using technology. We use technology for almost every task in our daily life. In the work context, technology is everywhere around us. Over the last few decades, developments in Information and Communication Technology (ICT) have brought about many changes in work, and these changes will continue as technologies evolve. By its very nature, technology is dynamic, and continuous developments in technology are changing working conditions, work demands, work processes, the content of jobs, where work is performed, how organizations relate to their employees and the delivery of education and training. These changes will continue as new technologies emerge. On the one hand, developments in ICT have ensured that we work more efficiently. On the other hand, new ICTs have their downsides, and the question is raised as to whether we can, as humans, keep up with these rapid technological changes, both mentally and physically. As new technologies come into use, we need continuously to balance their risks and benefits for health and wellbeing at work (Salvendy, 2012; Schwab, 2016; Manyika et al., 2013).

Recent technological developments, like persuasive technology, offer radically new possibilities as interventions for health and wellbeing at work. Persuasive technologies are interactive systems developed to change the attitudes or behaviours or both of users through persuasion and social influence. In addition to monitoring or (self-)tracking, persuasive technology uses an influencing algorithm and actuators to provide active feedback to the user (Fogg, 2003; Fogg et al., 2009; Orji and Moffatt, 2018).

Collected data form the basis for persuasive technology. Manual input or sensors installed on or embedded in people and their environments, provide data from which the subject's physiological state and behaviour can be derived. Persuasive technologies can quantify users behaviour, emotions, physical and mental activity and bodily functions. Smart software can analyse these data and discover patterns that are invisible to the user. By giving feedback to the user, these technologies give insights and recommendations and thus help the user to make everyday choices in a variety of areas such as lifestyle, health, financial housekeeping or environmental awareness. Persuasive technology has already shown promising results in a broad range of health behaviour change. However, within the context of work, it has yet to become standard practice (Kool et al., 2015; Orji and Moffatt, 2018; Van Den Broek, 2017).

This thesis explores the potential of persuasive technology for health and wellbeing at work. To gain insight, we need to understand the worker, the tasks, the interaction between worker and technology and the working environment in which the technology will be applied. Besides technological challenges, this specifically poses challenges in the field of human sciences such as human factors and ergonomics, work and organizational psychology, behavioural sciences, user-centred design, human computer interaction and design engineering.

The occupational context in which persuasive technologies are being applied poses additional constraints concerning the design and implementation of these technologies. Better occupational health and wellbeing requires different behaviour. To change worker attitudes or behaviours, it is important to decide what is monitored, which methods have to be used and how to interpret what is monitored. Individual workers also have different needs and personal goals. For instance, some workers suffer from sleeping problems because of their night shifts, while others need to better balance their work and private life. The working environment in which persuasive technologies are being applied also leads to additional constraints for the design of persuasive technologies. For example, design requirements differ for workers performing office tasks, working at assembly lines or working in clean rooms. Next, it has to be determined which actions are appropriate to take and how the worker can be personally motivated. Finally, we need to study the societal impact of persuasive technology and how to apply it in a responsible matter.

These challenges will be addressed in this thesis from four perspectives: (1) whether these technologies are theory based, (2) how best to assess the effectiveness of these technologies, (3) whether they are actually effective and (4) what the societal impact of these technologies is. This thesis thus puts the results of various studies in the broader context of persuasive technology for health and wellbeing at work, providing an overview what has been achieved and directions for future research.

1.1 Motivation

1.1.1 Background: trends in technology and work

Trends in technology

Technology has long been part of our work, and the impact of technology on working life has been an ongoing topic in science for decades (Manyika et al., 2013; Schwab, 2016). Since the start of the industrial revolution more than 250 years ago, technology has brought many changes and enabled the use of new methods for performing tasks. The World Economic Forum (Schwab, 2016) describes the current technological developments as the 'fourth industrial revolution'. The first industrial revolution spanned from about 1760 to around 1840 with the invention of the steam engine. The second industrial revolution, which started in the late 19th century into the early 20th century, made mass production possible through the use of electricity and assembly lines. This revolution also brought telephones and airplanes into wide use. The third industrial revolution began in the 1960s and can be called the computer or digital revolution, because it was catalysed by the development of semiconductors, mainframe computers (1960s), personal computing (1970s and 80s) and the internet (1990s). The fourth industrial revolution began at the start of the 21st century and builds on the third, digital revolution. It is characterized by a much more ubiquitous and mobile internet, by smaller and more powerful sensors and by artificial intelligence and machine learning. This 'fourth industrial revolution' is, unlike previous revolutions, evolving at an exponential rather than a linear pace (Manyika et al., 2013; Schwab, 2016). Technology has become increasingly 'personal'. Computers were first located in large rooms, then on desks and later moved onto people's laps. Technology can now be found in people's pockets in the form of mobile phones; the next step will be technology integrated directly into clothing and other accessories (Schwab, 2016).

Several important technological developments and trends could be described as significant drivers of change that affect the nature of work. These developments are characterized by three major advances: (1) in miniaturization and portability, (2) in computing power and speed and (3) in ICT services and infrastructure. This thesis contains work which has stretched out over about ten years. During that period, technology has changed and so has the terminology. The following interrelated visions and concepts have frequently been used to describe trends in technology; many of these visions and concepts are interrelated, overlapping and building upon each other.

› **Mobile Internet**

The mobile internet can be defined as a combination of mobile computing devices, high-speed wireless connectivity and mobile applications (Manyika et al., 2013). The basic properties of mobile computing can be summarized as: portable (small battery-operated handheld devices), remote wireless connectivity, networked (remote data and service access), location sensitive and secure (encryption based with authentication and conditional access; Aarts and De Ruyter, 2009). Today, smartphones and tablets are the main devices used to access the mobile internet, but new forms are constantly emerging. In 2016, almost three and a half billion people, or 46% of the world population, had an internet connection. The number of mobile phone users was even higher: more than four and a half million people own a mobile (or smart) phone (Calvo et al., 2016). In the coming years, mobile internet devices could well be smaller, far more powerful, more intuitive, wearable and packed with many types of sensors. For instance, smartphones and tablets contain multiple sensors, including accelerometers and location sensors. More recent models include sensors that monitor temperature, humidity and air pressure, as well as sensors that detect screen proximity (Manyika et al., 2013).

› **Pervasive/ Ubiquitous Computing**
Pervasive computing provides a new view on mobile computing. It is a vision that stresses issues related to interoperability and seamless interconnectivity. Ubiquitous computing can occur using any device and in any location, but the focus is shifted towards the software properties of services, rather than on the device properties as in mobile computing. The properties can be formulated as: ubiquitous (overly present); interactive (multi-modal user interfaces); interoperable (plug and play with seamless integration and access); distributed (simultaneous access to resources including databases and processing units); and scalable (adaptation of resources; Aarts and De Ruyter, 2009). Pervasive computing has been enabled by the development of **cloud computing**: technology which makes it possible to deliver any computer application or service over a network or the internet. Behind the scenes, this requires a complex system of servers and storage systems. With cloud technology, the bulk of the computational work can be done remotely and delivered online, reducing the need for storage and processing power on local computers and devices. Because apps rely on cloud resources, the cloud has also been a major driver of smartphone use. It is expected that future mobile services will become even lighter and faster through cloud computing. Cloud computing provides on-demand self-service and availability anytime and anywhere. This allows workers across the world to work together by sharing data and information, enabling flexible working and remote working (Manyika et al., 2015; Stacey et al., 2017).

› **Ambient Intelligence**
Ambient intelligence (AmI) builds further on the concepts of mobile and pervasive computing by involving the entire environment: electronic systems embedded in

the everyday environment that are sensitive and responsive to people in a seamless, unobtrusive and often invisible way. The word ambient refers to the unobtrusive integration of technology in every-day objects and environments. The term intelligence reflects specific forms of social interaction: technology should be able to recognize people, to personalize individual preferences and to adapt to users over time. The purpose of AmI is to improve productivity, creativity and pleasure through enhanced user-system interaction. Compared to mobile and pervasive computing, the emphasis is on greater user friendliness, more efficient service support, user-empowerment and support for human interactions (Aarts and De Ruyter, 2009; Cook and Song, 2009).

› **Internet of Things**

The internet of things (IoT) might be described in the same way as AmI, but is primarily concerned with the physical objects involved. Increasingly, the connected world includes physical objects. People, products, services and places are being equipped with networked sensors and actuators that enable them to monitor their environment, report their status, exchange data, receive instructions and even take action based on the information they receive. This is what is meant by the IoT, which is made possible by connecting technologies and various platforms and using sensors that are becoming increasingly smaller, cheaper and smarter. Today, there are billions of devices around the world connected to the internet, such as smartphones, tablets and computers, but also vehicles, manufacturing equipment, wearable heart monitors and railway tracks (to schedule maintenance activities). Their numbers are expected to increase dramatically over the next few years (Manyika et al., 2013; Schwab, 2016; Swan 2012). There are three stages in IoT applications: (1) capturing data from the object or environment with **sensors** (from simple location data to more complex data from different sources); (2) aggregating, processing and modelling data, for instance using machine learning or big data techniques (**reasoning**); and (3) **acting** on that information by taking immediate action or collecting data over time to improve processes or behaviour. IoT technology ranges from simple identification tags to complex sensors and actuators. Examples include movement (via accelerometer), location (via GPS), heart rate and heart rate variability, electromyography (EMG), hazardous substances, temperature and sound or combinations of several sensing elements (Kraaij et al., 2019; Manyika et al., 2013; Swan 2012, Stacey et al., 2017).

Within the IoT, we can distinguish the specific group of wearable devices, or **wearables**. This refers to devices which can be worn on the person or incorporated into clothing or even the body. Miniaturization and increased battery life have led to the development of devices like the Fitbit, Apple Watch and Microsoft Band.

Examples of sensors integrated in clothing are sport shirts developed to provide real time workout data, or sensors integrated in firemen's clothing to detect location or hazardous substances. A rather new product category is that of disposable patches and electronic tattoos that are worn for days that, for example, measure blood chemistry, hydration level or body temperature. Microsoft HoloLens is an example of a wearable with augmented reality (Schwab, 2016; Stacey et al., 2017; Swan 2012).

› **Big Data**

Big data refers to a combination of three trends: (1) the increasing rate of data generation (through increasing levels of global connectivity and networking); (2) improving data storage; and (3) advancing data analysis (new analytical techniques are being developed to manage large data sets and to derive new insights into behaviours). Big data technologies will enable faster decision-making and increase efficiency in work processes in a wide range of industries and applications. Big data analytics could also allow vastly improved analysis of historical and current occupational health and safety data, which might help to clarify the causes of (occupational) health problems and diseases and eventually to predict them (Stacey et al., 2017).

One particular way to analyse these data is **machine learning.** Machine learning is the process by which software applications learn to draw conclusions from patterns they recognize within massive data sets. Moreover, these algorithms can 'learn' more and get smarter as they go along; the more they process big data, the more refined their algorithms become. This makes it possible to automate workers' tasks, achieving performance near to that of a human or, indeed, a superhuman. Many current machine-learning approaches are simulated aspects of the human brain. Neural networks are inspired by brain structures via interconnected layers of artificial neurons, which adaptively strengthen or weaken their interconnections based on experience. Deep learning technologies make use of algorithms that form a learning hierarchy in which higher-level concepts are defined using layers of lower-level concepts. Some machine-learning techniques identify their own categories and concepts (e.g. by cluster analysis; Manyika et al., 2015). Machine-learning techniques are opening the way for new, much more customized and personalized services and predictions, which can benefit consumers, including workers (Schwab, 2016).

These technological trends have also led to developments in monitoring technology. Increasingly, self-monitoring or self-tracking has become popular, particularly to improve personal health and professional productivity, with behaviour change as important means.

- **Quantified Self**
  The quantified self (QS) is a movement of people who monitor or track their behaviour, thoughts, feelings and other aspects of their daily life. This is also called life-logging or self-monitoring. The process of recording one's own behaviour dates back to the 1970s and was traditionally employed in clinical and research settings to serve as an assessment or treatment within the course of therapy. At QS Meetups, people talk about their experiences organized along the so-called three prime questions: What did you do? How did you do it? What did you learn? These conversations are recorded and uploaded to the quantifiedself.com blog for sharing and exchanging knowledge. Because sensors have become smaller and integrated within mobile devices, self-monitoring has become mainstream. Wearables make it easy for people to track numerous types of data, inside and outside the clinical setting. Now, self-monitoring has been widely embodied in the design of sensing and monitoring applications (Choe et al., 2014).

- **Personal Informatics**
  Similar to the QS movement, personal informatics (PI, also known as personal analytics) allows users to collect data and review personally relevant information. PI focuses on systems that not only allow users to gather data, like the QS, but also seek to facilitate favourable changes in behaviour based on the logged data. The behaviour changes are thereby data driven: users self-track and examine their data and change their behaviour based on their personal insights. This is called the 'self-improvement hypothesis', and it represents the prevailing intention in designing such systems, as well as the most common reason for users to adopt them (Choe et al., 2014; Kersten-Van Dijk et al., 2017).

- **Persuasive Technology**
  Already in the 1970s and 1980s, computer systems were designed to motivate health behaviours and work productivity. However, it was only in the late 1990s – during the rise of the internet – that people began to make interactive systems capable of motivating and influencing users. Persuasive technology is being developed to change the attitudes or behaviours (or both) of users through persuasion and social influence. When compared to the QS movement and PI, persuasive technology uses additional influencing algorithms and actuators to provide active feedback to the user. Persuasive strategies are hard to invent and apply, but when achieved, they have proven to be successful in health behaviour change (Orji and Moffatt, 2018; Van den Broek, 2017). An important reason for this is that persuasive technology does not apply coercion: the intrinsic motivation of the user is crucial. Behaviour change can also be quite subtle, possibly without the users' complete awareness. Persuasive technology is applied in several domains, for example commerce (buying, branding), personal finance (adherence

to a personal budget), safety, preventive health and disease management, but not yet in the occupational setting. When applied in the occupational setting, it might increase the possibilities for self-management among workers to enhance health and wellbeing at work (Fogg et al., 2009; Orji and Moffatt, 2018; Van den Broek, 2017).

Alternative terms for persuasive technology are **behaviour change support system (BCSS)** or **digital behaviour change intervention (DBCI)**. BCSS is defined as a sociotechnical information system with psychological and behavioural outcomes designed to form, alter or reinforce attitudes, behaviours or an act of complying without using coercion or deception (Oinas-Kukkonen, 2013). DBCI is used to refer to an intervention that employs digital technology to promote and maintain health, through primary or secondary prevention and management of health problems. The technologies used can include the internet, wearables and IoT devices that can provide intelligent monitoring and feedback as and when needed (i.e. 'just-in-time adaptive interventions' or 'ecological momentary interventions'). DBCIs are typically automated, interactive and personalized and can be used to promote health by supporting behaviour change or decision-making or to enhance physical and mental wellbeing (West and Michie, 2016; Yardley et al., 2016).

› **mHealth**
Mobile health or mHealth covers medical and public health practices and might be seen as the mobile variation of eHealth (electronic health). mHealth involves mobile and wireless technology to support the achievement of health objectives. It includes mobile devices such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), smart watches and other wireless devices or wearables. It also has applications such as lifestyle and wellbeing apps that may connect to medical devices or sensors (e.g. bracelets or watches) as well as e-coaching, health information, medication reminders provided by SMS and telemedicine provided wirelessly (European Commission, 2014).

Trends in work – related to technology
Over the last decades, developments in ICT have brought about many disruptive changes in the types of jobs available, how we work, where we work, access to information, use of devices, organizational structures and delivery of education and training. These changes will continue as technology and demographic and social patterns evolve (Dul et al., 2012; Eurofound and the International Labour Office, 2017; Salvendy, 2012; Schwab, 2016; Stacey et al., 2017; Stacey et al., 2018). The new technology revolution will provoke more disruption than the previous industrial revolutions because of its speed (everything is happening at a much faster pace than

before), breadth and depth (many changes are occurring simultaneously) and the complete transformation of entire systems (i.e. platform economy; Schwab, 2016).

ICT has had a growing impact on work since personal computers first entered the workplace. The first computers were aimed at supporting existing jobs. Many different categories of work, particularly those that involve mechanically repetitive and precise manual labour, have already been automated. Many others will follow, as computing power continues to grow exponentially. Developments in ICT have caused society to change from an industrial to a knowledge economy. Increasing numbers of workers spend their days in front of a computer screen or a mobile device. Today, the Dutch working population works with a computer 4 hours a day on average, and 38.4% of the working population works with a computer 6 or more hours a day. These numbers are still rising (Hooftman et al., 2019; Manyika et al., 2013; Schwab 2016).

The emergence of new technologies, such as the IoT, big data and cloud computing, also enable new business models and offerings (e.g. the platform economy, which is economic and social activity facilitated by platforms such as Airbnb or Uber). ICT is thus no longer seen as a specific separate sector, but rather as a provider of essential services for all sectors of our economy and society. This has led to a blurring of the boundaries between different industries and sectors. In jobs where a physical presence is required – such as manufacturing – computer control, increased automation and the use of robots are changing the nature of work (Salvendy, 2012; Schwab, 2016, Stacey et al., 2017).

In other jobs, ICT has changed the necessity of physical presence, enabling remote and flexible work, including working from home or while travelling. People no longer need to be located in the same place to communicate and exchange documents and information. Their workplace can be anywhere as growing wireless networks allow people to carry out their work. This situation is also referred to as new ways of working (NWW; Pot et al., 2012). It leads to high flexibility in working hours, quasi-continuous availability and fading borders between work and private life. In addition, technology-mediated learning is emerging as the preferred method for training employees (Bailey and Kurland, 2002; Czaja and Sharit, 2009; Dul et al., 2012; Eurofound and the International Labour Office, 2017; Laihonen et al., 2012; Pot et al., 2012; Robertson and Vink, 2012; Salvendy, 2012; Schwab, 2016).

Trends in technology not only change what we do but also who we are. Technology has changed the perception of ourselves, including ourselves as workers. It affects our identity and its many related facets – our sense of privacy, our notions of ownership, the time we devote to work and leisure, as well as how we develop our careers, cultivate our skills and relate to others, for instance our colleagues and employers. Technology also provides us with opportunities for personal development. Should we use the advances in technology to make ourselves better workers? What does 'better' mean? To be free of diseases (Schwab, 2016)? Technology stimulates us to be constantly aware of our health status: will we ever reach ultimate 'health'? Only a decade ago, we did not question our health on such a frequent basis; we only questioned our health when we experienced a health problem. Although interactive technologies afford conveniences and efficiencies, the overall contribution of technology to wellbeing has been a topic of ongoing debate. Some have highlighted how new technologies inform, liberate and enrich our lives, whereas others suggest that the new technologies too often impoverish our experiences and self-regulation of behaviours, distracting us from relationships and compromising health-promoting activities such as physical activity and sleep (Calvo et al., 2016).

These trends all give rise to both challenges and opportunities for a number of areas, including occupational health and wellbeing. Technological developments pose new risks. However, there is also significant potential, such as technologies that could help to identify risks and better resolve them. As new technologies come into use, we need continuously to balance their risks and benefits to support these changes and to empower workers to deploy new work styles and patterns (Manyika et al., 2013; Salvendy, 2012; Schwab, 2016; Stacey et al., 2018).

1.1.2 Problems associated with trends in technology and work

Advances in technology lead to an increase in occupational health and wellbeing risks, because of the physical and psychosocial demands such advances can place on workers (Stacey et al., 2018).

Considering physical demands, it can be seen that advances in automation and (remote) computer control have resulted in a shift away from occupations that require moderate intensity physical activity to occupations that are largely sedentary (Church et al., 2011; Hallal et al., 2012; Stacey et al., 2018). In the early 1960s, almost half of private industry occupations in the US required at least moderate intensity physical activity, and now less than 20% demand this level of activity. Over the last 50 years in the US, it is estimated that work-related daily energy expenditure has decreased by more than 100 calories. Because the time spent engaged in physical activity during working time represents a large portion of the total hours in a week, this reduction in work-related energy expenditure accounts for a significant portion of the increase in body weight over that same period (Church et al., 2011). Worldwide, 31.1% of adults are physically inactive (Hallal et al., 2012). It is estimated that physical inactivity accounts for 6% of global deaths (Lee et al., 2012; Van der Ploeg et al., 2012). The work itself has also become more sedentary. Sedentary behaviour is an independent risk factor for all-cause mortality, independent of physical activity. Deleterious health effects such as cardiovascular disease, cancer, type 2 diabetes, obesity and overweight are associated with sitting for long periods (Bennie et al., 2013; Chau et al., 2010; Van der Ploeg et al., 2012). Sedentary behaviour is usually defined as the time spent sitting. Increasingly, many working adults spend large amounts of time sitting each day (Chau et al., 2011; Proper and Hildebrandt, 2006). Like physical activity, sedentary behaviour occurs in different domains: for example, at work, during leisure time or while commuting (Hallal et al., 2012). The average European person spends 309 minutes a day sitting (5.2 hours; Bennie et al., 2013). Of the Dutch working population, 32.1% report sitting for more than 7.5 hours per day (Loyen et al., 2016). Research has shown that employees with unhealthy lifestyle behaviours such as low physical activity levels and sedentary behaviour, are less productive at work (presenteeism), have decreased work ability and take more sick days (Proper et al., 2006; Robroek et al., 2011; Robroek et al., 2013; Rongen et al., 2013).

The increased use of computers has caused an increase in ergonomic risks as well (Stacey et al., 2018; Wahlström, 2005). Research shows a relationship between desktop computer use (duration and posture) and development of musculoskeletal symptoms (Gerr et al., 2004; IJmker et al., 2007; Wahlström, 2005), with reported prevalence rates for musculoskeletal symptoms between 10% and 62% (Wahlström, 2005). Ergonomic risks further increase with the mobile internet, which allows people to work anywhere, including at home, in public places or transport, and the use of mobile devices that are not suitable for use for long durations and therefore causing complaints to the neck and upper limbs (Stacey et al., 2018). Static postures and repetitive movements, physical work demands that are associated with computer work, are related to presenteeism, decreased work ability and sickness absence (Martimo et al., 2009; Van den Heuvel et al., 2010). Productivity losses of 15% have been reported among computer workers with musculoskeletal complaints while at work (Hagberg et al., 2002). In addition, effects on presenteeism may occur more frequently and may be larger than the effect on sickness absence (Hagberg et al., 2002; Van den Heuvel et al., 2010; Van den Heuvel et al., 2007).

In addition to physical demands, advances in technology have affected the psychosocial demands of workers: NWW initiated a revolution in the way employees interact with each other, including remote collaboration and increased possibilities for sharing information. Examples of the drawbacks of NWW include managing the large inflow of information, unanticipated tasks generated by new incoming messages, lack of control over incoming messages, interruptions and task switching associated with email, perceived pressure to respond quickly, continuous availability of mobile devices resulting in extended workdays, permanent connectivity and a disturbed work-life balance, a risk of misunderstandings between colleagues (or even cyber-bullying) and decreased perceived social contact and support because of increased electronic communication (Demerouti et al., 2014; Van den Heuvel et al., 2018). Some workers lack the necessary skills to be able to use advanced technology and cope with change. There is also the risk for task deprivation and cognitive underload with automated work processes, for instance when operators' roles become supervisory with only occasional intervention, resulting in losses of task variety, concentration and alertness (Stacey et al., 2018). Due to the interconnectedness of things and people, the algorithmic management of work and workers may lead to loss of control. For example, monitoring devices may have a negative impact if workers feel performance pressure. Also, the lack of transparency of deep learning algorithms makes it difficult for workers to understand and interact with the system and to respond in case of system failures (Stacey et al., 2018).

From the literature, it can be seen that these psychosocial demands are associated with negative work consequences such as health complaints, sickness-related absence, decreased work ability and productivity loss (Alavinia et al., 2009; Eurofound and EU-OSHA, 2014; Karlsson et al., 2010; Niedhammer et al., 2012; Nieuwenhuijsen et al., 2010; Van den Heuvel et al., 2010). In Europe, 25% of workers say they experience work-related stress due to high psychosocial and cognitive demands for all or most of their working time, and a similar proportion reports that work affects their health negatively (Eurofound and EU-OSHA, 2014; Van den Heuvel et al., 2018).

Challenges in occupational health and wellbeing are further increased because of the ageing workforce. In Europe, the proportion of older people in the working force is increasing more than in other continents (Dul et al., 2012). In Europe, the proportion aged 55 and over is 26% (Irastorza, 2019) and is estimated to reach 34% by 2025 and 52% by 2060 (Eurofound, 2011), leading to a shrinking labour pool (Dul et al., 2012; Eurofound, 2011; Eurostat Newsrelease, 2012; Salvendy, 2012). European member states have therefore increased the official pension age. However, this only will be successful if workers maintain their physical and mental health into retirement. As workers age, their physical, physiological and psychosocial capabilities change. They are also likely to experience a range of diseases associated with increasing age such as cardiovascular diseases and type 2 diabetes. In addition, due to cumulative exposure, the effects of demanding work develop over time; extending working years may therefore increase the risk for occupational health problems that will need to be managed in the workplace (OECD, 2014; Schwab, 2016). Although older workers may have reduced capabilities in some respects, they also have more developed capabilities such as mental growth (strategic thinking, language skills, motivation, commitment, work expertise) and social capabilities (ability to adjust their behaviour). However, there are large variations among older age groups and these can become more pronounced with age. In general, an individual's performance remains stable throughout their working career, and many workers compensate for losses in physical health and cognitive capabilities though more extensive work experience and knowledge (Dul et al., 2012). When retirement ages increase, it might be expected that it will become harder for workers to compensate for these losses. This leads to the overarching question of how workers in general – and specifically older workers – can work up until higher ages in good health (OECD, 2014; Schwab, 2016).

Next to physical and psychosocial risks, we also have to be aware of the practical challenges of technological developments. First, the pace of change due to advances in technology might make it difficult to keep occupational health and wellbeing regulations up to date. Second, changing business models and employment hierarchies due to increased online and flexible working arrangements and the introduction of algorithmic management (e.g. Uber) have the potential to disrupt current mechanisms for management of health and wellbeing (Stacey et al., 2018). Increasing numbers of workers are being treated as self-employed and might fall outside existing regulations for occupational health and wellbeing (Stacey et al., 2018). Finally, we need to address data privacy and ownership. On the one hand, monitoring requires data storage, processing and analysis. Most likely, when data concern our health and wellbeing, such data are very personal and not meant to be shared indiscriminately (Stacey et al., 2018; Van den Broek, 2017). This could include, for instance, data on sleep patterns to coach shift workers on dealing with negative effects of night shifts. On the other hand, workers may lack understanding of what data are collected and for what purpose, leading to feelings of lack of control over their data. This might negatively influence their willingness to use technology. Therefore, the security of data processing has to be combined with transparency in its use. Principles for secure and transparent data processing have been regulated by the EU General Data Protection Regulation (GDPR).

Advances in technology can create new risks and challenges for workers. However, technology in itself is neither good nor bad. It might offer new opportunities to reduce health and wellbeing risks or to better manage them (Stacey et al., 2018).

1.1.3 Persuasive technology as a potential solution to enhance health and wellbeing at work

Although new possible risk factors may appear when using the new technologies described earlier, they also have real potential to drive improvements in working life. For instance, these technologies create the opportunity to develop new kinds of interventions. Technology can play an important role in solving the question of how to motivate and stimulate workers to adopt healthy, productive and safe working behaviours (European Commission, 2014; Kraaij et al., 2019; Manyika et al., 2013; Manyika et al., 2015; Salvendy, 2012; Schwab, 2016; Stacey et al., 2017; Stacey et al., 2018).

A healthy workforce is a more productive workforce, and effective interventions that strengthen health and wellbeing of workers will lead to individual benefits as well as organizational profits (Berry et al., 2011; Robroek et al., 2011; Van den Heuvel et al., 2010). Besides that, the workplace is a fruitful setting for health promotion because of the presence of natural social networks, the possibility of reaching a large population, and the fact that people spend a great deal of their time at work (Hutchinson and Wilson, 2012; Rongen et al., 2013; Van der Klink et al., 2001). For these reasons, much effort has been put into the development and evaluation of interventions in the workplace setting. This includes activities to change individuals' risks, attitudes, behaviour and awareness, as well as comprehensive interventions such as workplace health promotion programmes (Chau et al., 2010; Eurofound and EU-OSHA, 2014; Rongen et al., 2013; Wierenga et al., 2013). However, research shows that workplace interventions may be beneficial, but not all interventions are useful or their overall effects are small (Bhui et al., 2012; Cancelliere et al., 2011; Chau et al., 2010; Eurofound and EU-OSHA, 2014; Hamberg-Van Reenen et al., 2012; Lamontagne et al., 2007; Richardson and Rothstein, 2008; Rongen et al., 2013; Speklé et al., 2010; Van der Klink et al., 2001; Wierenga et al., 2013). This calls for an exploration of new approaches for health and wellbeing at work and, in particular, **persuasive technology** as a potential intervention, which is studied in this PhD thesis. Figure 1 shows a graphical representation of persuasive technology for health and wellbeing at work.

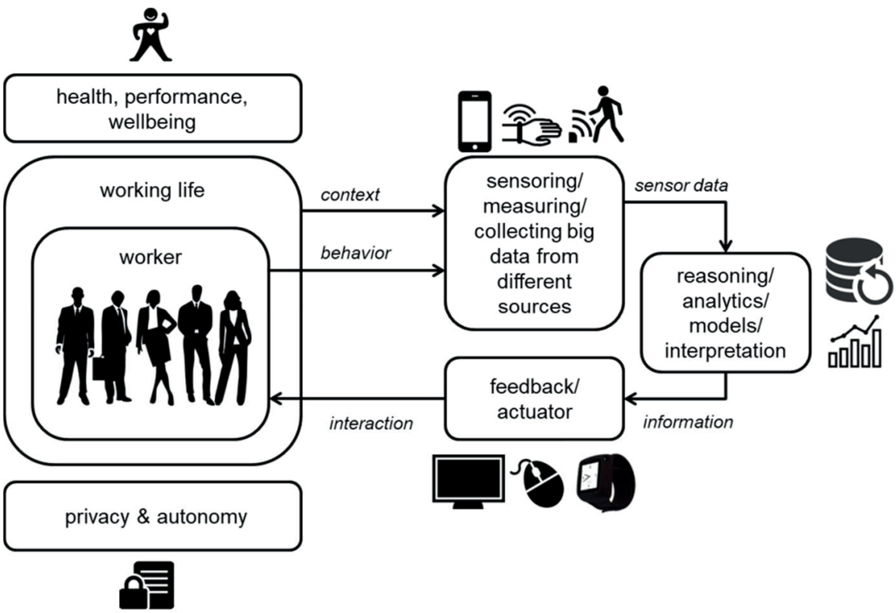

Figure 1 Graphical representation of persuasive technology for health and wellbeing at work

Persuasive technologies are developed to change users' attitudes or behaviours through persuasion and social influence as described above. One of the examples of persuasive technology that aims to change the behaviour of office workers is the multi-modal mouse with tactile feedback by means of a pin through a hole in the left mouse button and force-feedback using an electromagnet. This mouse is intended to optimize task performance, such as response time, precision positioning and reducing error rates (Akamatsu et al., 1995). Another example is the Feel It Mouse (Immersion Corporation, San Jose, CA, USA), a haptic force-feedback computer mouse (Dennerlein and Yang, 2001) designed to optimize performance as well as to reduce self-reported musculoskeletal pain and discomfort.

The Smart Chair (BMA Ergonomics) is an example of persuasive technology integrated within an office chair to measure sitting postures as well as time sitting and thus sedentary behaviour. Based on sensor data analytics and algorithms, the chair gives feedback by means of a tactile feedback signal as a sign to suspend sitting, for instance by standing up or moving around. In addition, feedback on postures is given via a small interface attached to the chair (Netten et al., 2013).

An example of technology aimed at changing the work-rest behaviour of office workers is rest break software (Slijper et al., 2007; Van den Heuvel et al., 2003). Rest break software introduces rest breaks and sometimes exercises when computer use has been too intensive or too prolonged. In general, the early versions of rest break software were not rated as user friendly, because they often disturbed the user during working tasks. Currently, research aims to discover the opportune moments in which a worker might be open for feedback tips to take a rest, such as by using context recognition to identify the closing of a computer program or document (Kaur et al., 2020; Kraaij et al., 2019; Sappelli et al., 2016).

Various features make persuasive technologies valuable for interventions at work. First, as portable devices (such as a smartphone or wearable) they can be switched on and remain with the owner throughout the day to self-track user's behaviour and feelings continuously and unobtrusively (Choe et al., 2014; Kraaij et al., 2019; Swan, 2012). In this way, they offer the opportunity to bring interventions into important real (working) life contexts where people make decisions about their health and encounter barriers. Second, persuasive technology may provide cheaper, more convenient interventions that are unavailable elsewhere, with a large reach. Third, connectedness facilitates the sharing of behavioural and health data with health professionals or peers. The increasing ability to use sensors to infer context such as user location, movement, emotion and social engagement has also raised the prospect of continuous and automated tracking of health-related behaviours and timely, tailored, adaptive and anticipatory interventions for specific contexts (Aarts and De Ruyter, 2009; Dennison et al., 2013; Middelweerd et al., 2014). Therefore, these technologies might allow workers to better understand their behaviour and improve it. In addition, technologies support a participative role for users, while enhancing their responsibility over their own health and performance, thus contributing to the empowerment of workers (European Commission, 2014).

In dangerous work environments, (wearable) sensors can prevent accidents and injury by sounding an alarm or shutting off machinery when a worker approaches danger. Sensors can also protect workers' health by tracking exposure to harmful chemicals, radiation, noise or vibration. In some cases, sensors might help to prevent injuries from happening by, for example, detecting possible back pain in a worker who is moving slowly after lifting a heavy object incorrectly or alerting a worker when working at height. By providing employers information on the safety of work environments, employers might be able to better manage risks and substantially reduce illness, injury and death (Manyika et al., 2015; Stacey et al., 2018).

Based on the characteristics of technology, preliminary findings in research and expected impact, using persuasive technology for health and wellbeing at work is promising. It might play an important role in answering the question of how workers can be motivated and encouraged to perform healthy behaviour, how they can learn and develop and how they can actively contribute to and control their own health. However, research is needed to examine its potential and to assess when, where and for whom applications are effective (Kumar et al., 2013; Klasnja et al., 2011; Pagoto and Bennett, 2013).

1.1.4 Scientific challenges

The field of persuasive technology for health and wellbeing at work is in development, and this will continue in the future with many topics to study. To move these developments forward and to increase the impact of persuasive technology in the context of work, there are four key challenges to responsible implementation.

Challenge: theory and evidence base

Persuasive technology seems convenient, but can we trust it? Are these technologies based on theories and a solid evidence base? Persuasive technologies gather data by means of sensors, questionnaires or self-reporting. The quality of these methods depends on a variety of factors, including validity (does it measure what it claims to measure?); reliability (does repeated use generate the same values?); and accuracy (how precise is the measurement?). Data quality has consequences for the interpretation of the users' behaviour and for the feedback based on these data. Low quality may result in users who quit using the tool. Along the steps from sensing to feedback, interpretation takes place, and the quality of these steps is unknown. There is also the question of whether it is applicable to use certain sensors to quantify certain behaviour. Next, collected data are being used to analyse users' behaviour to give feedback. To do this, persuasive technology draws on a variety of assumptions, theories and standards. The extent to which technologies tend to be built upon behaviour change theories and on the theoretical models related to occupational health and wellbeing remains questionable. Therefore, it might be useful to know which theories or models are being used in persuasive technology. Given the theories and models used, what is the scientific proven effect and is there consensus among findings (Kool et al., 2015)? In this thesis, we look into the incorporation of theory on behaviour change in persuasive technology.

Challenge: research methods

Persuasive technologies develop continuously; new versions appear within months or even weeks, which raises the question of whether traditional research methods can be used to evaluate these technologies. Persuasive technologies challenge the way we conduct research. Usually, the development of technology is characterized by a highly iterative process. The rapidly evolving nature of technologies and their uptake means that some components are continuously improved during a trial, which poses a threat to internal validity (Kumar et al., 2013; West and Michie, 2016). Therefore, evaluation methods are needed that fit with these development cycles. Besides that, one of the promises of technology is to develop tailored and personalized interventions. In addition, efficiency (time and effort used in relation to the results achieved), acceptability (perceived usefulness, ease of use) and satisfaction (the extent to which the users' physical, cognitive and emotional responses that result from the use of a system meet the users' needs and expectations) need to be considered. Technologies will only result in effects when they are being actually used by end users. We therefore need insight into whether and how users use the system, how well the systems fit into the daily lives and context of the end users and the users' responses to use of the system. What aspects of the system do participants find most helpful or frustrating? How do different components of the system work together? What things do participants wish the system could do? What problems do participants face? Why do participants decline to participate? Why do participants (not) remain engaged over time (ISO 9241-11, 2018; Klasnja et al., 2011)? Therefore, next to quantitative methods, insights from qualitative methods are needed. Within this thesis, we will compare different qualitative methods to evaluate persuasive technology.

Challenge: effectiveness

Are persuasive technologies indeed a powerful medium for delivering interventions at work? What are the effects shown and intended, and how do they affect us? Evaluation is essential, not only to estimate the magnitude of their outcomes, but also to ensure they do no harm (Pagoto and Bennett, 2013). Persuasive technologies are being developed and evaluated in a variety of domains such as physical activity (Anderson et al., 2007; Consolvo et al., 2006; Consolvo et al., 2008; Lin et al., 2006), obesity (Bexelius et al., 2010; Patrick et al., 2009) and stress management (Koldijk et al., 2016; Plarre et al., 2011). However, scientific evidence is still limited (Fanning et al., 2012; Free et al., 2013). For persuasive technology specifically aimed at the occupational context, only a few studies were found, for example a tailored intervention on physical activity, snacking behaviour and sleep among airline pilots (Van Drongelen et al., 2014) and a review on persuasive technologies to reduce prolonged sedentary behaviour at work, which included eight studies (Wang et al., 2018). The purpose of technologies is to help users achieve their goals. The questions are whether these technologies are in fact able to do that, which factors determine long-term use and how feedback features are influencing reaching these goals. Within this thesis, effectiveness of persuasive technology is being studied on behaviour change, health-related outcomes and performance.

Challenge: societal impact

How far can we permit technology to go in influencing our behaviour, who is actually profiting from the collected data and is the user informed about all this (Kool et al., 2015)? In the context of work, self-monitoring technology has influenced the relationship between employer and employee. Technology has increased the possibilities for workers to take responsibility concerning their health and wellbeing at work (self-management/do it yourself mentality) and has given employers a more motivational and facilitating role. The employer might use aggregated data from individuals to better understand the emergence of symptoms and effects of interventions. This might lead to worker empowerment, but the continuous gathering of information also raises questions about safeguarding privacy and the responsible use of personal data – all the more so because emerging technologies often cross the boundaries between work and private life (e.g. measurements of sleep quality of workers). In addition, can the employee and the employer trust the (aggregated) data? To illustrate this: a well-known problem for applications is how to determine the cut-off point for a worker to fall into a certain health or safety risk category. This is decided by the algorithm. Minor variations in sensing might suddenly have large implications for the feedback given. Careless interpretation might worry users unnecessarily (false positives) or ease their minds when they should worry (false negatives). Furthermore, does the worker have a real choice to use or not to use applications that are offered by the employer? Do we need new regulations for all of this (Kool et al., 2015)? Developments in technology move fast, which means that now is the time to think about how we can guide the introduction of technologies for health and wellbeing at work in a responsible matter. Within this thesis, light is shed on what might be the societal impact of persuasive technology by researching a case-study of a digital stress-coach at work.

1.2 Aim of this thesis

This thesis explores the potential of persuasive technology in the context of health and wellbeing at work.

1.3 Outline

To explore the potential of persuasive technology for health and wellbeing at work, it is important to address the challenges described above: whether these technologies are theory based, how the effectiveness of these technologies can be assessed, whether these technologies are effective and what the societal impact of these technologies is. This leads to the following research questions.

Table 2 Research questions and corresponding chapters

Research questions	Chapter
1. Do persuasive technologies for health and wellbeing at work incorporate theory or evidence-based principles and constructs?	2
2. Which types of research methods are appropriate and useful to evaluate persuasive technologies for health and wellbeing at work?	3
3. What are the effects of persuasive technologies on health and wellbeing at work?	4, 5, 6
4. What is the societal impact of persuasive technology for health and wellbeing at work?	7

To gain insight into these questions, this thesis bundles the results of various studies in the broader context of persuasive technology for health and wellbeing in the work setting. The studies were embedded in research projects at TNO and supported by the Dutch National COMMIT programme – grant number Project07/SWELL, the Dutch Ministry of Economic Affairs, the Dutch Ministry of Social Affairs and Employment, Hoverstop BV, ABN AMRO and the Rathenau Institute.

Chapter 2: Behaviour change techniques in mHealth apps for the mental and physical health of employees: systematic assessment

The review in Chapter 2 provides an overview of behaviour change theories incorporated in mHealth applications for the mental and physical health of employees. In particular, this study evaluates which behaviour change techniques can be identified and which combinations of behaviour change techniques are present (De Korte et al., 2018a).

Chapter 3: evaluating an mHealth app for health and wellbeing at work: mixed-method qualitative study

Within this chapter, three different qualitative research methods are compared: interviews with workers, focus groups with workers and a focus group with experts. The objectives of this study were to gain insight into (1) the opinions and experiences of employees and experts on drivers and barriers using an mHealth app in the working context and (2) the added value of three different qualitative methods that are available to evaluate mHealth apps in a working context related to user satisfaction and technology acceptance (De Korte et al., 2018b).

Chapter 4: Effects of a feedback signal in a computer mouse: laboratory experiment

The Hoverstop Mouse (Hoverstop BV, Amsterdam) is a computer mouse that aims to change unnecessary, unfavourable postures of the lower arm and wrist that can cause sustained muscle tension. This mouse was used to study the effects of persuasive technology on behaviour, short-term health effects, performance and user friendliness. In a laboratory setting, 15 subjects participated in a comparative, experimental study with repeated measures (De Korte et al., 2008).

Chapter 5: Effects of a feedback signal in a computer mouse: short-term RCT in the field

The study in Chapter 5 sought to determine the effects of persuasive technology on behaviour, performance and usability with a short-term randomized controlled trial in the field. Again, the Hoverstop Mouse was used for this. This study particularly evaluated whether the results of the laboratory study, described in Chapter 4, would hold true in the field and to gain better insight into how users become accustomed to feedback, task effects and acceptability during the initial phase of working with the mouse (De Kraker et al., 2008)

Chapter 6: Effects of four types of non-obtrusive feedback: laboratory experiment

Building on the results of Chapters 4 and 5, the study in Chapter 6 investigated the effects of different types of feedback on behaviour, task performance and usability. The study was conducted with 24 subjects in a laboratory setting. Four types of feedback (two visual signals and two tactile signals) were compared with a no-feedback condition. The Hoverstop Mouse was also used for this study (De Korte et al., 2012).

Chapter 7: The digital stress coach. Total control over your mental health, or 'big brother is watching you'?

In Chapter 7, the impact of persuasive technology in the work context is described in a position paper using the example of a digital stress coach. The paper describes the changing relationship between employer and employee. In addition, these developments are measured against the current acts and guidelines on working conditions, privacy and personal data protection. Finally, policy implications and recommendations are given (Van Lieshout et al., 2015).

Chapter 8: General discussion and conclusion

Chapter 8 concludes with a general discussion of the potential of persuasive technology for the health and wellbeing of workers based on the results of the research described in this thesis, including a reflection, recommendations for future research, implications for design and development and implications for practice.

1.4 Literature

Aarts E., & De Ruyter B. (2009) New research perspectives on ambient intelligence. *Journal of Ambient Intelligence and Smart Environments, 1*, 5-14.

Akamatsu, M., MacKenzie, I.S., & Hasbroucq, T. (1995). A comparison of tactile, auditory and visual feedback in a pointing task using a mouse-type device. *Ergonomics, 38,* 816-827.

Alavinia, S.M., Molenaar, D., & Burdorf, A. (2008). Productivity loss in the workforce: associations with health, work demands, and individual characteristics. *American Journal of Industrial Medicine, 52*(1), 49-56.

Anderson, I., Maitland, J., Sherwood, S., Barkhuus, L., Chalmers, M., Hall, M., Brown, B., & Muller, H. (2007). Shakra : tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile networks and applications, 12*(2-3), 185-199.

Baileyn D.E., & Kurland, N.B. (2002(). A review of telework research: findings, new directions, and lessons for the study of modern work. *Journal of Organizational Behavior, 23,* 383-400.

Bennie, J.A., Chau, J.Y., Van der Ploeg, H.P., Stamatakis, E., Do, A., & Bauman, A. (2013). The prevalence and correlates of sitting in European adults – a comparison of 32 Eurobarometer-participating countries. *International journal of behavioral nutrition and physical activity, 10,* 1-7.

Berry, L.L., Mirabito, A.M., & Baun, W. (2011). What's the hard return on employee wellness programs? *Harvard Business Review, 89*(3), 20-1.

Bexelius, C., Lof, M., Sandin, S., Lagerros, U.T., Forsum, E., & Litton, J.E .(2010). Measures of physical activity using cell phones: validation using criterion methods. *J Med Internet Res, 12*(1), e2.

Bhui, K.S., Dinos, S., Stansfeld, S.A., & White, P.D. (2012). A synthesis of the evidence for managing stress at work: a review of the reviews reporting on anxiety, depression and absenteeism. *Journal of Environmental and Public Health,* doi: 10.1155/2012/515874.

Calvo, R.A., Vella-Brodrick, D., Desmet, P., & Ryan, R.M. (2016). Editorial for "Positive computing: a new partnership between psychology, social sciences and technologists. *Psychology of Wellbeing 6,* 10. DOI: 10.1186/s13612-016-0047-1.

Cancelliere, C., Cassidy, J.D., Ammendolia, C., & Côté, P. (2011). Are workplace health promotion programs effective at improving presenteeism in workers ? A systematic review and best evidence synthesis of the literature. *BMC Public health, 11,* 395.

Chau, J.Y., Van der Ploeg, H.P., Van Uffelen, J.G.Z., Wong, J., Riphagen, I., Healy, G.N., Gilson, N.D., Dunstan, D.W., Bauman, A.E., Owen, N., & Brown, W.J. (2010). Are workplace interventions to reduce sitting effective? *A systematic review. Preventive Medicine; 51*, 352-356.

Chau, J.Y., Van der Ploeg, H.P., Dunn, S., Kurko, J., Bauman, A.E. (2011). A tool for measuring workers' sitting time by domain: the Workforce Sitting Questionnaire. *British Journal of Sports Medicine*, 45(15): 1216-1222. Doi: 10.1136/bjsports-2011-090214.

Choe, E.K., Lee, N.B., Lee, B., Pratt, W., & Kientz, J.A. (2014). Understanding Quantified-Selfers' practices in collecting and exploring personal data. *CHI 2014*, April 26- May 01 2014, Toronto, ON, Canada.

Church, T.S., Thomas, D.M., Tudor-Locke, C., Katzmarzyk, P.T., Earnest, C.P., Rodarte, R.Q., Martin, C.K., Blaire, S.N., & Bouchard, C. (2011). Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PLoS ONE, 6*(5), e19657. Doi: 10.1371/journal.pone.0019657.

Consolvo, S., Everitt, K.M., Smith, I., & Landay, J.A. (2006). Design requirements for technologies that encourage physical activity. In *Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems* (pp. 457-466). New York: ACM Press.

Consolvo, S., Klasnja, P., McDonald, D.W., Avrahami, D., Froehlich, J., LeGrand, L., Libby, R., Mosher, K., & Landay, J.A. (2008). Flowers or a robot army ?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th international conference on ubiquitous computing* (pp. 54-63). New York: ACM.

Cook, D.J., & Song, W. (2009). Ambient intelligence and wearable computing: sensors on the body, the home and beyond. *Journal of Ambient Intelligence and Smart Environments, 1*, 83-86.

Czaja, S.J., & Sharit, J. (2009). Preparing organizations and older workers for current and future employment: training and retraining issues. In S.J. Czaja, & J. Sharit (Eds), *Ageing and work: assessment and implications for the future* (pp. 259-278). Baltimore MD: John Hopkins University Press.

De Korte, E.M., De Kraker, H., Bongers, P.M., & Van Lingen, P. (2008). Effects of a feedback signal in a computer mouse on movement behavior, muscle load, productivity, comfort and user friendliness. *Ergonomics, 51,* 1757-1775.

De Korte, E.M., Huysmans, M.A., De Jong, A.M., Van de Ven, J.G.M., & Ruijsendaal, M. (2012). Effects of four types of non-obtrusive feedback on computer behavior, task performance and comfort. *Applied Ergonomics, 43,* 344-353.

De Korte, E., Wiezer, N., Bakhuys Roozeboom, M., Vink, P., Kraaij, W. (2018a). Behavior change techniques in mHealth apps for the mental and physical health of employees: systematic assessment. *JMIR Mhealth Uhealth*, 6(10): e167. Doi: 10.2196/mhealth.6363.

De Korte, E., Wiezer, N., Janssen, J., Vink, P., Kraaij, W. (2018b). Evaluating an mHealth app for health and wellbeing at work: mixed-method qualitative study. *JMIR Mhealth Uhealth,* 6(3): e72. Doi: 10.2196/mhealth.6335.

De Kraker, H., De Korte, E.M., Van Mil, F.L.K., Rijs, B.P., & Bongers, P.M. (2008). The effect of a feedback signal in a computer mouse on hovering behaviour, productivity, comfort and usability in a field study. *Ergonomics, 51,* 140-155.

Demerouti, E., Derks, D., Ten Brummelhuis, L.L., & Bakker, A. (2014). New Ways of Working: impact on working conditions, work-family balance, and wellbeing. In C. Korunka, & P. Hoonakker (Eds.), *The Impact of ICT on Quality of Working Life*. Dordrecht: Springer Science + Business Media Dordrecht. DOI 10.1007/978-94-017-8854-0_8.

Dennerlein, J.T., & Yang, M.C. (2001). Haptic force-feedback devices for the office computer: performance and musculoskeletal loading issues. *Human Factors, 43*(2), 278-286.

Dennison, L., Morrison, L., Conway, G., & Yardley, L. (2013). Opportunities and Challenges for Smartphone Applications in Supporting Health Behavior Change: *Qualitative Study. J Med Internet Res, 15*(4), e86. URL: http://www.jmir.org/2013/4/e86/

Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W.S., Wilson, J.R., & Van der Doelen, B. (2012). A strategy for human factors/ ergonomics : developing the discipline and profession. *Ergonomics, 55*(4), 377-395. DOI: 10.1080/00140139.2012.661087.

Eurostat Newsrelease (2012). *Active ageing in the EU.* Luxembourg: Publications Office of the European Union.

Eurofound (2011). *Living longer, working better - Europe's coming of age, Factsheets*. Luxembourg: Office for Official Publications of the European Communities.

Eurofound & EU-OSHA (2014). *Psychosocial risks in Europe: prevalence and strategies for precentions*. Luxembourg: Publications office of the European Union.

Eurofound and the International Labour Office (2017). *Working anytime, anywhere: The effects on the world of work*. Publications Office of the European Union, Luxembourg, and the International Labour Office, Geneva.

European Commission (2014). *Green paper on mobile Health (mHealth).* Brussels: European Commission, 10.4.2014.

Fanning, J.F., Mullen, S.P., & McAuley, E. (2012). Increasing physical activity with mobile devices: a meta-analysis. *Journal of Medical Internet Research, 14*(6), e161. DOI: 10.2196/jmir.2171.

Fogg, B.J. (2003). *Persuasive Technology: using computers to change what we think and do.* Morgan Kaufmann Publisher, San Fransisco.

Fogg, B.J., Cuellar, G., & Danielson, D. (2009). Motivating, influencing and persuading users: an introduction to captology. In J. Jacko, & A. Sears (Eds.), *The human-computer interaction handbook: fundamentals* (pp. 133-147). Boca Raton, Fl.: CRC Press, Taylor and Francis Group.

Free, C., Phillips, G., Galli, L., Watson, L., Felix, L., Edwards, P., Patel, V., & Haines, A. (2013). The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. *PLoS Medicine, 10*(1), 1-45.

Gerr, F., Marcus, M., & Monteilh, C. (2004). Epidemiology of musculoskeletal disorders among computer users: lesson learned from the role of posture and keyboard use. *J Electromyogr Kinesiology, 14*(1), 25-31.

Hagberg, M., Tornqvist, E.W., & Toomingas, A. (2002). Self-reported reduced productivity due to musculoskeletal symptoms: associations with workplace and invididual factors among white-collar cumputer users. *J. Occup. Rehabil., 12,* 151-162.

Hallal, P.C., Andersen, L.B., Bull, F.C., Guthold, R., Haskell, W., & Ekelund, U. (2012). Clobal physical activity levels: surveillance progress, pitfalls and prospects. *Lancet, 380,* 247-257.

Hamberg-van Reenen, H.H., Proper, K.I., & Van den Berg, M. (2012). Worksite mental health interventions: a systematic review of economic evaluations. *Occupational Environmental Medicine, 69*(11), 837-845.

Hooftman, W.E., Mars, G.M.J., Janssen, B., De Vroome, E.M.M., Janssen, B.J.M., Pleijers, A.J.S.F., Ramaekers, M.M.M.J. & Van den Bossche, S.N.J. (2019). *Nationale enquête arbeidsomstandigheden 2018 (NEA).* Leiden/Heerlen: TNO/CBS.

Hutchinson, A.D., Wilson, C. (2011). Improving nutrition and physical activity in the workplace: a meta-analysis of intervention studies. *Health Promotion International*, 27(2): 238-249. Doi: 10.1093/heapro/dar035.

Irastorza, X. (2019). *Third European Survey of Enterprises on New and Emerging Risks (Esener 3) – First Findings.* Bilbao, Spain: European Agency for Safety and Health, Prevention and Research Unit.

ISO 9241-11:2018 (2018). *Ergonomics of human-system interaction -- Part 11: Usability: Definitions and concepts.* International Organization for Standardization (ISO), 2nd Edition.

Karlsson, M.L., Björklund, C., & Jensen, I. (2010). The effects of psychosocial work factors on production loss and the mediating effect of employee health. *Journal of occupational and environmental medicine, 52*(3), 310-317.

Kaur, H., Williams, A.C., McDuff, D., Czerwinski, M., Teevan, J., Iqbal, S.T. (2020). Optimizing for happiness and productivity: modeling opportune moments for transitions and breaks at work. *CHI 2020*, April 25-30, Honolulu, HI, USA.

Kersten-Van Dijk, E., Westerink, J.H.D.M., Beute, F., Ijsselsteijn, W.A. (2017). Personal informatics, self-insight, and behavior change: a critical review of current literature. *Human Computer Interaction*, 0, 1-29. Doi: 10.1080/07370024.2016.1276456.

Klasnja, P., Consolvo, S., & Pratt, W. (2011). How to evaluate technologies for health behavior change in HCI research. *CHI 2011*, May 7-12, Vancouver, Canada.

Koldijk, S., Kraaij, W., & Neerincx, M.A. (2016). Deriving Requirements for Pervasive Well-Being Technology From Work Stress and Intervention Theory: Framework and Case Study. *JMIR Mhealth Uhealth, 4*(3), e79.

Kool, L., Timmer, J., & Van Est, R. (Eds.) (2015). *Sincere support, the rise of the e-coach*. Den Haag: Rathenau Instituut 2015.

Kraaij, W., Verberne, S., Koldijk, S., De Korte, E.M., Van Dantzig, S., Sappelli, M., Shoaib, M., Bosems, S., Achterkamp, R., Bonomi, A., Schavemaker, J., Hulsebosch, B., Wabeke, T., Vollenbroek-Hutten, M., Neerincx, M., Van Sinderen, M. (2019) Personalized support for well-being at work: an overview of the SWELL project. *User Model User-Adap Inter*. https://doi.org/10.1007/s11257-019-09238-3

Kumar, S., Nilsen, W.J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W.T., Shar, A., Spring, B., Spruijt-Metz, D., Hedeker, D., Honavar, V., Kravitz, R., Lefebre, R.C., Mohr, D.C., Murphy, S.A., Quinn, C., Shusterman, V., & Swendeman, D. (2013). Mobile health technology evaluation: the mhealth evidence workshop. *American Journal of Preventive Medicine, 45*(2), 228-236.

Laihonen, H., Jääskeläinen, A., Lönnqvist, A., & Ruostela, J. (2012). Measuring the productivity impacts of new ways of working. *Journal of Facilities Management, 10*(2), 102-113.

Lamontagne , A,D,, Keegel, T., Louie, A.M., Ostry, A., & Landsbergis, P.A. (2007). A systematic review of the job-stress intervention evaluation literature, 1990-2005. *Int J Occup Environ Health, 13,* 268-280.

Lee, I.-M., Shiroma, E.J., Lobelo, F., Puska, P., Blair, S.N., & Katzmarzyk, P.T. (2012). Effect of physcial inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet, 380*, 219-229.

Lin, M., Lane, N.D., Mohammod, M., Yang, X., Lu, H., Cardone, G., Ali, S., Doryab, A., Berke, E., Campbell, A.T., & Choudhury, T. (2012). BeWell+: Multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization. In *WH '12. Proceedings of the conference on Wireless Health October 23-25, San Diego, USA* (Article 10). New York, ACM.

Loyen, A., Van der Ploeg, H.P., Bauman, A., Brug, J., Lakerveld, J. (2016). European sitting championship: prevalence and correlates of self-reported sitting time in the 28 European Union member states. *Plos One* 11(3): e0149320 https://doi.org/10.1371/journal.pone.0149320

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). *The internet of things; mapping the value beyond the hype*. [S.l.]: McKinsey & Company.

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., & Marrs, A. (2013). *Disruptive technologie: Advances that will transform life, business, and the global economy*. [S.l.]: McKinsey Global Institute.

Martimo, K.-P., Shiri, R., Miranda, H., Ketola, R., Varonen, H., & Viikari-Juntura, E. (2009). Self-reported productivity loss among workers with upper extremity disorders. *Scandinavian Journal of Work, Environment & Health, 35*(4), 301-308.

Middelweerd, A., Mollee, J.S., Van der Wal, N., Brug, J., & Te Velde, S.J. (2014). Apps to promote physical activity among adults: a review and content analysis. *International Journal of Behavioral Nutrition and Physical Activity, 11,* 97.

Netten, M.P., Van der Doelen, L.H.M., & Goossens, R.H.M. (2013). Chair based measurements of sitting behavior, a field study of sitting postures and sitting time in office work. In V.G. Duffy (Ed.), *Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management Human Body Modeling and Ergonomics* (pp. 261-268). Berlin, Heidelberg: Springer-Verlag. LNCS 8026.

Niedhammer, I., Chastang, J.-F., Sultan-Taïeb, H., Vermeylen, G., & Parent-Thirion, A. (2012). Psychosocial work factors and sickness absence in 31 countries in Europe. *European Journal of public health, 23*(4), 622-628.

Nieuwenhuijsen, K., Bruinvels, D., & Frings-Dresen, M. (2010). Psychosocial work environment and stress-related disorders, a systematic review. *Occupational Medicine, 60,* 277-286. DOI: 10.1093/occmed/kqq081.

OECD (2014). *Ageing and employment policies: Netherlands 2014: working better with age*. Paris: Ageing and Employment Policies, OECD Publishing.
DOI: 10.1787/9789264208155-en.

Oinas-Kukkonen, H. (2013). A foundation for the study of behaviour change support systems. *Pers Ubiquit Comput*, 17: 1223-1235.

Orji, R., Moffatt, K. (2018). Persuasive technology for health and wellness: state-of-the-art and emerging trends. *Health Informatics Journal*, 24(1), 66-91. Doi: 10.1177/1460458216650979.

Pagoto, S., & Bennett, G.G. (2013). How behavioral science can advance digital health. *Transl Behav Med, 3,* 271-276. DOI: 10.1007/s13142-013-0234-z.

Patrick, K., Raab, F., Adams, M.A., Dillon, L., Zabinski, M., Rock, C.L., Griswold, W.G., & Norman, G.J. (2009). A text-message based intervention for weight-loss: randomized controlled trial. *J Med Internet Res, 11*(1), 1-9.

Plarre, K., Raij, A., Hossain, S.M., Ali, A.A., Nakajima, M., al'Absi, M., Ertin, E., Kamarck, T., Kumar, S., Scott, M., Siewiorek, D., Smailagic, A., & Wittmers, L.E. (2011). Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks* (pp. 97-108). New York: ACM.

Pot, F., Dhondt, S., De Korte, E.M., Oeij, P., Vaas, F. (2012). *Workplace innovation in the Netherlands*. Chapter 8. In: I. Houtman (Ed.), Work Life in The Netherlands (pp 173-190). TNO, Hoofddorp, The Netherlands.

Proper, K.I. & Hildebrandt VH. (2006). Physical activity among Dutch workers - Differences between occupations. *Preventive Medicine, 43,* 42-45.

Proper, K.I. ,Van den Heuvel, S.G., De Vroome, E.M., Hildebrandt, V.H., & Van der Beek, A.J. (2006). Dose-repsonse relation between physical activity and sick leave. *British journal of sports medicine, 40*(2), 173-178.

Richardson, K.M., & Rothstein, H.R. (2008). Effects of occupational stress management intervention programs: a meta-analysis. *Journal of occupational health psychology, 13*(1), 69-93.

Robertson, M.M., & Vink, P. (2012). Examining new ways of office work between the Netherlands and the USA. *Work, 41*, 5086-5090.

Robroek, S.J.W., Schuring, M., Croezen, .S, Stattin, M., & Burdorf, A. (2013). Poor health, unhealthy behaviors, and unfavorable work characteristics influence pathways of exit from paid employment among older workers in Europe: a four year follow-up study. *Scandinavian Journal of Work, Environment & Health, 39*(2), 125-133.

Robroek, S.J.W., Van den Berg, T.I.J., Plat, J.F., & Burdorf, A. (2011). The role of obesity and lifestyle behaviours in a productive workforce. *Occupational Environmental Medicine, 68,* 134-139.

Rongen, A., Robroek, S.J.W., Van Lenthe, F.J., & Burdorf, A. (2013). Workplace health promotion- a meta-analysis of effectiveness. *American Journal of Preventive medicine, 44*(4), 406-415.

Salvendy, G. (Ed.) (2012). *Handbook of human factors and ergonomics*. Hobroken, NJ: John Wiley and Sons.

Sappelli, M., Verberne, S., Kraaij, W. (2016). Adapting the interactive activation model for context recognition and identification. ACM Trans. Interact. Intell. Syst. 6, 3, Article 22, Doi: http://dx.doi.org/10.1145/2873067

Schwab, K. (2016). *The fourth industrial revolution*. Geneva: World Economic Forum.

Slijper, H.P., Richter, J.M., Smeets, J.B.J., & Frens, M.A. (2007). The effects of pause software on the temporal characteristics of computer use. *Ergonomics, 50*(2), 178-191.

Speklé, E.M., Hoozemans, M.J., Blatter, B.M., Heinrich, J., Van der Beek, A.J., Knol, D.L., Bongers, P.M., & Van Dieën, J.H. (2010). Effectiveness of a questionnaire based intervention programme on the prevalence of arm, shoulder and neck symptoms, risk factors and sick leave in computer workers: a cluster randomised controlled trial in an occupational setting. *BMC Musculoskelet Disord., 11*(99). doi: 10.1186/1471-2474-11-99.

Stacey, N., Ellwood, P., Bradbrook, S. (2017). *Key trends and drivers of change in information and communication technologies and work location. Foresight on new and emerging risks in OSH.* Working report EU-OSHA, Luxembourg: publications office of the European Union. Doi: 10.2802/807562.

Stacey, N., Ellwood, P., Bradbrook, S.l, Reynolds, J., Williams, H., Lye, D.(2018). *Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025.* European Risk Observatory Report, Luxembourg: publications office of the European Union. Doi:10.2802/515834.

Swan, M. (2012). Sensor Mania! The internet of things, wearable computing, objective metrics, and the Quantified Self 2.0 - Review. *Journal of Sensor and Actuator Networks, 1,* 217-253. doi: 10.3390/jsan1030217.

Van den Broek, E.L. (2017) *Monitoring technology: The 21st century's pursuit of well-being?* Report number: EU-OSHA Discussion paper, affiliation 21860. DOI: 10.13140/RG.2.2.26713.01123.

Van den Heuvel, S.G., De Looze, M.P., Hildebrandt, V.H., & Thé, K.H. (2003). Effects of software programs stimulating regular breaks and exercises on work-related neck and upper-limb disorders. *Scan J Work Environ Health, 29*(2), 106-116.

Van den Heuvel, S.G., Geuskens, G.A., Hooftman, W.E., Koppes, L.L.J., & Van den Bossche, S.N.J. (2010). Productivity loss at work; health related and work-related factors. *Journal of Occupational Rehabilitation, 20,* 331-339. DOI 10.007/s10926-009-9219-7.

Van den Heuvel, S.G., IJmker, S., Blatter, B.M., & De Korte, E.M. (2007). Loss of productivity due to neck/shoulder symptoms and hand/arm symptoms: results from the PROMO study. *Journal of Occupational Rehabilitation, 17*(3), 370-82.

Van den Heuvel, S.G., Bakhuys Roozeboom, M.C., Eekhout, I., Venema, A. (2018). *Management of psychosocial risks in workplaces – evidence from the second European survey of enterprises on new and emerging risks (Esener-2).* European Agency for Safety and Health, Luxembourg: Publication Office of the European Union. Doi: 10.2802/5030

Van der Klink, J.J.L., Blonk, R.W.B., Schene, A.H., & Van Dijk, F.J.H. (2001). The benefits of interventions for work-related stress. *American Journal of Public Health, 91*(2), 270-276.

Van der Ploeg, H.P., Chey, T., Korda, R.J., Banks, E., & Bauman, A. (2012). Sitting time and all-cause mortality risk in 222.497 Australian Adults. *Arch Intern Med, 172*(6), 494-500.

Van Drongelen, A., Boot, C.R.L., Hynek, H., Twisk, J.W.R., Smid, T., & Van der Beek, A.J. (2014). Evaluation of an mHealth intervention aiming to improve health-related behavior and sleep and reduce fatigue among airline pilots. *Scand J Work Environ Health, 40*(6), 557-568.

Van Lieshout, M., Wiezer, N., De Korte, E. (2015). The Digital Stress coach. Total control over your mental health, or 'Big Brother is watching you'? Chapter 6. In: Kool, L., Timmer, J., Van Est, R. (Eds.), *Sincere Support - The rise of the e-coach* (pp. 155-178) The Hague: Rathenau Instituut. ISBN/EAN: 978-90-77364-64-2

Wahlström, J. (2005). Ergonomics, musculoskeletal disorders and computer work. *Occup Med, 55*(3), 168-176.

Wang, Y., Wu, L., Lange, J-P., Fadhil, A., Reiterer, H. (2018). Persuasive technology in reducing prolonged sedentary behavior at work: a systematic review. *Smart Health* 7-8: 19-30. Doi: 10.1016/j.smhl.2018.05.002.

West, R., & Michie, S. (2016). *A guide to development and evaluation of digital behavior change interventions in healthcare*. London: Silverback Publishing.

Wieringa, D., Engbers, L.H., Van Empelen, P., Duijts, S., Hildebrandt, V.H., & Van Mechelen, W. (2013). What is actually measured in process evaluations for worksite health promotion programs: a systematic review. *BMC Public Health, 13,* 1190.

IJmker, S., Huysmans, M., Blatter, B.M., Van der Beek, A.J., Van Mechelen, W., & Bongers, P.M. (2007). Should office workers spend fewer hours at their computer? A systematic review of the literature. *Occupational and Environmental Medicine, 64,* 211-222.

Yardley, L., Choudhury, T., Patrick, K., Michie, S. (2016). Current issues and future directions for research into digital behaviour change interventions. *American Journal of Preventive Medicine* 51(5): 814-815. DOI: 10.1016/j.amepre.2016.07.019

2

Behavior change techniques in mHealth apps for the mental and physical health of employees: systematic assessment[1]

[1] Published in *JMIR Mhealth Uhealth 2018;6(10):e167*

Elsbeth de Korte[1,2], MSc
Noortje Wiezer[1], PhD
Maartje Bakhuys Roozeboom[1], PhD
Peter Vink[2], PhD
Wessel Kraaij[1,3], PhD

[1] Netherlands Organization for Applied Scientific Research, Leiden, Netherlands
[2] Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
[3] Leiden Institute of Advanced Computer Science, Faculty of Science, Leiden University, Leiden, Netherlands

Abstract

Background: Employees remain at risk of developing physical and mental health problems. To improve the lifestyle, health, and productivity many workplace interventions have been developed. However, not all of these interventions are effective. Mobile and wireless technology to support health behavior change (mobile health [mHealth] apps) is a promising, but relatively new domain for the occupational setting. Research on mHealth apps for the mental and physical health of employees is scarce. Interventions are more likely to be useful if they are rooted in health behavior change theory. Evaluating the presence of specific combinations of behavior change techniques (BCTs) in mHealth apps might be used as an indicator of potential quality and effectiveness.

Objective: The aim of this study was to assess whether mHealth apps for the mental and physical health of employees incorporate BCTs and, if so, which BCTs can be identified and which combinations of BCTs are present.

Methods: An assessment was made of apps aiming to reduce the risk of physical and psychosocial work demands and to promote a healthy lifestyle for employees. A systematic search was performed in iTunes and Google Play. Forty-five apps were screened and downloaded. BCTs were identified using a taxonomy applied in similar reviews. The mean and ranges were calculated.

Results: On average, the apps included 7 of the 26 BCTs (range 2-18). Techniques such as “provide feedback on performance,” “provide information about behavior-health link,” and “provide instruction” were used most frequently. Techniques that were used least were “relapse prevention,” “prompt self-talk,” “use follow-up prompts,” and “provide information about others’ approval.” “Stress management,” “prompt identification as a role model,” and “agree on behavioral contract” were not used by any of the apps. The combination “provide information about behavior-health link” with “prompt intention formation” was found in 7/45 (16%) apps. The combination “provide information about behavior-health link” with “provide information on consequences,” and “use follow-up prompts” was found in 2 (4%) apps. These combinations indicated potential effectiveness. The least potentially effective combination “provide feedback on performance” without “provide instruction” was found in 13 (29%) apps.

Conclusions: Apps for the occupational setting might be substantially improved to increase potential since results showed a limited presence of BCTs in general, limited use of potentially successful combinations of BCTs in apps, and use of potentially unsuccessful combinations of BCTs. Increasing knowledge on the effectiveness of BCTs in apps might be used to develop guidelines for app developers and selection criteria for companies and individuals. Also, this might contribute to decreasing the burden of work-related diseases. To achieve this, app developers, health behavior change professionals, experts on physical and mental health, and end-users should collaborate when developing apps for the working context.

Keywords: behavior change techniques; mHealth; mental health; physical health; lifestyle; workplace; app; employee; work.

2.1 Introduction

Despite increased awareness and growing efforts to develop measures to effectively manage work-related risk factors and promote workers' healthy behavior, employees are still at risk of developing physical and mental health problems [1,2]. This is caused by physical and psychosocial work demands and unhealthy lifestyle behaviors, such as low physical activity levels and sedentary behavior. This is often provoked by the way current work and working environments are arranged.

The development of new technologies has brought about many changes in the way people work, resulting in a shift away from occupations that require moderate-intensity physical activity to occupations that are composed of sitting [3,4]. Physical inactivity and sedentary behavior (defined as time spent sitting [4]) are associated with deleterious health effects such as cardiovascular diseases, cancer, type 2 diabetes, and obesity [5-8]. Research has shown that employees with low physical activity levels and sedentary behavior are less productive at work (presenteeism), have decreased workability, and take more sick days [9-12].

Furthermore, the number of employees working with computers has increased over the past decades [13]. Research shows a relationship between computer use and the development of musculoskeletal symptoms [13-15]. Static postures and repetitive movements, physical work demands that are associated with computer work, are related to presenteeism, decreased work ability, and sickness absence [16,17].

During the past decade organizations started to organize work flexibly [18]. Employees decide for themselves where, when, and with which (digital) tools they work. This brings advantages such as autonomy, remote collaboration, and increased possibilities for sharing information. However, there are also drawbacks, such as struggling with managing the inflow of information, interruptions and task switching, perceived pressure to respond quickly, decreased perceived social support, and a disturbed work-life balance [18]. High psychosocial work demands are associated with health complaints, sickness absence, decreased workability, and productivity loss [1,3,19-22].

Improved working conditions are needed to create a healthy and productive working population [10,16]. Besides that, the workplace is a fruitful setting for health promotion because of the presence of natural social networks, the possibility of reaching a large population, and the fact that people spend a great deal of their lifetime at work [9,23,24]. For these reasons, much effort has been put into the development and evaluation of interventions in the workplace setting. This includes selective activities to change the individuals' risks, attitudes, behavior, and awareness as well as comprehensive interventions such as workplace health promotion programs [1,9,25,26]. However, research shows that workplace interventions may be beneficial, but not all these interventions are useful, or their overall effects are small [1,9,24-32].

Research shows that workplace interventions are more effective when they involve evidence-based principles that (1) offer a variety of engagement modalities, (2) start with a needs assessment of participants, (3) offer higher intensity of contacts to keep participants actively involved, (4) are tailored to address participants' needs, (5) address multiple risk factors, (6) support self-management, (7) use incentives, (8) provide easy access and easy follow-up, (9) use social support, and (10) are grounded in behavior theory [9,24,28,31,33]. Mobile and wireless technology is a growing area in supporting health behavior change and might offer a promising approach as a workplace intervention since it could offer many of these elements [34-37]. Mobile health, also known as mHealth, covers medical and public health practice supported by mobile devices, such as mobile (smart) phones, personal digital assistants and other wireless devices. It also includes lifestyle and wellbeing apps that may connect to wearable sensors and personal guidance systems [38]. Various features make them good candidates for the delivery of interventions supporting health behavior change. First, as portable devices, they can continuously monitor the users' behavior using sensors. They offer the opportunity to bring behavioral interventions into important real-life and working contexts where people make decisions about their health and encounter barriers to behavior change. Second, they may provide cheaper, more convenient interventions. Third, the connectedness facilitates the sharing of data with health professionals or peers. Finally, the increasing ability to use sensors to infer context, such as user location, movement, emotion, and social engagement. This has raised the prospect of timely, tailored interventions for specific contexts [39-43]. As a result, these technologies support a participatory role by users, while enhancing their responsibility for their health and performance [38].

mHealth apps are being developed and evaluated to support behavior change of the general population in a variety of domains, such as physical activity [44-48], obesity [49], and stress management [50-52]. Even with the recent proliferation of apps, research evidence regarding their effectiveness is scarce [53]. The vast majority of commercial apps have not been evaluated using scientific methods, and these apps tend not to be grounded explicitly in theories of health behavior [54]. In recent years, mHealth apps have been developed to target the occupational setting [55-57], a context characterized by its specific barriers. Physical working contexts might put additional constraints on the use of mHealth apps, for instance when working in cleanrooms or high-security settings. Likewise, the organizational working context has specific focus points, such as the fit of an app with working schedules, embedding an app within prevention programs, and the role of management in implementation and adoption of an app. However, despite their potential, little research has been published on mHealth apps for employees. Only 1 study was found showing the positive effects of a tailored mHealth intervention on physical activity, snacking behavior, and sleep among airline pilots [58]. Insight is needed to determine whether mHealth apps are a powerful medium for delivering interventions in the workplace setting. Therefore, these apps need to be evaluated on (1) their potential to support healthy work behavior, (2) their consistency with evidence-based practices, and (3) their effectiveness in improving mental and physical health. The aim of this study is to examine the first step: to assess whether mHealth apps for employees use principles and constructs underlying the processes of behavior change to enhance their mental and physical health.

Research on internet interventions (electronic health) and mHealth shows that they are more likely to be useful if they are firmly rooted in health behavior change theory [34,36,40,59]. Understanding which behavior change techniques (BCTs) are implemented can illuminate mechanisms by which using an app might facilitate behavior change as well as the types of persons for whom a given app may work best [60]. Abraham and Michie [61] and Michie et al [62] suggested several BCTs common to many health behavior theories and developed several versions of a taxonomy to identify BCTs in a range of health promotion interventions [61,62]. The taxonomies have been used to identify techniques or combinations of techniques that might enhance effectiveness [36,40].

A large body of research has been published using the taxonomy in traditional health promotion interventions [40], but few have quantified the extent to which specific BCTs are included in apps. To date, studies have evaluated whether apps for physical activity [40,54,60,63] or apps for physical activity and diet [53] incorporate BCTs. The most frequently applied BCTs in traditional health promotion interventions are “goal-setting,” “prompt intention formation,” “provide feedback on performance,” “self-monitoring,” and “review of behavioral goals” [40,61,64]. Studies report inconclusive evidence regarding the number of BCTs that are associated with effectiveness. A systematic review by Webb et al [59] on Web-based interventions reported that interventions that include a larger number of BCTs, using a taxonomy adapted from Hardeman et al [65], are more likely to be effective. In contrast, another meta-analysis by Michie et al [64] using the Abraham and Michie’s taxonomy [61], suggested that the number of included BCTs is not associated with a larger effect. The study showed that interventions were most likely to be effective when “self-monitoring” was used as a technique, or when “self-monitoring” plus an additional self-regulation technique were used [64].

When interventions involve multiple BCTs, the effects might be additive, neutral (ie, cancel each other out), or amplified [66]. Accordingly, the inclusion of specific combinations of BCTs appears to be critical. Dusseldorp et al [66] used meta-analysis to conclude that specific combinations of BCTs increase the chances of achieving a change in health behavior, while other combinations decrease them. Specific combinations were more successful than average, and the strongest effects were found with motivation-enhancing BCTs. Most effective combinations were “provide information about behavior-health link” with “prompt intention formation” and “provide information about behavior-health link” with “provide information on consequences” and “use follow-up prompts.” Least effective were interventions using “provide feedback on performance” without “provide instruction.”

In summary, studies on traditional health promotion interventions show that not only the presence of BCTs, but also specific combinations of BCTs might explain intervention success. Up until now, none of the studies on the inclusion of BCTs in apps for physical activity and diet [36,40,54,60,63] have evaluated the presence of specific combinations of BCTs. Although this has not yet been confirmed in studies on mHealth in general, and specifically for the occupational setting, it can be suggested that certain combinations of BCTs also serve as an indicator for potential effectiveness in mHealth. This study aims to evaluate whether apps for the mental and physical health of employees incorporate BCTs and, if so, which ones can be identified, and which combinations are present.

2.2 Methods

2.2.1 Overview

A comparative assessment was made of apps aimed at reducing physical and psychosocial risks at work including stress prevention or coping with stress and to promote a healthy workstyle (ie, prevention of sedentary behavior or promotion of physical activity) for individual workers. Three independent reviewers undertook the assessment of the presence of BCTs and combinations of BCTs in apps: 1 scientist in ergonomics and human factors (EK), and 2 experts on mental health (NW, MBR).

2.3 Search strategy

Since app stores differ in their acceptance policy and therefore might offer different apps, the study sample was identified through systematic searches in 2 app stores: iTunes and Google Play. The algorithms within Google Play and iTunes work differently in how they classify and rank apps and make matches for specific keywords. For instance, the Google Play algorithm considers the keywords from the description of an app, and it will rank the app in the search results accordingly. The first results listed are the most relevant. In iTunes, the app description does not influence the app store algorithm in ranking the apps.

Between December 2014 and April 2015 apps were searched, screened, and downloaded. Search terms were based on Boolean logic and included combinations for domain (work, worksite, workplace, worker, employee), health (activity, health, lifestyle, stress, mental, physical, behavior, risk, sitting, posture, shift work, vitality, resilience, wellbeing), and intervention (coach, intervention, assistant, motivation, support, program). Searches were performed without using the app stores' categories.

2.3.1 Inclusion

To be included, apps had to meet the following criteria: (1) be work-related, (2) be aimed at stress prevention or coping and/or psychosocial risk reduction and/or physical risk reduction and/or prevention of sedentary behavior and/or promotion of physical activity, (3) be aimed at healthy adults, (4) provide individually tailored feedback, and (5) be English or Dutch. Apps that contained handbooks, product catalogues or Occupational Safety and Health incident reporting were excluded. Apps that focused on older adults, students or individuals with health problems (e.g. depression) were also excluded.

2.3.2 *Screening and assessment*

Figure 1 shows an overview of the selection and screening procedure.

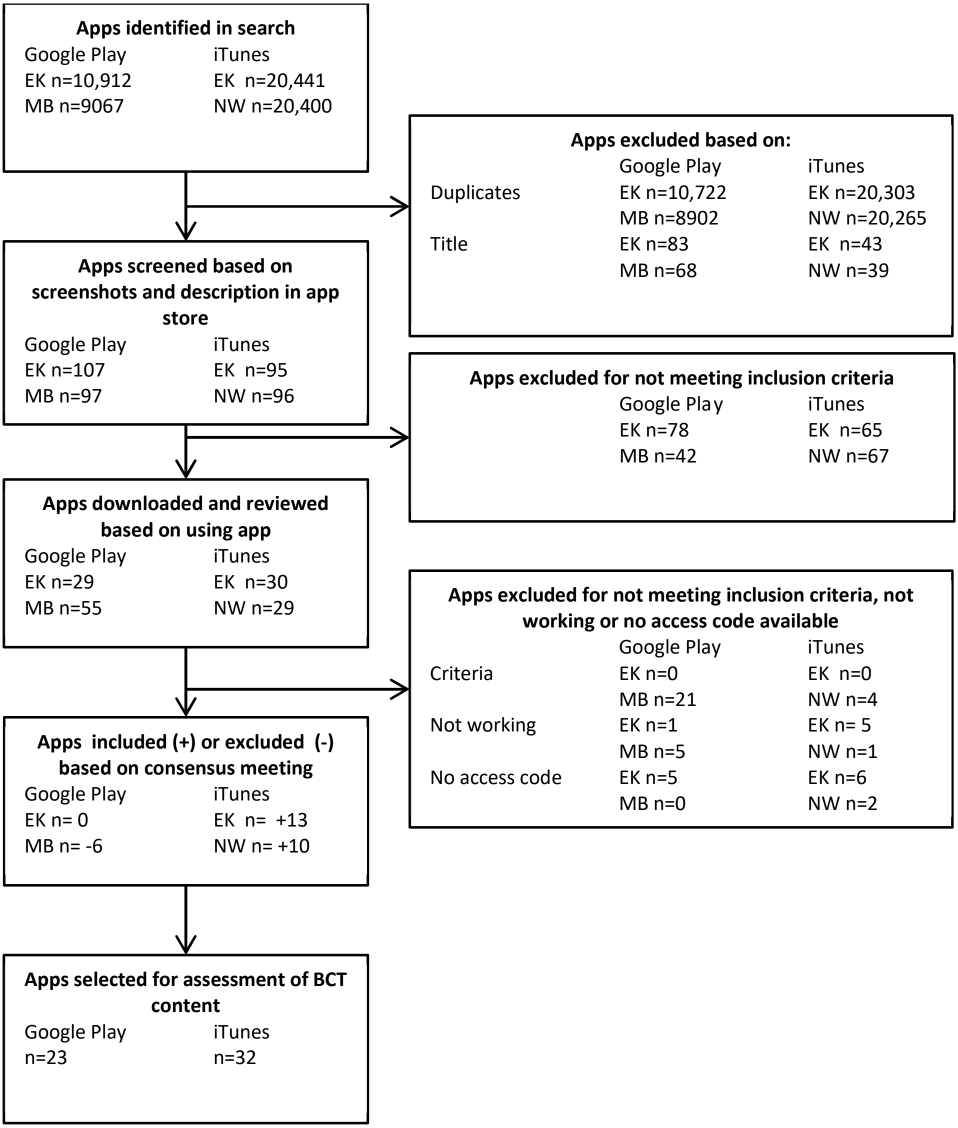

Figure 1 Overview of the selection and screening procedure of apps for assessment of behavior change techniques (BCTs).

Search, inclusion, screening of apps, and assessment of BCTs were performed by 3 researchers (EK and NW for iTunes; EK and MBR for Google Play). Any differences were resolved by discussion with the 2 reviewers and if necessary with the third reviewer.

First, search terms were entered in the app stores and apps were searched based on their title. Second, using the screenshots and description in the app store, the apps were screened using the inclusion criteria. Third, if an app seemed to be suitable for inclusion, it was downloaded to an iPhone 4 or a Samsung Galaxy S2. If there were doubts whether an app met the inclusion criteria, it was downloaded. If an app had a free version and a paid version, the free version was downloaded first to be reviewed. If the paid version contained additional features, it was downloaded and used for further analysis. Some apps required a unique access code. In this case, the app providers were contacted by email or phone to request a temporary access code or a demo version. While some app providers cooperated, others did not respond. These apps were not included for further analysis. Fourth, the downloaded apps were again assessed based on the inclusion criteria. Some apps appeared to be not working—these were excluded. In a consensus meeting, the final set of apps for assessment on BCTs was selected.

The reviewers used the included apps until they felt that they were familiar with the details. This varied from one hour (for very basic apps) to four weeks (for extensive apps or apps that took time before the user received feedback). The apps were assessed using the taxonomy of BCTs used in interventions, developed by Abraham and Michie [61]. This taxonomy consists of 26 BCTs and has been previously used to identify BCTs in apps [36,40]. For practical reasons we chose not to use the recent and comprehensive taxonomy by Michie et al with 93 BCTs [62]. This involved a high sensitivity of techniques which were considered too sensitive for the evaluation of apps. The taxonomy of the 26 BCTs formed the basis of the more elaborate taxonomy. In this approach, some of the BCTs used in the earlier taxonomy were specified into more detailed BCTs. To fully understand the content of the 26 BCTs, we studied the 93 BCTs. Before evaluation, all reviewers examined the coding manual and discussed each technique carefully, until a consensus was reached on definitions. Some definitions of BCTs from the taxonomy of Abraham and Michie [61] were adapted to be used for the assessment of apps (Multimedia Appendix 1). For each app, the researchers evaluated and provided a score if the 26 BCTs were present (1) or not (0). In addition to the BCTs, the researchers assessed whether the app was aimed at physical risk prevention, psychosocial risk prevention (including stress prevention or coping) or lifestyle promotion (prevention of sedentary behavior or promotion of physical activity). The apps were scored independently, and Krippendorff's alpha was used to evaluate interrater reliability since it can be used regardless of the number of observers, levels of measurement, sample sizes, and the presence or absence of missing data [67]. Also, the app name, a short description, the name of the app store, and the price for each app were collected, and stored in Excel for further analysis. The means and frequencies were calculated for the BCTs and the price of the app. Krippendorff alpha for nominal data was used to evaluate interrater reliability.

2.4 Results

2.4.1 General findings

The reviewers detected 10,912 (EK) and 9,067 (MBR) apps in Google Play and 20,441 (EK) and 20,400 (NW) in iTunes. The difference between Google Play and iTunes is because, for each search, Google Play generates a maximum of 250 results per search term, in contrast to iTunes, which has no maximum.

After the inclusion procedure, 45 apps were selected for the assessment of BCTs (see Table 1 for a general overview of the apps). Thirteen apps were found in Google Play, 22 in iTunes, and 10 were found in both app stores. Of the apps found in both stores, iChange2 and Wellmo were evaluated by NW and EK on an iPhone 4. The other 8 found in both stores were evaluated by MBR and EK on a Samsung Galaxy S2. Thirty-two apps were reviewed by NW and EK on an iPhone 4 while MBR and EK reviewed 23 apps on a Samsung Galaxy S2. In total, 45 different apps were evaluated.

Reliability data is shown in Table 1. Krippendorff alpha coefficients ranged from .23 to 1.00. Of the 45 reliability tests, 34 (76%) apps yielded alphas of at least .61 indicating good reliability. Fair reliability was found for 9 (29%) apps, which yielded alphas ranging from .41 to .60. Inferior reliability was assessed for 2 (4%) apps that scored below .41 [68]. Of the 45 apps, 13 (29%) had to be paid for with a mean price of €2.40 (range €0.99-4.99). Twenty-nine apps (64%) were free, and 3 (7%) apps had an access code. This access code was used when the app was offered as part of a company program. These apps are not free; however, the cost of these apps is unknown.

Fifteen (33%) apps were targeted at physical risk prevention, 23 (51%) at psychosocial risk prevention (including stress prevention or coping with stress), and 34 (76%) at lifestyle promotion (prevention of sedentary behavior or promotion of physical activity). Twenty-three (51%) apps were directed at a minimum of two categories, and 22 (49%) at just 1.

2.4.2 Behavior Change Techniques

The average number of BCTs was 7 (range 2-18). Most BCTs were used in iChange2 (18) and Wellmo (16). Table 1 shows that the least BCTs were identified in Positive Me (2), Ergometer (3), Office health alarm clock (3), and Stress Check by AIIR consulting LLC (3).

Figure 2 shows the BCTs identified most frequently and which BCTs were not. All 45 apps “provided feedback on performance”. This was no surprise since it was one of the inclusion criteria. Other techniques that were used more often were “provide information about behavior-health link” in 37 (82%) apps and “provide instruction” in 32 (71%) apps. Techniques that were used least were “relapse prevention” found in 3 (7%) of the apps, “prompt self-talk” in 2 (4%) apps, “use follow-up prompts” in 2 apps, and “provide information about others approval” in 1 app. “Stress management,” “prompt identification as a role model,” and “agree on behavioral contract” were not used by any of the apps (Figure 3).

Finally, combinations of techniques were analyzed. The combination “provide information about behavior-health link” with “prompt intention formation” was found in 7 (16%) apps (Brightr, iChange2, Move More, Office Buzz, Wellmo, 48-hour stress relief and Office exercise & stretch). The combination “provide information about behavior-health link” with “provide information on consequences” and “use follow-up prompts” was found in 2 (4%) apps (iChange2 and Wellmo). These combinations were found to be the most effective in health behavior change in the meta-analysis by Dusseldorp et al [66] indicating potential effectiveness in mHealth apps. The least effective combination “provide feedback on performance” without “provide instruction,” according to the meta-analysis of Dusseldorp et al [66], was found in 13 (29%) apps (Break Reminder, Darma, Fitlab, iSteplog, My Wellbeing App: Psycare Assist, Office Buzz, Office health alarm clock, Positive Me, Stand-up!, Standing desk companion [Varidesk], Stress Check [AIIR consulting LLC], Walk to Work and Workonit).

Table 1 Descriptive data of the apps that were evaluated for the presence of behaviour change techniques.

Name of the app	Krippendorff’s alpha	App Store purchased	Price (€)/ Code	Category of risk prevention or lifestyle promotion that apply to the apps			BCTs[1]
				Physical risk prevention	Psychosocial risk prevention	Lifestyle promotion	
1-minute desk workout	0.59	iTunes	0	Yes	-	Yes	8
48-hour stress relief	0.59	iTunes	4.99	-	Yes	-	8
Aetna Resources for Living	0.83	Google Play/ iTunes	0	-	Yes	-	8

Name of the app	Krippendorff's alpha	App Store purchased	Price (€)/ Code	Category of risk prevention or lifestyle promotion that apply to the apps			BCTs[1]
				Physical risk prevention	Psychosocial risk prevention	Lifestyle promotion	
Balance Coach Report Pro	0.63	iTunes	2.99	-	Yes	Yes	6
Break Reminder	1.00	Google Play	0	Yes	Yes	Yes	4
Brightr	0.95	Google Play/ iTunes	access code	-	Yes	Yes	10
Carecall	0.32	iTunes	0	-	Yes	Yes	6
Chair Yoga	0.84	iTunes	2.99	Yes	Yes	Yes	6
CNV mijn loopbaan app	0.79	iTunes	0	-	Yes	Yes	5
Darma	0.76	iTunes	0	-	-	Yes	4
Desk Workout	0.73	iTunes	0	Yes	-	Yes	7
Ergo@WSH	0.77	Google Play/ iTunes	0	Yes	-	Yes	7
ErgoCom	0.72	Google Play	0	Yes	-	-	4
Ergometer	0.63	Google Play	0	Yes	-	-	3
Ergonomics	0.86	iTunes	0	Yes	-	Yes	8
Fatigue Score Calculator	0.90	iTunes	1.29	-	-	Yes	5
Fitlab	0.79	Google Play	0	-	Yes	Yes	4
Get Off Your Butt!	0.91	iTunes	1.99	-	-	Yes	6
Happy@work	1.00	Google Play	3.99	-	Yes	-	5

Name of the app	Krippendorff's alpha	App Store purchased	Price (€)/ Code	Category of risk prevention or lifestyle promotion that apply to the apps			BCTs[1]
				Physical risk prevention	Psychosocial risk prevention	Lifestyle promotion	
Headspace.com meditation	0.79	Google Play	0	-	Yes	Yes	11
Ichange2	0.78	Google Play/ iTunes	Access code	-	Yes	Yes	18
iStepLog	0.63	iTunes	0	-	-	Yes	11
Ladies' Office Workout	0.55	Google Play/ iTunes	0	-	-	Yes	9
Measure Workplace Stress	0.64	Google Play/ iTunes	0	-	Yes	-	4
Minute Stretches	0.84	iTunes	0.99	Yes	-	Yes	7
Move More	0.62	iTunes	0.99	-	-	Yes	11
My Wellbeing App: Psycare Assist	0.23	Google Play/ iTunes	0	-	Yes	Yes	4
Office Buzz	0.63	Google Play	0	-	Yes	Yes	7
Office exercise & stretch	0.62	Google Play	1.18	-	-	Yes	9
Office health alarm clock	0.43	iTunes	0.99	Yes	-		3
Office Wellness	0.92	Google Play	0	Yes	-	Yes	8
Positive Me	0.43	iTunes	0	-	Yes	Yes	2
Provider resilience	0.65	iTunes	0	-	Yes	-	12
Salute the Desk	0.78	iTunes	3.99	Yes	-	Yes	9

Name of the app	Krippendorff's alpha	App Store purchased	Price (€)/ Code	Category of risk prevention or lifestyle promotion that apply to the apps			BCTs[1]
				Physical risk prevention	Psychosocial risk prevention	Lifestyle promotion	
Stand up!	0.63	Google Play/ iTunes	0	-	-	Yes	8
Standing desk companion (Varidesk)	0.59	iTunes	0	-	-	Yes	7
Stop Sitting virtual weight loss	0.75	iTunes	0.99	-	-	Yes	8
Stress Check (wisdomathand/ office harmony)	0.61	Google Play	0	-	Yes	-	4
Stress Check (AIIR consulting LLC)	0.61	iTunes	0	-	Yes	Yes	3
Stress Releaser Meditation	0.61	Google Play	3.82	-	Yes	-	5
VGZ Mindfulness Coach	0.51	Google Play	0	-	Yes	-	8
Voom	0.63	iTunes	0	Yes	-	Yes	11
Walk to Work	0.53	Google Play	0	-	-	Yes	6
Wellmo	0.65	Google Play/ iTunes	Access code	Yes	Yes	Yes	16
Workonit	0.50	Google Play/ iTunes	0	Yes	Yes	Yes	8

[1]BCT: behavior change technique

2

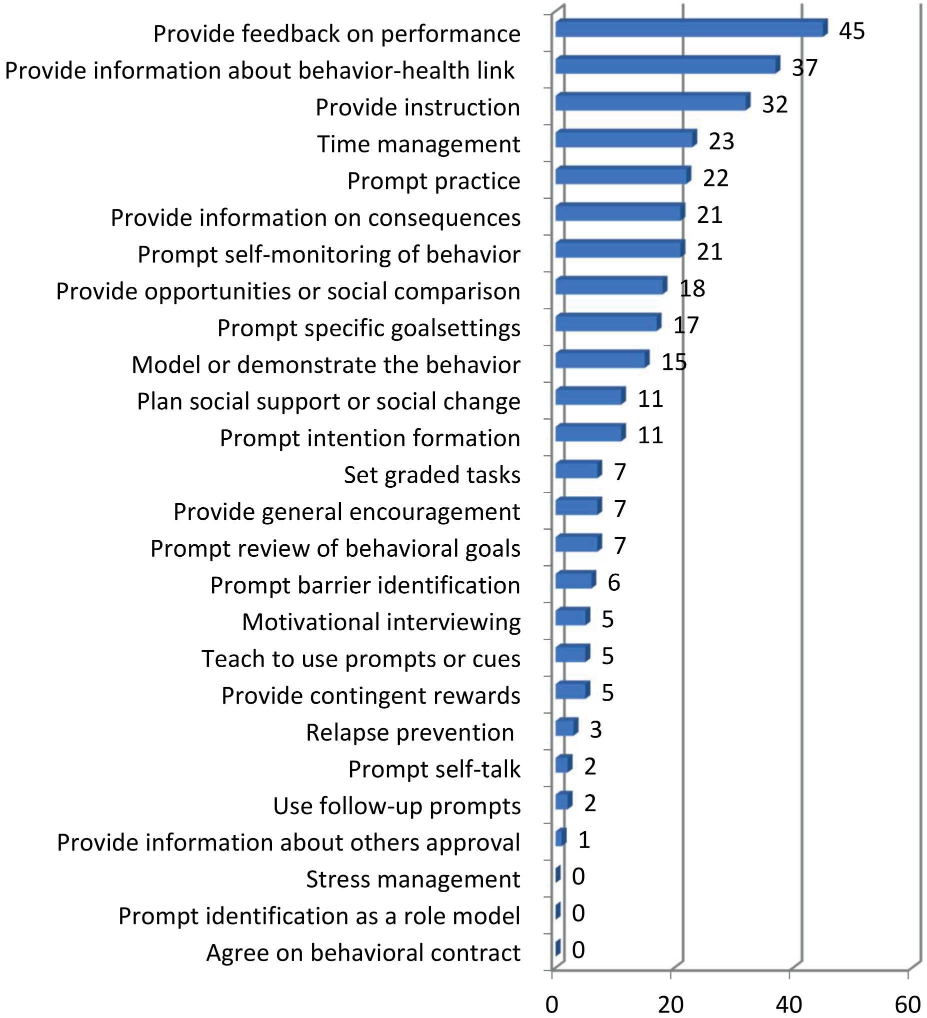

Figure 2 Frequencies of the behavior change techniques found in apps using the taxonomy by Abraham and Michie [61].

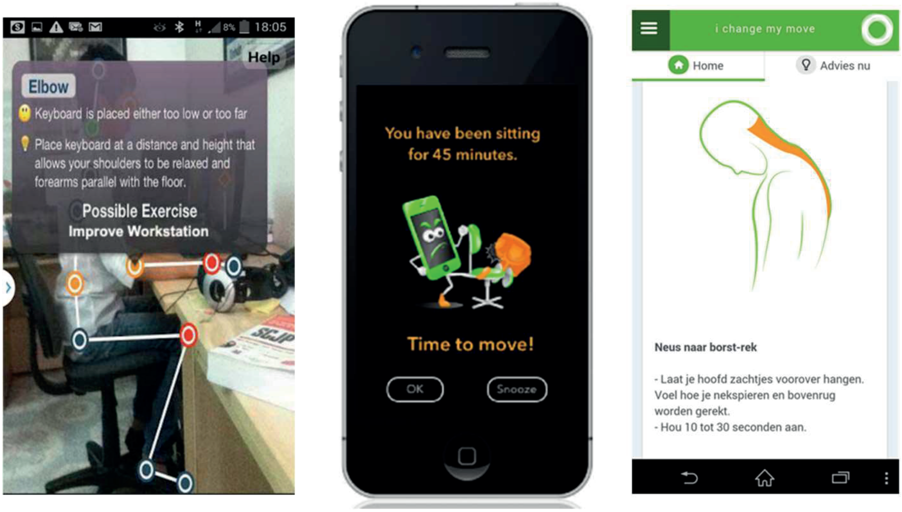

Figure 3 Examples of behavior change techniques used in apps, from left to right: “provide feedback on performance” (Ergo@WSH), “prompt practice” (Get Off Your Butt!) and “model or demonstrate behavior” (iChange2). Pictures have been taken from app descriptions in Google Play store (Ergo@WSH), iTunes (Get Off Your Butt!) and from app provider (iChange2).

2.5 Discussion

2.5.1 Principal findings

In this study, the presence of BCTs was identified in apps for the mental and physical health of employees. Previously, researchers have studied the presence of BCTs in apps, such as physical activity apps [36,40,54,60,63], dietary apps [53], medication adherence apps [69] or cancer survivorship apps [70]. Others have studied the presence of BCTs in wearable lifestyle activity trackers [71,72]. However, this study was the first to assess BCTs in apps aimed at improving the mental and physical health of employees. Also, this app assessment was the first to look at specific combinations of BCTs in apps, which might serve as an indicator of potential effectiveness.

The majority of the apps (34/45, 76%) in this study aimed to improve the health of employees targeted lifestyle promotion, while the number of apps directed at psychosocial risk prevention (23/45, 51%) and physical risk prevention (15/45, 33%) was much lower. About half (22/45, 49%) of the apps targeted just 1 of these categories. Reviewers noticed that lifestyle apps used sensors more often (eg, the accelerometer of the mobile phone used for step counting). In contrast, apps aiming at psychosocial risk prevention rarely used sensors to monitor; these apps generally used questions or questionnaires to gather data. One of the main advantages of mobile technology compared to the traditional nondigital interventions in the workplace setting is the ability to monitor the user's behavior with sensors continuously. This offers the opportunity to bring behavioral interventions into an important working context where people make decisions about their health and encounter barriers to behavior change. Differences in technical possibilities might influence the sort of apps that are being developed and the kind of behaviors they target.

The results of this study showed a limited presence of theoretical behavior change constructs. Previous research has highlighted the shortage of the application of behavior change theory in digital interventions, such as websites and apps designed to promote health behavior change [36,40]. Cowan et al [63] suggest that the general lack of theoretical constructs on behavior change included in apps might not be entirely unexpected, given that app developers' expertise relates to software development and may not include health behavior theory. Therefore, they might not thoroughly incorporate health behavior change theory into their apps [63]. Another explanation for these findings might be that, as per Dusseldorp et al [66], it is the type, quality, and combinations of BCTs, and how they are implemented, rather than the quantity of the techniques that matter. Finally, some techniques might not be detected by the researchers. The low Krippendorff alpha values found in some apps, as well as the discussions, emerged during the consensus meetings showed that reviewers did not always discover all features of the apps. Some features were not explicit during use. For example, reminders, updates, and feedback might have occurred for one reviewer, but not for another. Some BCTs were not easily traceable, for instance only via pop-up messages. This resulted in a different assessment of BCTs and might explain the interrater variability for some of the apps, with the lowest Krippendorff alphas belonging to Carecall and My Wellbeing App: Psycare Assist. However, it is important to note that low Krippendorff alphas might also exist in the case of rare values, especially with binary variables (ie, BCT present or BCT not present) with 1 rare value. Krippendorff alpha compares the "observed" and "expected" disagreements and to satisfy this it takes into account the prevalence of the categories coded for the variable. Nevertheless, one of the strengths of this study is that all apps have been screened and identified by at least two reviewers and in general, reasonable to good interrater reliability has been established.

In apps for mental and physical health, 7 BCTs were identified on average. Also, the number of applied BCTs showed a large variation between apps (range 2-18). These results are in line with those of Middelweerd et al [40] who found an average of 5 (range 2-8), Conroy et al [54] (average 4, range 1-13), Yang et al [60] (average 7, range 1-21), and Direito et al [36] (average 8, range (2-18), although these studies targeted physical activity and nutrition apps).

In this study, it was shown that the most common BCTs in apps for the health promotion of employees were “feedback on performance,” “providing information about the behavior-health link and provide instruction.” Middelweerd et al [40], Direito et al [36], and Conroy et al [54] also showed that “provide feedback on performance” and “provide instruction” were among the most identified BCTs. “Provide feedback on performance” was also found by Middelweerd et al [40] to be the most applied technique, although this was, similar to the current study, one of the inclusion criteria.

The current study showed that BCTs “relapse prevention,” “use follow-up prompts,” “prompt self-talk,” and “provide information about others’ approval” were identified the least. “Relapse prevention” and using “follow-up prompts” are important for sustained behavior change, but in the current study, these were applied in 3 apps only, which might question the value of these apps for changing behavior in the long-term [36]. However, it is unclear why these BCTs have been found in only a limited number in the sample of apps. For instance, these techniques might work well for interventions targeting addictive behaviors (eg, smoking) but might not be relevant for interventions promoting work style or habit formation.

“Stress management,” “prompt identification as a role model,” and “agree on behavioral contract” were not applied at all in any of the apps, which is in line with the work of Middelweerd et al [40] and Direito et al [36]. Further findings were not in line with the work of others: “prompt identification as a role model” was the fourth most applied technique in the study by Direito et al [36] but was not applied in that of Middelweerd et al [40], nor in the current study. “Prompt identification as a role model” was found by Direito et al [36] and there seems to be no technical obstacles to also applying “stress management” and “prompt identification as a role model” in apps. It appears that app developers might lack expertise in health behavior theory and therefore not include these techniques in their apps.

Compared to nondigital interventions in the workplace setting, one of the advantages of apps is the ability to monitor users’ behavior continuously and to deliver context-aware, personalized interventions. Consequently, these technologies support a participative role of users, while enhancing their responsibility for their health and performance [38,41-43]. For this reason, it was expected that many apps in the current study would have applied “prompt self-monitoring” in 21 (47%) apps, “plan social support or change” in 11 (24%) apps, and “prompt barrier identification” in 6 (13%) apps as a technique. The results did not quite confirm these expectations.

Applying certain combinations of BCTs is also essential. Dusseldorp et al [66] concluded from their meta-analysis that specific combinations of BCTs increase the likelihood of achieving change in health behavior, whereas other combinations decrease the possibility. The results of the current study showed that only a few apps applied most effective combinations and many apps applied the least effective. The meta-analyses by Dusseldorp et al [66] were performed with data on nondigital interventions. It is unclear whether this applies to digital interventions as well, but app developers should at least be conscious on how the number, the use, and combinations of BCTs might influence the effectiveness of an app. Therefore, future research should focus on the evaluation of which BCTs and combinations of BCTs are likely to be successful in effectively changing unhealthy behavior. Also, the present study shows that knowledge on effective BCTs might currently be underused in app development and suggests the need for multidisciplinary collaboration between app developers and behavior change experts. Others have concluded this as well [36,63,73]. Besides, to design tailored and targeted app-based interventions, insight into the preferences of the target population for certain BCTs is of importance. This has been shown by Belmon et al [74] for young adults in physical activity apps. Some BCTs were rated as more positive to apply than others. Ratings of BCTs differed according to personality traits and exercising self-efficacy. This may apply to apps for employees, and therefore, preferably employees should also be engaged in the development.

This study on BCTs in apps for the mental and physical health of employees had certain limitations. The procedure to search, identify, and review apps is susceptible to bias. Reviewers searched, screened, and downloaded apps on different days. Generally, apps are developed very fast and what is offered in app stores varies daily. This might have influenced the search results, especially those based on algorithm ranking (Google Play).

These fast developments also became apparent when some apps that were selected for download appeared to be untraceable. Presumably, many new apps have also appeared in the meantime. Still available apps have likely been changed, and new versions are available in the app stores since apps are updated continuously. This is illustrated by the study of Larsen et al [75] on the availability of mental health apps in iTunes and Google Play stores. They found 50% of search results changing within 4 months and an app being removed every 2.9 days. Therefore, conclusions on the apps that participated in the current study have to be interpreted with caution.

The taxonomy of Abraham and Michie [61] has not been developed specifically for apps. Therefore, reviewers had to translate the BCTs to app characteristics, which might have led to different interpretations than initially intended. For instance, stress management appeared to be a difficult BCT to interpret. It is defined as “may involve a variety of specific techniques (eg, progressive relaxation) that do not target the behavior but seek to reduce anxiety and stress.” However, in many apps in this study, management of stress was the targeted behavior, which was confusing. After a consensus meeting, it was decided to identify this technique only in cases where advice was given on ways to facilitate performance of the targeted behavior.

In addition to methodological limitations, there are also limitations in interpreting the results. As stated in the introduction, the extent to which apps are built upon theoretical models of the themes they address is essential (ie, stress management apps making use of evidence-based stress models). The current study focused on the presence of specific combinations of behavior change theories in apps. However, this is not necessarily an indication of good quality. Some of the apps in the study applied BCTs but also gave feedback that was not in line with current scientific insights. This raises the question of the value of these apps in supporting the user to enhance mental and physical health. Although an app might use principles and constructs underpinning the processes of behavior change, it also needs to be consistent with evidence-based practices. Therefore, designing useful apps requires the application of expertise from diverse fields and would benefit from interdisciplinary collaboration. While there is a consensus among software developers on the importance of engaging users, an mHealth app for employees would also benefit from collaboration with behavior change experts and experts in mental and physical health [76].

Moreover, the current study does not answer the question of whether apps are effective in changing behavior and thereby in the prevention of physical and mental health risk or promotion of a healthy lifestyle. To determine effectiveness, controlled trials are necessary, preferably using evaluation methods that fit with the fast, iterative development processes of apps (eg, a stepped wedge design) [35,37]. To date, the evidence base of apps is still scarce. Many apps are not based on solid evidence or evaluated with scientific methods [54,63,73].

Despite these limitations, this study provides the first analysis of health behavior theory applied in apps for the mental and physical health of employees. This research method cannot establish effectiveness and usability of these apps. Further research is needed to assess the effectiveness and usability of apps as intervention means for employees.

2.5.2 Conclusion

The findings of this study suggest that apps might be substantially improved to bring behavioral interventions into the working context where employees make decisions about their health and encounter barriers to behavior change. This study might be a first step toward implementing BCTs in a manner that is likely to increase behavior change potential.

The results, in general, showed a limited presence of BCTs, limited use of potentially successful combinations of BCTs in apps, and the use of potentially unsuccessful combinations of BCTs. Current knowledge on potentially effective combinations of BCTs seems to be underused in app development for the occupational setting. Knowledge of BCTs should be incorporated more in the development of apps. Combining behavior change theory and providing content with a robust evidence base and taking into account the specific context of the occupational setting could contribute to the development of effective mHealth-based interventions for employees and decrease the burden of work-related diseases. Although BCTs have been shown to be effective in face-to-face or online behavior change interventions, it is still unclear whether they are effective mHealth interventions. Future research should, therefore, focus on evaluating which BCTs and combinations of BCTs are effective in changing health behavior of employees when used in apps. For this evaluation, quantitative and qualitative methods should be used.

To increase potential and effectiveness, a collaboration between app developers, health behavior change professionals, experts on physical and mental health, and end-users is suggested. Combinations of expertise could provide higher quality apps. Until now, it is unclear which criteria could be used by organizations when selecting apps to offer to their employees. Furthermore, for employees, it remains unclear which app would help them best to improve their physical and mental health at work. An increase in knowledge on the effectiveness of BCTs in apps could be used to develop guidelines for app developers and the development of selection criteria for companies and individuals.

Acknowledgments

We would like to thank Ernest de Vroome, statistical expert at the Netherlands Organization for Applied Scientific Research, for reviewing the methods of this study.

Conflicts of interest

None declared.

Abbreviations

BCT: behavior change technique
mHealth: mobile health

2.6 References

1. Publications Office of the European Union, Luxembourg. 2014. Psychosocial risks in Europe: prevalence and strategies for prevention URL: https://www.eurofound.europa.eu/publications/report/2014/eu-member-states/working-conditions/ [accessed 2018-02-18] [WebCite Cache ID 6xJD5VVZi]
2. Eurofound. Fifth European Working Conditions Survey: Overview Report. Publications Office of the European Union, Luxembourg 2012. [doi: 10.2806/34660]
3. Church TS, Thomas DM, Tudor-Locke C, Katzmarzyk PT, Earnest CP, Rodarte RQ, et al. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. PLoS One 2011 May;6(5):e19657 [doi: 10.1371/journal.pone.0019657] [Medline: 21647427]
4. Hallal P, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012:380-257. [doi: 10.1016/S0140-6736(12)60646-1]
5. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012;380:219-229. [doi:10.1016/S0140-6736(12)61031-9]
6. Van der Ploeg HP, Chey T, Korda RJ, Banks E, Bauman A. Sitting time and all-cause mortality risk in 222 497 Australian adults. Arch Intern Med 2012 Mar 26;172(6):494-500. [doi: 10.1001/archinternmed.2011.2174] [Medline: 22450936]
7. Chau JY, Van der Ploeg HP, Van Uffelen JGZ, Wong J, Riphagen I, Healy GN, et al. Are workplace interventions to reduce sitting effective? A systematic review. Preventive Medicine 2010;51:352-356. [doi: 10.1016/j.ypmed.2010.08.012]
8. Bennie JA, Chau JY, Van der Ploeg HP, Stamatakis E, Do A, Bauman A. The prevalence and correlates of sitting in European adults? A comparison of 32 Eurobarometer-participating countries. International journal of behavioral nutrition and physical activity 2013;10:1-7. [doi: 10.1186/1479-5868-10-107]
9. Rongen A, Robroek SJW, Van Lenthe FJ, Burdorf A. Workplace health promotion: a meta-analysis of effectiveness. Am J Prev Med 2013 Apr;44(4):406-415. [doi: 10.1016/j.amepre.2012.12.007] [Medline: 23498108]
10. Robroek SJW, Van den Berg TIJ, Plat JF, Burdorf A. The role of obesity and lifestyle behaviours in a productive workforce. Occup Environ Med 2011 Feb;68(2):134-139. [doi: 10.1136/oem.2010.055962] [Medline: 20876556]
11. Robroek SJW, Schuring M, Croezen S, Stattin M, Burdorf A. Poor health, unhealthy behaviors, and unfavorable work characteristics influence pathways of exit from paid employment among older workers in Europe: a four-year follow-up study. Scandinavian Journal of Work, Environment & Health 2013;39(2):125-133. [doi: 10.5271/sjweh.3319]
12. Proper KI, Van den Heuvel SG, De Vroome EM, Hildebrandt VH, Van der Beek AJ. Dose-response relation between physical activity and sick leave. Br J Sports Med 2006 Feb;40(2):173-178 [doi: 10.1136/bjsm.2005.022327] [Medline: 16432007]
13. Wahlström J. Ergonomics, musculoskeletal disorders, and computer work. Occup Med 2005;55(3):168-176. [doi:10.1093/occmed/kqi083]
14. IJmker S, Huysmans M, Blatter BM, Van der Beek AJ, Van Mechelen W, Bongers PM. Should office workers spend fewer hours at their computer? A systematic review of the literature. Occupational and Environmental Medicine 2007;64:211-222. [doi: 10.1136/oem.2006.026468]

2

15. Gerr F, Marcus M, Monteilh C. Epidemiology of musculoskeletal disorders among computer users: lesson learned from the role of posture and keyboard use. J Electromyogr Kinesiology 2004;14(1):25-31. [doi: 10.1016/j.jelekin.2003.09.014]
16. Van den Heuvel SG, Geuskens GA, Hooftman WE, Koppes LLJ, Van den Bossche SNJ. Productivity loss at work; health related and work-related factors. Journal of Occupational Rehabilitation 2010;20:331-339. [doi: 10.1007/s10926-009-9219-7]
17. Martimo KP, Shiri R, Miranda H, Ketola R, Varonen H, Viikari-Juntura E. Self-reported productivity loss among workers with upper extremity disorders. Scand J Work Environ Health 2009 Jul;35(4):301-308 [Medline: 19471843]
18. Demerouti E, Derks D, Ten Brummelhuis LL, Bakker A. New Ways of Working: impact on working conditions, work-family balance,wellbeing. In: Korunka C, Hoonakker P, editors. The Impact of ICT on Quality of Working Life. Dordrecht Netherlands: Springer Science + Business Media; 2014.
19. Niedhammer I, Chastang J, Sultan-Taïeb H, Vermeylen G, Parent-Thirion A. Psychosocial work factors and sickness absence in 31 countries in Europe. Eur J Public Health 2013 Aug;23(4):622-629. [doi: 10.1093/eurpub/cks124] [Medline:23002241]
20. Karlsson ML, Björklund C, Jensen I. The effects of psychosocial work factors on production loss, and the mediating effect of employee health. J Occup Environ Med 2010 Mar;52(3):310-317. [doi: 10.1097/JOM.0b013e3181d1cda2] [Medline: 20190652]
21. Nieuwenhuijsen K, Bruinvels D, Frings-Dresen M. Psychosocial work environment and stress-related disorders, a systematic review. Occup Med 2010 Jun;60(4):277-286. [doi: 10.1093/occmed/kqq081] [Medline: 20511268]
22. Alavinia SM, Molenaar D, Burdorf A. Productivity loss in the workforce: associations with health, work demands, and individual characteristics. American Journal of Industrial Medicine 2008;52(1):49-56. [doi: 10.1002/ajim.20648]
23. Hutchinson AD, Wilson C. Improving nutrition and physical activity in the workplace: a meta-analysis of intervention studies. Health promotion international 2011;27(2):238-249. [doi: 10.1093/heapro/dar035]
24. Van der Klink JJL, Blonk RWB, Schene AH, Van Dijk FJH. The benefits of interventions for work-related stress. Am J Public Health 2001 Feb;91(2):270-276. [Medline: 11211637]
25. Wierenga D, Engbers LH, Van Empelen P, Duijts S, Hildebrandt VH, Van Mechelen W. What is actually measured in process evaluations for worksite health promotion programs: a systematic review. BMC Public Health 2013 Dec 17;13:1190 [doi: 10.1186/1471-2458-13-1190] [Medline: 24341605]
26. Chau JY, Van der Ploeg HP, Vn Uffelen JGZ, Wong J, Riphagen I, Healy GN, et al. Are workplace interventions to reduce sitting effective? A systematic review. Preventive Medicine 2010 Nov;51(5):352-356. [doi: 10.1016/j.ypmed.2010.08.012] [Medline: 20801153]
27. Lamontagne AD, Keegel T, Louie AM, Ostry A, Landsbergis PA. A Systematic Review of the Job-stress Intervention Evaluation Literature, 1990-2005. International Journal of Occupational and Environmental Health 2013 Jul 19;13(3):268-280. [doi: 10.1179/oeh.2007.13.3.268] [Medline: 17915541]
28. Richardson KM, Rothstein HR. Effects of occupational stress management intervention programs: a meta-analysis. J Occup Health Psychol 2008 Jan;13(1):69-93. [doi: 10.1037/1076-8998.13.1.69] [Medline: 18211170]
29. Bhui KS, Dinos S, Stansfeld SA, White PD. A synthesis of the evidence for managing stress at work: a review of the reviews reporting on anxiety, depression, and absenteeism. Journal of Environmental and Public Health 2012. [doi:10.1155/2012/515874]

30. Hamberg-van Reenen HH, Proper KI, Van den Berg M. Worksite mental health interventions: a systematic review of economic evaluations. Occup Environ Med 2012 Nov;69(11):837-845. [doi: 10.1136/oemed-2012-100668] [Medline: 22864248]
31. Cancelliere C, Cassidy JD, Ammendolia C, Côté P. Are workplace health promotion programs effective at improving presenteeism in workers? A systematic review and best evidence synthesis of the literature. BMC Public Health 2011 May 26;11(1). [doi: 10.1186/1471-2458-11-395] [Medline: 21615940]
32. Speklé EM, Hoozemans MJ, Blatter BM, Heinrich J, Van der Beek AJ, Knol DL, et al. Effectiveness of a questionnaire based intervention programme on the prevalence of arm, shoulder and neck symptoms, risk factors and sick leave in computer workers: a cluster randomised controlled trial in an occupational setting. BMC Musculoskelet Disord 2010 May 27;11(99):99 [doi: 10.1186/1471-2474-11-99] [Medline: 20507548]
33. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. Annual Rev Public Health 2008;29:303-323. [doi: 10.1146/annurev.publhealth.29.020907.090930]
34. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013 Jan;10(1):e1001362 [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
35. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
36. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [doi:10.1186/1471-2458-14-646] [Medline: 24965805]
37. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. 2011 May Presented at: CHI 2011; May 7-12; Vancouver, Canada.
38. European Commission. Green paper on mobile Health (mHealth). Brussels; 2014 Apr 10. URL: https://ec.europa.eu/ digital-single-market/en/news/green-paper-mobile-health-mhealth
39. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr;15(4):e86 [doi: 10.2196/jmir.2583] [Medline: 23598614]
40. Middelweerd A, Mollee JS, Van der Wal N, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11(97) [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
41. Swan M. Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. JSAN 2012 Nov 08;1(3):217-253. [doi: 10.3390/jsan1030217]
42. Choe EU, Lee NB, Lee B, Pratt W, Kientz JA. Understanding Quantified-Selfers practices in collecting and exploring personal data. In: Proceedings of the SIGHI Conference on Human Factors in Computer Systems. 2014 Presented at: CHI 2014; April 26 - May 1 2014; Toronto, ON. Canada p. 1143-1152. [doi: 10.1145/2556288.2557372]
43. Aarts E, De Ruyter B. New research perspectives on ambient intelligence. Journal of Ambient Intelligence and Smart Environments 2009;1(1):5-14. [doi: 10.3233/AIS-2009-0001]
44. Anderson I, Maitland J, Sherwood S, Barkhuus L, Chalmers M, Hall M, et al. Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. Mobile networks and applications 2007;12(2-3):185-199

45. Consolvo S, Everitt KM, Smith I, Landay JA. Design requirements for technologies that encourage physical activity. In: Proceedings of CHI 2006. 2006 Presented at: CHI 2006; April 22-27 2006; Montreal, Québec, Canada p. 457-466.
46. Consolvo S, Klasnja P, McDonald DW, Avrahami D, Froehlich J, LeGrand L, et al. Flowers or a robot army? Encouraging awareness and activity with personal, mobile displays. In: Proceedings of the 10th international conference on ubiquitous computing. 2008 Presented at: UbiComp 2008; September 21-24 2008; Seoul, Korea p. 54-63.
47. Lin J, Mamykina L, Lindtner S, Delajoux G, Strub HB. Fish'n'steps: encouraging physical activity with an interactive computer game. In: Proceedings of Ubicomp 2006. 2006 Presented at: 8th international conference on ubiquitous computing; 2006; Orange County, USA p. 261-278.
48. Bexelius C, Löf M, Sandin S, Trolle LY, Forsum E, Litton J. Measures of physical activity using cell phones: validation using criterion methods. J Med Internet Res 2010 Jan;12(1):e2 [doi: 10.2196/jmir.1298] [Medline: 20118036]
49. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009 Jan;11(1):e1 [doi: 10.2196/jmir.1100] [Medline: 19141433]
50. Plarre K, Raij A, Hossain S, Ali A, Nakajima M, al'Absi M, et al. Continous inference of psychological stress from sensory measurements collected in the natural environment. In: Proceedings of ACM/IEEE 10th Conference on Information Processing in Sensor Networks. 2011 Presented at: 10th Conference on Information Processing in Sensor Networks (IPSN); 12-14 April 2011; Chicago, IL, USA p. 97-108.
51. Koldijk S, Kraaij W, Neerincx MA. Deriving Requirements for Pervasive Well-Being Technology From Work Stress and Intervention Theory: Framework and Case Study. JMIR Mhealth Uhealth 2016 Jul 05;4(3):e79 [doi:10.2196/mhealth.5341] [Medline: 27380749]
52. Christmann CA, Hoffmann A, Bleser G. Stress Management Apps With Regard to Emotion-Focused Coping and Behavior Change Techniques: A Content Analysis. JMIR Mhealth Uhealth 2017 Feb 23;5(2):e22 [doi:10.2196/mhealth.6471] [Medline: 28232299]
53. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [doi:10.1186/1471-2458-14-646] [Medline: 24965805]
54. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
55. Nield, D. 2014. In corporate wellness programs, wearables take a step forward URL: http://fortune.com/2014/04/15/ in-corporate-wellness-programs-wearables-take-a-step-forward/ [accessed 2016-09-21] [WebCite Cache ID 6kgsSWeJe]
56. Gray, R. 2015. How to use wearable technology for health and wellbeing URL: http://www.hrmagazine.co.uk/article-details/ how-to-use-wearable-technology-for-health-and-wellbeing [accessed 2016-09-21] [WebCite Cache ID 6kgr4yNZ9]
57. Steigner G, Doarn CR, Schütte M, Matusiewicz D, Thielscher C. Health Applications for Corporate Health Management. Telemed J E Health 2017 Dec;23(5):448-452. [doi: 10.1089/tmj.2016.0162] [Medline: 27854179]
58. Van Drongelen A, Boot CRL, Hynek H, Twisk JWR, Smid T, Van der Beek AJ. Evaluation of an mHealth intervention aiming to improve health-related behavior and sleep and reduce fatigue among airline pilots. Scand J Work Environ Health 2014 Nov;40(6):557-568. [doi: 10.5271/sjweh.3447] [Medline: 25121620]

59. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [doi: 10.2196/jmir.1376] [Medline: 20164043]
60. Yang C, Maher JP, Conroy DE. Implementation of behavior change techniques in mobile applications for physical activity. Am J Prev Med 2015 Apr;48(4):452-455. [doi: 10.1016/j.amepre.2014.10.010] [Medline: 25576494]
61. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
62. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
63. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations? A content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
64. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
65. Hardeman W, Griffin S, Johnston M, Kinmonth A. Interventions to prevent weight gain: A systematic review of psychological models and behavior change methods. International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity 2000;24:131-143. [doi: 10.1038/sj.ijo.0801100]
66. Dusseldorp E, Van Genugten L, Van Buuren S, Verheijden MW, Van Empelen P. Combinations of techniques that effectively change health behavior: evidence from Meta-CART analysis. Health Psychol 2014 Dec;33(12):1530-1540. [doi:10.1037/hea0000018] [Medline: 24274802]
67. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Communication Methods and Measures 2007 Apr;1(1):77-89. [doi: 10.1080/19312450709336664]
68. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
69. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior Change Techniques in Apps for Medication Adherence: A Content Analysis. Am J Prev Med 2016 May;50(5):e143-e146. [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
70. Dahlke VD, Fair K, Hong YA, Beaudoin CE, Pulczinski J, Ory MG. Apps seeking theories: results of a study on the use of health behavior change theories in cancer survivorship mobile apps. JMIR Mhealth Uhealth 2015;3(1):e31 [doi: 10.2196/mhealth.3861] [Medline: 25830810]
71. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. JMIR Mhealth Uhealth 2016;4(2):e40 [doi: 10.2196/mhealth.4461] [Medline: 27122452]
72. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug;16(8):e192 [doi: 10.2196/jmir.3469] [Medline: 25131661]

73. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
74. Belmon LS, Middelweerd A, te Velde SJ, Brug J. Dutch Young Adults Ratings of Behavior Change Techniques Applied in Mobile Phone Apps to Promote Physical Activity: A Cross-Sectional Survey. JMIR mHealth uHealth 2015 Nov 12;3(4):e103. [doi: 10.2196/mhealth.4383]
75. Larsen ME, Nicholas J, Christensen H. Quantifying App Store Dynamics: Longitudinal Tracking of Mental Health Apps. JMIR Mhealth Uhealth 2016 Aug 09;4(3):e96 [doi: 10.2196/mhealth.6020] [Medline: 27507641]
76. Pagliari C. Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J Med Internet Res 2007 May 27;9(2):e15 [doi: 10.2196/jmir.9.2.e15] [Medline: 17537718]

2.7 Multimedia appendix 1. Definitions of Behavior Change Techniques

This is a Multimedia Appendix to a full manuscript published in the JMIR mHealth and uHealth. For full copyright and citation information see http://mhealth.jmir.org doi:10.2196/mhealth.6363

Below, definitions of Behavior Change Techniques, used to evaluate the apps in the current study. Definitions are based on the work of Abraham & Michie (2008) and adapted to be used for evaluation apps. Adaptions are shown in *italics*

Definitions of Behavior Change Techniques (BCTs)	
1. Provide information about behavior health link	General information about behavioral risk, for example, susceptibility to poor health outcomes or mortality risk in relation to the behavior.
2. Provide information on consequences	Information about the benefits and costs of action or inaction, focusing on what will happen if the person does or does not perform the behavior.
3. Provide information about others' approval	Information about what others think about the person's behavior and whether others will approve or disapprove of any proposed behavior change.
4. Prompt intention formation	Encouraging the person to decide to act or set a general goal, for example, to make a behavioral resolution such as "I will take more exercise next week".
5. Prompt barrier identification	Identify barriers to performing the behavior and plan ways of overcoming them.
6. Provide general encouragement	Praising or rewarding the person for effort or performance without this being contingent on specified behaviors or standards of performance.
7. Set graded tasks	Set easy tasks, and increase difficulty until target behavior is performed.
8. Provide instruction	Telling the person how to perform a behavior and/or preparatory behaviors.
9. Model/ demonstrate the behavior	An expert shows the person how to correctly perform a behavior, for example, in class or on video, *or by visualizations in the app*.
10. Prompt specific goal setting	Involves detailed planning of what the person will do, including a definition of the behavior specifying frequency, intensity, or duration and specification of at least one context, that is, where, when, how, or with whom.
11. Prompt review of behavioral goals	Review and/or reconsideration of previously set goals or intentions

2

Definitions of Behavior Change Techniques (BCTs)	
12. Prompt self-monitoring of behavior	The person is asked to keep a record of specified behavior(s) (e.g., in a diary). *An app that requires pressing a button at every intake would also count as an instance of self report. Multiple time-points.*
13. Provide feedback on performance	Providing data about recorded behavior or evaluating performance in relation to a set standard or others' performance, i.e., the person received feedback on their behavior. *Includes providing an overview of recorded behavior*
14. Provide contingent rewards	Praise, encouragement, or material rewards that are explicitly linked to the achievement of specified behaviors.
15. Teach to use prompts/ cues	Teach the person to identify environmental cues that can be used to remind them to perform a behavior, including times of day or elements of contexts.
16. Agree behavioral contract	Agreement (e.g., signing) of a contract specifying behavior to be performed so that there is a written record of the person's resolution witnessed by another.
17. Prompt practice	Prompt the person to rehearse and repeat the behavior or preparatory behaviors.
18. Use follow up prompts	Contacting the person again after the main part of the intervention is complete.
19. Provide opportunities for social comparison	Facilitate observation of non-expert others' performance for example, in a group class or using video or case study *or via app.*
20. Plan social support/ social change	Prompting consideration of how others could change their behavior to offer the person help or (instrumental) social support, including 'buddy' systems and/or providing social support.
21. Prompt identification as role model	Indicating how the person may be an example to others and influence their behavior or provide an opportunity for the person to set a good example.
22. Prompt self-talk	Encourage use of self-instruction and self-encouragement (aloud or silently) to support action.
23. Relapse prevention	Following initial change, help identify situations likely to result in readopting risk behaviors or failure to maintain new behaviors and help the person plan to avoid or manage these situations.
24. Stress management	May involve a variety of specific techniques (e.g. progressive relaxation) that do not target the

Definitions of Behavior Change Techniques (BCTs)	
	behavior but seek to reduce anxiety and stress.
25. Motivational interviewing	Prompting the person to provide self-motivating statements and evaluations of their own behavior to minimize resistance to change.
26. Time management	Helping the person make time for the behavior (e.g., to fit it into a daily schedule).

Reference

Abraham C & Michie S. (2008). A taxonomy of behavior change techniques used in interventions. Health Psychology; 27(3):379–87

3

Evaluating an mHealth app for health and wellbeing at work: mixed-method qualitative study[1]

[1] Published in *JMIR Mhealth 2018, 6*(3), e72

Elsbeth Marieke de Korte[1,2], MSc
Noortje Wiezer[1],PhD
Joris H Janssen[3], PhD
Peter Vink[2], PhD
Wessel Kraaij[4,5], PhD

[1] Netherlands Organisation for Applied Scientific Research, Leiden, Netherlands
[2] Faculty Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
[3] FocusCura, Amsterdam, Netherlands
[4] Netherlands Organisation for Applied Scientific Research, Den Haag, Netherlands
[5] Faculty of Science, Leiden Institute fo Advanced Computer Science, Leiden University, Leiden, Netherlands

Abstract

Background: To improve workers' health and wellbeing, workplace interventions have been developed, but utilization and reach are unsatisfactory, and effects are small. In recent years, new approaches such as mobile health (mHealth) apps are being developed, but the evidence base is poor. Research is needed to examine its potential and to assess when, where, and for whom mHealth is efficacious in the occupational setting. To develop interventions for workers that actually will be adopted, insight into user satisfaction and technology acceptance is necessary. For this purpose, various qualitative evaluation methods are available.

Objective: The objectives of this study were to gain insight into (1) the opinions and experiences of employees and experts on drivers and barriers using an mHealth app in the working context and (2) the added value of three different qualitative methods that are available to evaluate mHealth apps in a working context: interviews with employees, focus groups with employees, and a focus group with experts.

Methods: Employees of a high-tech company and experts were asked to use an mHealth app for at least 3 weeks before participating in a qualitative evaluation. Twenty-two employees participated in interviews, 15 employees participated in three focus groups, and 6 experts participated in one focus group. Two researchers independently coded, categorized, and analyzed all quotes yielded from these evaluation methods with a codebook using constructs from user satisfaction and technology acceptance theories.

Results: Interviewing employees yielded 785 quotes, focus groups with employees yielded 266 quotes, and the focus group with experts yielded 132 quotes. Overall, participants muted enthusiasm about the app. Combined results from the three evaluation methods showed drivers and barriers for technology, user characteristics, context, privacy, and autonomy. A comparison between the three qualitative methods showed that issues revealed by experts only slightly overlapped with those expressed by employees. In addition, it was seen that the type of evaluation yielded different results.

Conclusions: Findings from this study provide the following recommendations for organizations that are planning to provide mHealth apps to their workers and for developers of mHealth apps: (1) system performance influences adoption and adherence, (2) relevancy and benefits of the mHealth app should be clear to the user and should address users' characteristics, (3) app should take into account the work context, and (4) employees should be alerted to their right to privacy and use of personal data. Furthermore, a qualitative evaluation of mHealth apps in a work setting might benefit from combining more than one method. Factors to consider when selecting a qualitative research method are the design, development stage, and

implementation of the app; the working context in which it is being used; employees' mental models; practicability; resources; and skills required of experts and users.

Keywords: mHealth; work; qualitative research methods; interview; focus group; technology acceptance; user satisfaction; usability; wellbeing; prevention.

3.1 Introduction

3.1.1 Mobile health apps for health and wellbeing at work

Workers' health is of importance to the individual, as well as to the organization in which a person is employed. As healthy workers perform better, workplace interventions are being developed to improve performance, health, and wellbeing of workers [1-5]. However, research shows that interventions are often not effective, or overall effects are small [3-13]. This calls for exploring new approaches for health and wellbeing at work.

Mobile and wireless technology (mobile health, mHealth), defined as wireless devices and sensors, including mobile phones worn by persons during their daily activities, is a growing area in supporting health behavior change [14-20].

Various features make mHealth a good candidate for workplace interventions. For example, mobile technology offers the ability to continuously and unobtrusively monitor user's behavior. Thereby, these technologies can better assess the user's needs and preferences to deliver context-aware, personalized, adaptive, and anticipatory interventions. In addition, it offers the opportunity to bring interventions into situations where people make decisions about their health and encounter barriers to behavior change. It might also offer cheaper and more convenient interventions with a high penetration and a large reach. Finally, it can support a participative role of users, while enhancing their responsibility over their own health and performance [18-23]. On the other hand, problems have been reported as well, such as quickly declining engagement after usage onset of mHealth apps [24].

3.1.2 Evidence base for mobile health

Studies on *Web-based interventions* show that they can have positive effects on health knowledge and behavior (eg, [25,26]). These effects also have been shown for Web-based interventions aimed at workers' health (eg, [27]). However, scientific evidence of *mobile apps* (mHealth) is still limited [28,14].

mHealth apps are being developed and evaluated in a variety of domains such as physical activity (PA) [29-33], obesity [34], and stress management [35]. A lot of these apps have poor or zero evidence base and have not been evaluated with scientific methods [24,36,37]. In recent years, mHealth apps are being developed specifically aimed at risk prevention and healthy behavior in the work setting [38,39], but despite its potential, hardly any research has been published on the content and the effectiveness. Only one study on mobile apps targeting the working population was found, which showed positive effects of a tailored mHealth intervention on PA, snacking behavior, and sleep among airline pilots [40].

Evaluation of mHealth is important, not only to estimate the magnitude of their outcomes but also to ensure they do no harm. Research is not only lacking on health outcomes but also on whether apps actually increase adherence to the behaviors they target and whether apps perform better compared with traditional interventions, either as a stand-alone strategy or integrated within a program [24]. However, technologies can only be effective when they are actually being used by end users. To advance technology design, we therefore need insight into end users' real-life experiences. Hence, evaluation must involve more than effectiveness evaluation. Testing acceptability and satisfaction of end users plays an essential role as well; this is widely recognized as critical to the success of interactive health applications [17,41]. How is the system used by participants? How well does the system fit into daily (working) lives and context? Which aspects of the system do participants find most helpful or frustrating? How do different components of the system work together? What things do participants wish the system could do? What problems do participants face? Why do participants decline to participate? Why do participants (not) remain engaged over time? [17]. To answer such questions, qualitative methods are needed.

To sum up, despite its great promise, evidence is sparse for mHealth in general [15,17,24] and specifically for risk prevention and healthy behavior at work. Insight is needed whether mobile apps are indeed a powerful medium to deliver interventions at work, a context characterized with its own specific barriers. This is a major scientific knowledge gap and might hamper the adoption of mHealth by the working population. Research is needed to examine its potential and to assess when, where, and for whom mHealth is efficacious, specifically for the working context.

3.1.3 Evaluating mobile health

To study the potential of mHealth apps, quantitative as well as qualitative studies are needed. However, mHealth interventions challenge the way we conduct research. What types of evaluations are appropriate and useful for mHealth apps?

An important challenge is to ensure that an evaluation method matches with the development cycles of technology, which is characterized by a highly iterative process. For instance, to convincingly demonstrate that mHealth apps are effective in changing behavior, often large-scale, long-term studies with control groups such as randomized controlled trials are used [15,17,42-44]. However, in mHealth research, the time it takes to perform high-quality effectiveness studies is critical because technology may be obsolete before a trial is completed. The rapidly evolving nature of both mHealth apps and their uptake means that some components are continuously improved during a trial, though changes to an intervention during an evaluation pose a threat to internal validity [15,43,44].

In addition, it is a challenge to conduct research in an occupational setting [45]. Common examples of challenges are as follows: (1) the organization wants to target all employees with an intervention, although workers might have different needs and goals (eg, some workers suffer from sleeping problems and others need to better balance their work-private life balance); (2) organizations provide only few departments to participate in the research (which might question whether the results represent all employees); (3) the outcomes of interventions depend on the context in which they are delivered, which might be different within an organization (eg, employees performing office tasks or working at an assembly line); and (4) organizations prefer research among their employees to have minimal effect on the daily production processes [45]. The occupational context leads to additional constraints concerning the design of an mHealth intervention and additional constraints concerning the choice of methodologies.

The first step when evaluating novel technologies already starts at the earlier stages of development and consists of gaining a deep understanding of how and why a system is used (or not) [17]. Understanding how technology interacts with other important factors that affect behavior change, such as people's attitudes and preferences, their relationships, and the context in which they work and live, is critical for the development and adoption of apps [17,45-49].

The focus of this study was to gain insight in users' real-life experiences of mHealth apps in the working context and the added value of different qualitative methods that might be applied to assess this within this context.

Various qualitative evaluation methods to collect this information are available to apply in one or more stages of an iterative design process [17,46-50]. *Expert-based* methods are commonly used for reasons of practicability, because they are reported to be cheap, fast, and one does not have to recruit users [41,46,47,51]. However, results may not reflect mHealth app use in real practice, as the context in which experts use an app differs from the context of targeted workers. Commonly applied *user-based* methods to gain insight in end users' real life experiences are focus groups, interviews, surveys, and loggings [41,47,50,52,53]. Focus groups give a quick overview of users' opinions, and they give insights into the needs of the target group. Part of its value lies in the unexpected findings that can come from free-flowing discussion in the group [50,52,54]. Focus groups require less time burden for an organization than interviews, another frequently adopted method in mHealth evaluation studies [47]. Interviews can be useful to understanding perceptions, opinions, motivation, context of use, and behavior. Generally, compared with the focus group method, interviews take more time but provide deeper insight [54].

3.1.4 Aim

This study aims to:

- Gain insight in the opinions and experiences of employees and experts on drivers and barriers for using an mHealth app for health and wellbeing in the working context to develop recommendations for design and implementation
- Gain insight into the added value of different qualitative methods that might be applied within a working context through comparing three different qualitative evaluation methods and assessing whether they yield the same issues evaluating an mHealth app

For this purpose, an mHealth app specifically developed to improve health and wellbeing of workers at a high-tech company is used as a case study. Three different qualitative methods are used to gain insight in the opinions and experiences of employees and experts on drivers and barriers for using an mHealth app: (1) interviews with end users, (2) focus groups with end users, and (3) focus group with experts. Usability studies have shown that the types of issues revealed by end users' and experts' evaluations and by different evaluation strategies only slightly overlap [41,46,47]. Therefore, it is hypothesized that (1) issues revealed by end users' (employees) and experts' evaluations only slightly overlap and (2) issues revealed by end users' interviews and users' focus groups only slightly overlap. Issues are important topics or points, either neutral, positive, or negative, brought forward by the participants in this study on the use of the mHealth app.

3.2 Methods

3.2.1 Brightr, a mobile health app for health and wellbeing at work

For this study, the Brightr app (version 1.0, Sense Health) was evaluated (Figure 1). Brightr is an mHealth app especially developed for workers at a high tech company to improve their health and wellbeing. Brightr continuously monitors worker's behavior, with modules for mental resilience, sleep, PA, nutrition, and shift work. Brightr aims to provide tailored and personalized feedback at the time and place when it matters the most: it offers the possibility to set personal goals that are monitored by short questionnaires (ie, in the mental resilience module) and incorporated sensor data of the mobile phone (ie, to monitor PA and sleep). The collected raw data is then being transformed into real-time human and environmental behavior measurements. On the basis of intelligent algorithms, Brightr provides tailored feedback and advice. In addition, it is possible to compare individual performance with the organization's average.

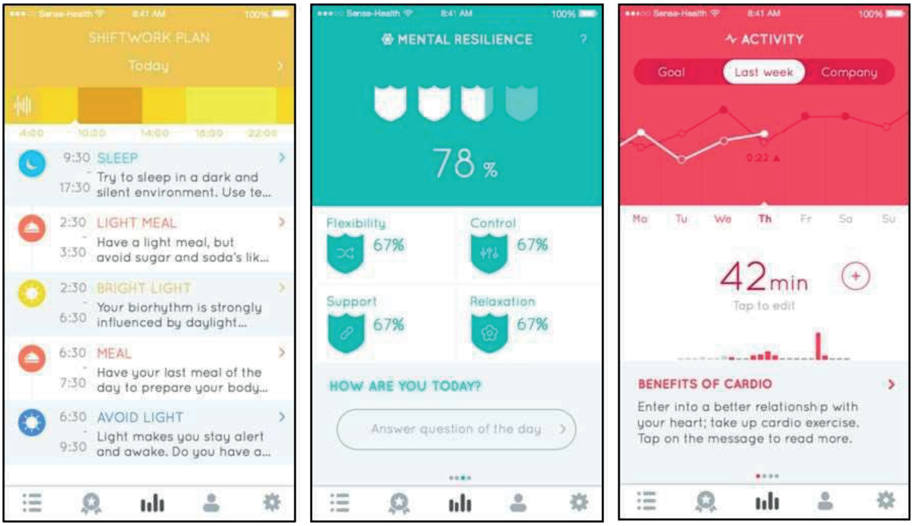

Figure 1 Brightr, examples of the shiftwork, mental resilience, and physical activity modules

3.2.2 *Qualitative evaluation methods*

This study included end user as well as expert evaluation methods. To get insight in users' real-life experiences with Brightr, three qualitative methods were used: interviews with end users, focus groups with end users, and a focus group with experts. These methods were applied as is customary in practice, and group sizes of each method were based on what was found in literature. It was planned to conduct between 20 and 25 interviews. In scientific literature, the guideline for the number of interviews is not clear. Some studies show that for an assessment of needs, 10 to 15 interviews will reveal about 80% of the needs [54]. Other studies advice to conduct interviews until saturation is reached and to stop when additional interviews will not yield new information [54,55]. Researchers advice to conduct between 6 and 200 interviews; most of them lie between 5 and 35 [55]. Therefore, aiming to conduct between 20 and 25 interviews was decided to be sufficient to get good results.

Semi standardized telephone interviews were conducted by two experienced interviewers (researcher EK with a background in human factors and ergonomics and researcher NW with a background in social sciences). They worked with an interview guide that contained a list of topics that should be addressed in every interview. After an introduction to the procedures, engagement questions on personal experiences with health and wellbeing interventions at work were asked. Second, exploration questions were asked on personal experiences with the use of general health and wellbeing apps and ideas on what kind of features an ideal app for health and wellbeing at work should have. Then, the Brightr app was evaluated using questions on general impression (eg, "What appeals to you, what not, and why?"), goal (eg, "Could you tell in your own words what the app aims to achieve?"), target group (eg, "For whom do you think this app was developed?"), potential (eg, "What would this app change for you?"), use (eg, "Do you (still) use the app and why (not)?"), outcome expectations (eg, "To what extent does this app fit your needs as a user?"), and information quality (eg, "What do you think about the amount of information to the users?"). The interview ended with general closing questions (eg, "Is there anything else you would like to say about Brightr?"). Before the start of the interview, participants signed an informed consent form. Interviews lasted up to 60 min. The interviews were transcribed verbatim and audio recorded to fix incomplete data during transcription.

The aim was to plan focus groups with a recommended size of 6 to 8 participants [54]. Three focus groups were conducted with end users (duration 90 min) at their company and one with experts (duration 120 min, at the research institute of EK and NW) by two experienced focus group facilitators: researchers EK and NW. Both researchers facilitated two focus groups and transcribed verbatim two times during the group discussions. The facilitator used a focus group guide that covered the same topics as the interview guide. Before the start of the focus group, participants signed an informed consent form. The focus group discussions were also audio recorded to fix incomplete data during transcription.

3.2.3 Participants

Brightr was offered to all employees of a high tech company, and they were able to download the app on a voluntary basis. Before recruitment for the evaluation study started, employees had the opportunity to use the app for at least 3 weeks. Employees were recruited for this study by a message on the company website and by messages on the information screens in the hallways that contained a link to the message on the company website. The message contained information on the aim, the setup, and data privacy of the study. To get insight in reasons for declining to use Brightr, employees were asked to follow a link in case they stopped using Brightr. This link directed to a questionnaire (Survalyzer) with two questions on the reasons for not using Brightr and on conditions or situations under which they would like to use an app such as Brightr. Employees using Brightr who were interested to participate in the study were asked to follow a link to another questionnaire (Survalyzer). It contained questions on gender, age, function group (operations and order fulfillment, sales and customer support, development and engineering, or support staff), hours working per week (flexible contract, 24 hours or less, 24-32 hours, or more than 32 hours), work experience at the company, and email address. This information was used to plan homogenous interview groups and focus groups. The email addresses were used to contact the participants to plan interviews and focus groups. Participants who declined an invitation for a focus group, for example, because the focus group was planned on an unfavorable timeslot for them, were asked to participate in an interview.

The experts were recruited by sending them an email with an invitation to participate in the study along with information about the aim and the setup of the study. They were asked to use the Brightr app for 3 weeks before they participated in a focus group. A total of 15 experts were recruited among the personal networks of two researchers (EK and NW) and consisted of behavioral scientists, psychologists, ergonomists, designers, human-computer interaction researchers, and policy makers. Upon acceptance of the invitation, experts received the Brightr app. To ensure a psychologically safe atmosphere, in which participants felt no barriers to speak freely, developers of the Brightr app (eg, researcher JJ) were excluded from the expert focus group.

3.2.4 Analysis

Qualitative data analysis was aimed to assess and compare issues addressed by end users in interviews and focus groups and by experts in a focus group. Data were collected from March 2015 to July 2015.

A codebook was constructed to analyze all transcripts. The codebook uses constructs from user satisfaction and technology acceptance models to understand and evaluate factors explaining users' perception about information systems to assess actual usage of these systems. Definitions used in the codebook of this study are adapted from the framework of Wixom and Todd [48], Bailey and Pearson [56], and Vosbergen et al [46] and specified further to the mHealth app that was used in this study. The final codebook can be found in Multimedia Appendix 1.

Data were categorized according to the following scheme: domain from the codebook, topic from the codebook, and whether the quote was positive, negative, neutral, or a recommendation, comparable to the analysis performed by Vosbergen et al [46]. In case a quote addressed multiple topics, it was categorized multiple times using different codes.

Two researchers (EK and NW) independently coded transcripts. After each transcript, they resolved discrepancies in discussion meetings up to the point they reached 80% matching codes, which was at the sixth transcript. The remaining transcripts were then evenly divided between researchers. Coded transcripts were included in Excel (Microsoft). Descriptive statistics were used to assess whether the three different qualitative analyses yielded the same issues evaluating Brightr and to gain insight in experiences and opinions that were obtained in general on drivers and barriers using Brightr in the working context.

3.3 Results

3.3.1 Nonparticipants

In the recruitment phase, 79 employees who declined to use Brightr filled in the two questions in Survalyzer on reasons for not using Brightr and conditions under which they would consider using an app such as Brightr. This group consisted of employees who never started using Brightr and employees who stopped using Brightr after a short period of time. How many employees never started to use Brightr is not known, nor is it known how long employees used Brightr before they stopped using it. This may have varied between just having a look at the app to using it for about 3 weeks. Figure 2 shows the main reasons of employees for not using Brightr. The most important reasons for not starting or quitting with Brightr were the large battery consumption of the app, not having a mobile phone, and the app had no relevance for the person. A total of 51 employees indicated that they would consider using Brightr under certain conditions. Most important conditions to consider using Brightr were improvements in battery use, clearer relevance for the user, and when the app would function on their mobile phone. A total of 28 employees would not consider using Brightr at all.

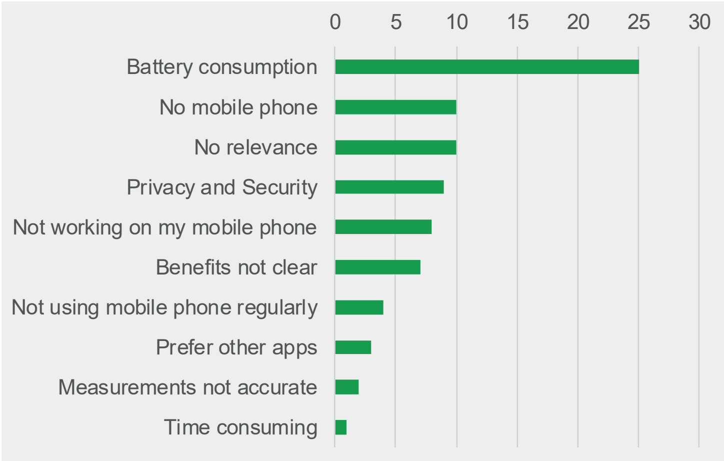

Figure 2 Main reasons (number of times mentioned) of 79 employees on why they declined to use Brightr and therefore, did not participate in the study

3.3.2 Participants

Reminders to participate in the study were sent twice via a pop-up message in the Brightr app to all users. After recruitment, 59 employees agreed to participate in the study. They received an invitation to plan an appointment for an interview or focus group. With 41 employees, an interview or focus group was planned. With 18 employees, it was not possible to plan an appointment because they did not respond to email messages or were absent from work because of sickness or vacation. Due to difficulties to recruit employees for the study, it was not possible to create homogeneous groups for interviews and focus groups.

With 22 employees, interviews were planned. The three focus groups with employees consisted of 4, 5, and 6 participants, respectively. Six more people were planned to participate in a focus group but declined, and 2 of them participated in an interview later on. Employee characteristics are shown in Table 1. Six experts (1 male, 5 female) participated in the focus group for experts. All participants obtained a university MSc and/or PhD in artificial intelligence, computer science, public administration, social sciences, or human movement sciences. They had expertise in the areas of behavior change, machine learning, big data and sensor data analysis, work-related stress, shiftwork, sustainable employability, electronic health or mHealth, mental resilience, PA, and intervention methods. All of the experts used Brightr for 3 weeks.

Table 1 Employee characteristics

Characteristics	Interviews	Focus groups
Number of employees, n	22	15
Years working at company, mean (SD)	6.6 (5.6)	10.4 (6.6)
Age (years), mean (SD)	39.0 (8.7)	45.2 (11.1)
Gender, n		
Male	17	13
Female	5	2
Function, n		
Operations and order fulfillment	7	5
Sales and Customer support	1	1
Development and engineering	9	5
Support function	5	4
Working hours, n		
Flexible or 0 hours	0	0
24 hours or less	1	0
24-32 hours	2	3
More than 32 hours	19	12

3.3.3 *Issues yielded with three qualitative methods*

Interviewing employees yielded 785 quotes, focus groups with employees yielded 266 quotes, and the focus group with experts yielded 132 quotes (Table 2).

Table 2 Number of participants in interviews and focus groups and number of quotes that were yielded with three different qualitative methods

Qualitative method characteristics	Interviews employees, n	Focus groups employees, n	Focus group experts, n
Number of participants	22	15	6
Number of quotes	785	266	132

3.3.4 *Overview of similarities and differences per domain*

Table 3 gives an overview of issues (neutral, positive, or negative) per domain. Interviews with employees yielded the highest percentage of issues within the domain of usefulness (25.5%, 200/785), followed by information quality (23.3%, 183/785). Focus groups with employees yielded also the most issues in the usefulness domain (27.4%, 73/266), which was followed by system quality (21.1%, 56/266). The focus group with experts yielded most issues in the system quality domain (23.5%, 31/132), followed by usefulness (22.7%, 30/132). In general, least issues were yielded on service quality.

Table 3 Overview of issues per domain (number and percentage)

Domain	Issues					
	Interviews employees		Focus groups employees		Focus group experts	
	n	%	n	%	n	%
System quality	98	12.5	56	21.1	31	23.5
Information quality	183	23.3	47	17.7	19	14.4
Service quality	8	1.0	3	1.1	0	0.0
Usefulness	200	25.5	73	27.4	30	22.7
Ease of use	48	6.1	8	3.0	11	8.3
Outcome expectations	126	16.1	39	14.7	17	12.9
Organizational factors	121	15.4	40	15.0	24	18.2

3.3.5 Overview of the value of issues per domain

Table 4 shows the number and percentage of positive, negative, or neutral issues and recommendations per domain.

Interviews yielded mostly recommendations within the domain of information quality and organizational factors. This method generated mainly negative issues in the domains of system quality, usefulness, ease of use, and outcome expectations. In contrast to both other methods, employee focus groups yielded mostly neutral issues within two domains: service quality and organizational factors. This method also generated mainly positive issues in the usefulness domain. Employee focus groups only yielded mostly recommendations in the domain of outcome expectations. This method generated mostly negative issues in the domains of system quality, information quality, and ease of use.

Experts gave mostly recommendations within the domains of system quality, outcome expectations, and organizational factors. No issues were yielded within the domain of service quality. In all other domains, experts mainly generated negative issues.

Table 4 Number and percentage of positive (+), negative (−), and neutral (0) issues or recommendations (R) within each domain

Domain and value	Issues within domain					
	Interviews employees		Focus groups employees		Focus group experts	
	n	%	n	%	n	%
System quality						
+[a]	11	11	2	4	0	0
–[b]	51	52	31	55	5	16
0[c]	11	11	9	16	8	26

Domain and value	Issues within domain					
	Interviews employees		Focus groups employees		Focus group experts	
	n	%	n	%	n	%
R[d]	25	26	14	25	18	58
Information quality						
+	43	23.5	6	13	6	32
–	59	32.2	24	51	10	53
0	10	5.5	1	2	0	0
R	71	38.8	16	34	3	16
Service quality						
+	3	38	1	33	0	0
–	2	25	0	0	0	0
0	0	0	2	67	0	0
R	3	38	0	0	0	0
Usefulness						
+	53	26.5	29	40	6	20
–	70	35.0	16	22	14	47
0	39	19.5	7	10	7	23
R	38	19.0	21	29	3	10
Ease of use						
+	20	42	0	0	0	0
–	21	44	6	75	7	64
0	0	0	2	25	2	18
R	7	15	0	0	2	18
Outcome expectations						
+	32	25.4	8	21	1	6
–	55	43.7	11	28	3	18
0	12	9.5	6	15	5	29
R	27	21.4	14	36	8	47
Organizational factors						
+	13	10.7	4	10	0	0
–	27	22.3	8	20	2	8
0	37	30.6	15	38	3	13
R	44	36.4	13	33	19	79

[a] + symbol signifies positive. [b] – symbol signifies negative. [c] 0 signifies neutral. [d] R signifies recommendations.

3.3.6 *Similarities and differences per topic*

Table 5 for each domain the underlying topics that were yielded by the employees (interviews and focus groups) and experts (focus group) are shown. An overview of illustrative In examples of quotes is shown in Multimedia Appendix 2.

System quality

Within the domain of system quality, issues of experts mostly focused on the topic "tailoring" (42%, 13/31), about 3 to 4 times as many as addressed by employees in respective interviews and focus groups (Table 5.). Experts especially stress the importance to tailor the app to the goals of the user and to personalize behavior change techniques, preferably using learning algorithms. Employees typically recommend tailoring the app to age, condition, and functioning type (ie, heavy work or desk work).

Employees mostly focused on the topic "performance of the system" (55/98, 56% and 21/56, 38% of the quotes in interviews and focus groups, respectively), whereas only 7% (2/7) of the quotes of experts were about this topic. Employees' quotes mainly focused on the high battery use; this was often a reason for quitting the use of Brightr. None of the experts made a quote on batteries. The system not working properly was another important issue on system performance for employees; it was either working too slow or having bugs.

In addition, in both focus groups, "time lines" was the second most addressed topic, 21% (12/98) and 23% (7/31) for employees and experts, respectively, almost twice as many as in the interviews (10%, 10/98). Issues on time lines mainly addressed the moments people use the app. An employee from the focus group stated:

> *When I receive a message, I take a look at the app. However, I take a look less often now, mostly in the evening or when I am at the toilet.* [Neutral quote]

Table 5 Topics of issues (number and percentage within domain)

	Issues within domain					
	Interviews with employees		Focus groups employees		Focus group experts	
Domain and topic	n	%	n	%	n	%
System quality						
Accessibility	0	0	0	0	0	0
Time lines (responsiveness)	10	10	12	21	7	23
Flexibility	10	10	5	9	5	16
Integration	5	5	8	14	2	7
Efficiency	5	5	1	2	0	0
Tailoring	13	13	6	11	13	42
Language	0	0	1	2	1	3
Errors or error prevention	0	0	2	4	1	3
Performance	55	56	21	38	2	7
Information quality						

Domain and topic	Issues within domain					
	Interviews with employees		Focus groups employees		Focus group experts	
	n	%	n	%	n	%
Accuracy	34	18.6	16	34	5	26
Precision	5	2.7	0	0	0	0
Reliability	1	0.5	4	9	1	5
Currency	5	2.7	2	4	0	0
Completeness	15	8.2	4	9	2	11
Format	45	24.6	4	9	3	16
Volume	9	4.9	1	2	2	11
Content	69	37.7	16	34	6	32
Visibility of system status	0	0.0	0	0	0	0
Service quality						
Relationship with app provider	2	25	0	0	0	0
Communication with app provider	2	25	2	67	0	0
Technical competence of app provider	1	13	0	0	0	0
Attitude of app provider	0	0	0	0	0	0
Schedule of products or services	2	25	0	0	0	0
Processing of change requests	0	0	0	0	0	0
Response time	0	0	0	0	0	0
Means of input with app provider	1	13	1	33	0	0
Usefulness						
Usefulness	14	7.0	18	25	4	13
Relevancy	110	55.0	42	58	13	43
Adherence	76	38.0	13	18	13	43
Ease of use						
User-friendly	21	44	3	38	1	9
Easy to use	8	17	0	0	0	0
Learnability	18	38	5	63	10	91
Memorability	1	2	0	0	0	0
Outcome expectations						
Expectations	30	23.8	5	13	0	0
Understanding of system	4	3.2	0	0	0	0
Confidence in the system	14	11.1	5	13	1	6
Feelings of participation	2	1.6	5	13	2	12
Feelings of control	23	18.3	11	28	6	35
Degree of training	0	0.0	0	0	1	6
Accuracy	12	9.5	4	10	00	0
Health and performance effects	41	32.5	9	23	7	41
Organizational factors						
Management involvement	6	5.0	4	10	5	21
Organizational competition	5	4.1	3	8	6	25
Security of data	39	32.2	15	38	7	29

Domain and topic	Issues within domain					
	Interviews with employees		Focus groups employees		Focus group experts	
	n	%	n	%	n	%
Documentation	0	0.0	4	10	3	13
Timing	22	18.2	6	15	0	0
Communication	49	40.5	8	20	3	13

Information quality

For the information quality domain, about one-third of the issues were yielded on the topic "content" of the app; this was similar for all types of methods. Employees mainly addressed the topics they would like to see in the app, for instance, food, sports, or work-rest schedules. An interviewed employee gave the following recommendation (see Multimedia Appendix 2):

> *I would like more information about food, what you should eat. Shift workers have to eat very fast at times (and therefore, the choices are not always healthy). I would like tips about food that is healthy and that you can eat fast.* [Recommendation]

Experts were mainly positive about the different aspects that were addressed by the app and gave recommendations on the content of the feedback:

> *I think of an app that shows the effects of your behavior, for example to show visually "what you have done now leads to this effect.* [Recommendation]

Next to the topic "content," interviews with employees yielded much issues on "format" (45/183, 24.6%; most employees liked the look and feel of the app). For both focus groups, "accuracy" of the app was an important topic (16/47, 34% with employees and 5/19, 26% with experts). Often, people doubted accuracy of the sleep measurements.

Service quality

Service quality was the least mentioned domain. Experts did not mention this domain and its topics at all. Interviews, as well as focus groups with employees, yielded the topics 'communication with the app provider and "means of input with app provider." In addition, in interviews, extra topics were addressed compared with the focus groups with employees: relationship with app provider, technical competence of app provider, and schedule of products and services.

Usefulness

Within the domain of usefulness, "relevancy" was the most addressed topic for each of the evaluation methods: 55.0% (110/200) of the quotes in employee interviews, 58% (42/73) in employee focus groups, and 43% (13/30) in expert focus groups.

All groups mainly focused on the extent to which the app or different aspects of the app helped to solve their problems (eg, sleep, stress, and healthy eating) or whether it addressed interests (eg, sports and food). An illustrative quote from an employee interview is as follows:

> *The best part, for me, is the shiftwork part (I work morning, evening, night shift). Since I try to follow the advices about maintaining a healthy lifestyle and working with shift hours. It helped me to keep down the stress in my body. I felt that I could focus better on the task during the daily (nightly) work.* [Positive quote]

An employee in a focus group stated:

> *The mental resilience part is doing absolutely nothing for me. I often think: for what reason am I doing this? If you are doing well, it has no added value*. [Negative quote]

An example of an expert quote is as follows:

> *Mental resilience also triggered...well, it yielded only frustration, I did not receive any tips*. [Negative quote]

For employees in the focus groups, "usefulness" was the second most addressed topic (25%, 18/73). One of the employees expressed:

> *It triggers to do things better in your behavior. The fact that I saw that I pretty quickly reached my physical activity goals was good, to see that it was not a problem for me*. [Positive quote]

For the other two groups, this was "adherence" (76/200, 38.0% for interviewed employees, 13/30, 43% for experts). Results showed that employees quit using the app mainly because of system failures, extensive battery use, or absence of relevancy, while push messages stimulate the use. Overall, many employees mentioned a decrease in use over time. Experts also mentioned system failures as a reason for attrition and stressed the importance of addressing user motivation.

Ease of use

Within the domain "ease of use," employees as well as experts experienced problems with discovering certain content and features of the app. One expert stated:

> *I found out only after a week that there was more than just physical activity. I swiped once accidentally and there they were: all sorts of modules!* [Neutral quote]

Interviewed employees focused mainly on the topic "user friendliness" (21/48, 44% of the issues, concerning positive as well as negative user experiences), followed by "learnability" (38%, 18/48). Results were opposite for both focus groups: the topic "learnability" was most important for employees and experts in focus groups (5/8, 63% and 10/11, 91%, respectively). In contrast with interviews, within both focus groups, the topics "ease of use" and "memorability" were not mentioned at all.

Outcome expectations

The topic "health and performance effects" was mentioned most often within the domain of outcome expectations; for interviewed employees and experts, it was the number one most addressed topic (41/126, 32.5% and 7/17, 41%, respectively), and for employees in focus groups, it was the second most addressed topic (23%, 9/39). The opinions on health and performance effects of interviewed employees were mixed: some declared that the app actually helped them to behave healthier, some think that an app such as Brightr is able to raise at least awareness, and others have doubts about the ability to change behavior or affect health. One of the interviewed employees stated:

> *There are many different kinds of workers in our company, some need physical activity advice (eg, they lift weights a lot at work), others have to exercise more (eg, sitting at desk to much). An app can help them to become aware. Goal of such an app is to try to get people think whether they are in balance. Do they have sufficient activity? I think it is possible that an app could help to reach goals. Reaching some goals must be possible*. [Positive quote]

Employees in focus groups showed similar opinions. Experts also showed mixed opinions about health and performance effects, but they focused more on different types of intervention functions apps might have, such as raise awareness, provide insight, give instruction, or change behavior and whether Brightr was able to do that (some agreed, some disagreed). For both focus groups, feelings of control appeared to be the second most important topic. Experts mainly stressed the importance of giving control to app users, for instance to set personal goals. They also discussed whether a user is able to decide for himself what he needs from a health perspective. Although employees also mentioned the significance of user autonomy, they were more focused on the possibilities to adjust missing data (eg, when they did not carry their mobile phone with them) or incorrect data (eg, app measuring walking instead of cycling).

Organizational factors

"Organizational factors" is an important domain to assess issues that influence uptake and implementation of mHealth apps in the working context. For the interviewed employees, "communication" was the most addressed topic within the domain of organizational factors (49/121, 40.5% of issues). It mainly addressed the way the app was implemented within their organization and how this was influenced by the relationship between employer and employee. Often they focused on whether management should play a role in implementation (management setting an example) or not (an organization should keep a certain distance when it comes to such personal data). For focus groups, security of data was most important; it was the second most addressed topic for the interviewed employees. Employees in interviews as well as in focus groups showed mixed opinions on data privacy and security. For some it is an important issue, for others it is not. Some employees mentioned that giving feedback to managers on an aggregated level might provide useful information for management. Experts mostly stressed the importance of being very concise and transparent on what happens to the data. "Management involvement" and "organizational competition" (congruence between assessment and feedback provided by the system and an external health professional or system [eg, coach, other app, and other system]) were least addressed by employees in interviews and focus groups, but gained much more attention by the experts. Experts mainly recommended organizations to embed an app such as Brightr in a bigger health or vitality program:

> *App should be a part of a bigger program, in terms of intervention. It is supportive within an intervention.* [Recommendation]

3.4 Discussion

3.4.1 Drivers and barriers using mobile health in the working context

The findings in this study suggest a number of valued characteristics, as well as challenges that organizations might consider for mHealth app and implementation and developers might use for design to enhance user satisfaction and technology acceptance. Overall, participants muted enthusiasm about the app. This is in line with the research of Dennison et al [20] who found similar results in their qualitative study on mobile phone apps supporting health behavior change among young adults. However, Dennison et al [20] found context sensing and social interaction features to be unnecessary and off-putting. This is in contrast with our study in which participants recommended to develop these features in future versions of the app, for example, the interest to compare personal data with organizational means or tailoring the personal advice to the shift work schedules. Apparently, to take context into account is very important for the application of mHealth in the working context but might be less important for other contexts of use. Combining results from the three evaluation methods that we used in our study, results show the following recommendations when designing mHealth apps for health and wellbeing at work.

Technology

System failures or poor performance (eg, high battery use) does influence adoption and adherence to mHealth apps negatively. Accuracy of measurements largely influences the confidence of users in the app and thereby influences its use. Accuracy (actual as well as perceived) but also the quality of the advice largely influences the possibility to reach behavior change and in line with that health and performance effects. It should therefore be based on solid evidence.

User characteristics

Relevancy and benefits of the app should be made clear to the (potential) user, within the app itself, as well as in communication guiding the implementation of the app. Furthermore, the app has to address users' characteristics (age, condition, health, function, and [work] activities), motivation, and needs (eg, health [risks] and wellbeing). A next step in developing apps should aim at using machine learning and learning algorithms to tailor the app to user characteristics automatically. A point of attention is giving users much autonomy, for instance, in ways to use the app, setting and adjusting goals, and when and how to receive feedback. Giving users autonomy in what they should need from a health point of view should be considered carefully, as users might not be aware of their health behaviors.

Context

It is very important to take into account the work context in which the app is being used. For instance, sometimes it is not possible to use a mobile phone in specific work contexts (eg, clean rooms), which affects the accuracy of the measurements. A suggestion might be to combine mobile phone apps with a wearable sensor that is possible to wear continuously in all (work) contexts. This suggestion is in line with Coursaris and Kim [57] who suggest to design interfaces and apps that fit particular contextual settings, while being flexible to accommodate others: "focus beyond the interface when designing applications" [57]. Furthermore, implementation plays a large role in the adoption and use of an app; this should thus be planned carefully, of which considering how and to what extent the management should be involved is an important factor. Experts suggest to embed such apps within a larger intervention to improve opportunities for success.

Privacy and autonomy

Results showed that for different end users privacy was either not an issue or an important issue. Van Lieshout et al [58] give some implications for dealing with apps that are offered by employers to their employees: an app offered by the employer always has to be used on a voluntary basis. Employees always should be alerted to their right to privacy and before apps are offered, and employees must be properly informed. In addition, within an organization, it should be very clear what happens to the data. Moreover, users should be given autonomy in deciding what happens to the data; various tools offer guidelines, for instance, Privacy by Design or Privacy Impact assessment [58,59].

3.4.2 Applying qualitative methods within a working context

Although studies have used qualitative evaluation methods in testing mobile apps [47,51,60-62] or compared qualitative evaluation methods in other apps, such as testing websites (eg, [46,63-67]), to our knowledge, this study was the first to assess whether different qualitative methods yield the same or different issues when testing an mHealth app for health and wellbeing at work.

The results of this study showed that issues revealed by experts only slightly overlapped with those expressed by employees. In addition, it was seen that interviews yielded different results compared with those from focus groups. These results are in line with conclusions from other studies comparing different qualitative evaluation methods: different methods identify unique issues, often more than common issues (eg, [47,51,63]).

Our study showed that the *type of evaluators* influences the kinds of issues an evaluation yields. The differences were seen in the attention that was given to the higher level of domains, as well as on the underlying topics that were addressed. For instance, the usefulness domain was given most attention by employees, whereas experts gave most attention to system quality. Moreover, differences were found in the values of remarks: positive, negative, neutral, or a recommendation. Although it was expected that experts would give many recommendations for improvement, they also yielded many negative remarks. Finally, analyzing the remarks itself, it was seen that even similar coded remarks were different in nature.

Employees gave insight into immediate practical experiences. The degree to which the app meets the needs of the employees and addressed their problems or interests is important for starting or continuing the use of the app. Furthermore, they described what motivated them to use the app, what prevented them from using it (such as system failures), and whether privacy of data played a role in using the app. This is in line with the findings of Vosbergen et al [46], but less in line with Lathan et al [67], who examined a Web-based system and found that users were mainly interested in efficient and effective use of the system. Results of this study are also in line with the work of Nielsen and Randall [68] on evaluating organizational-level interventions, who argue that insight into employee experiences is important to match an intervention to identified problems and to match it with the specific individual working context.

In this study, experts were more focused on higher level issues, building on their knowledge of theories and models, and using approaches derived from scientific knowledge and expertise. This is in line with Vosbergen et al [46] and Jaspers [41]. The experts in this study emphasized quality and evidence base of the information and ways to enhance adoption and continuous use by employees: accuracy of measurements, tailoring the app to user needs and providing users with autonomy (within certain boundaries), addressing user motivation, implementation of the app within the organization, and embedding in larger health or vitality programs. When implementing an app such as Brightr, they stressed that the intervention function of an app should be clear: raising awareness, providing insights, giving instructions, or changing behavior, as this influences the design of the app. Finally, they stressed the importance of transparency of data. According to Nielsen and Randall [68], who developed a model for evaluating organizational-level interventions, expert opinions are important as they are focused on the broader context of interventions and the use of theories. They understand the links between work and health and the underlying mechanisms, which is necessary to develop and implement effective interventions such as mHealth apps.

Tan et al [63] conclude that methods using experts or using end users complement each other and that neither method could be replaced by the other. They suggest using experts especially in the early design stages of development as they address user issues on a higher level, whereas user testing should be conducted in later stages as it needs a well-developed test bed. Vosbergen et al [46] concluded that an evaluation cannot be performed without end users, and the results of our study subscribe these conclusions. Vermeeren et al [51] and Adams and Cox [69] describe the importance of recruiting experts with required expertise, preferably with the right domain expertise.

Our study also revealed that the *type of evaluation* influences the kinds of issues an evaluation yields: issues addressed by employees in interviews differed from the issues addressed by employees in focus groups. This was seen in the attention that was given to certain domains, the values of the remarks within the domains, as well as the topics within each domain. Zapata et al [47] found four different evaluation methods in their systematic review that were used in mHealth evaluations: questionnaires, interviews, logs, and “think out loud” method. Questionnaires were the most applied method, followed by interviews. They did not find studies that used focus groups as an evaluation method for mHealth apps. This study shows that conducting focus groups for evaluating mHealth apps in the working context provides valuable information.

Often, a method is chosen on the basis of practicability [51,69]. Focus groups seem efficient because it gives a quick overview of opinions of multiple users at the same time [54]. Conducting interviews is a time-consuming process but offers the possibility of obtaining detailed and thorough information compared with, for instance, a questionnaire [69]. Some issues are for ethical and privacy reasons better dealt with in interviews, whereas a focus group will allow for easier reflection on common experiences [69]. This study did not confirm the idea that interviews lead to deeper insights or more detailed information as Van Boeijen et al [54] state; in this study, differences were found in the domains and underlying topics that were addressed, and results seemed of similar level of detail. Nor did this study confirm that ethical and privacy issues were better dealt with in interviews compared with focus groups [69]. In both settings, interviews and focus groups, employees in this study felt free to speak. From a practical point of view, our study showed that conducting focus groups is a more efficient qualitative method to evaluate an mHealth app than conducting interviews. Although both evaluation methods address overlapping issues, a focus group might offer more information on common or different experiences, for example, on factors such as (middle) management support, employee support, participation, information, and communication. In interviews, detailed individual experiences might have a more prominent role, such as the individual working conditions and individual factors such as readiness for change, perceptions, and appraisals.

3.4.3 Limitations

Several limitations of our study have to be discussed. Due to difficulties to recruit employees for the study, it was not possible to create homogeneous groups for interviews and focus groups. Results on the analysis of the questionnaire data of nonparticipants showed that our final group of employees probably has been biased. Within our sample of employees, individuals who were more motivated to respond (for instance, because they have strong opinions on the mHealth app) might have been overrepresented, as we used a self-selection protocol during recruitment.

Furthermore, although the total number of participants is larger than in most studies on user experiences, all three evaluator groups differed in size: 22 employees were interviewed, 15 employees participated in three focus groups, and 6 experts participated in one focus group. As a consequence, large differences between the number of remarks yielded by each method were found. To compare between methods, we therefore used percentages.

In addition, the three methods differed in the evaluation technique and the instructions that were given. These variations influenced results and made it difficult to examine the causes of the differences that were found between the three evaluation methods. However, the goal was to compare three different methods in the way they are commonly used in practice, not to compare them in an experimental setting with controlled variations. For this purpose, three methods were compared using one case study with the Brightr app to make a systematic comparison of methods; the study would have to be repeated in more settings.

Moreover, an early version of the Brightr app was used while conducting the study. On one hand, this might have skewed the responses to focus more on system quality and accuracy as compared with an app that has been developed further. On the other hand, this might have provided extra points of feedback that might otherwise not have been compared between qualitative methods.

Finally, a limitation of our study lies within the rating process. For time efficiency reasons, 2 raters independently coded remarks and resolved discrepancies in discussion meetings up to the point they reached 80% matching codes, at the sixth transcript. The remaining transcripts were then evenly divided between researchers. Although this procedure has been followed to reach a certain degree of reliability, no interrater reliability tests have been performed, and raters might have used different interpretations in rating the remaining transcripts.

3.4.4 Conclusions

Findings in this study provide the following recommendations for organizations planning to provide mHealth apps to their workers, as well as for developers of mHealth apps: (1) system performance influences adoption and adherence, (2) relevancy and benefits of the mHealth app should be clear to the user and should address users' characteristics, (3) app should take into account the work context, and (4) employees should be alerted to their right to privacy and use of personal data.

When considering which qualitative method to apply in a work setting, findings in this study showed that the type of evaluators as well as type of evaluation method influences which kinds of issues will be generated. The results revealed that different evaluation methods are complementary and therefore, evaluation processes might advantage from combining more than one method, which is also concluded by others [47,51,62-64]. Factors to consider when selecting methods for a qualitative evaluation of mHealth apps in the occupational setting are as follows: required information on the design and implementation of the mHealth app, the working contexts in which it is being used and participants' mental models on the mHealth app and context; the development stage of the app; practicability; resources; and skills required of experts and/ or users.

However, more scientific insight on these issues is still necessary. Furthermore, which methods work best in what situation and which methods work well together are still questions under research.

Acknowledgements

The authors would like to thank all the employees of the high-tech company and all the experts for taking part in this study. This research was supported by the Dutch National COMMIT program, grant number Project07/ SWELL, awarded to Wessel Kraaij.

Conflicts of interest

JJ was working at Sense Health at the time of the study where he worked on the development of the Brightr app. He was responsible for implementing the Brightr app at the high-tech company and providing the app to the experts. Though he helped with the practical setup of the study and writing the manuscript, he did not take part in any focus groups or interviews, nor did he take part in analyzing the data. Brightr product development has been able to benefit from this research by using the feedback to further develop Brightr. Currently, Brightr version 3.6 is being deployed at several organizations.

Abbreviations

mHealth: mobile health

PA: physical activity

3.5 References

1. Van den Heuvel SG, Geuskens GA, Hooftman WE, Koppes LL, van den Bossche SN. Productivity loss at work; health-related and work-related factors. J Occup Rehabil 2010 Sep;20(3):331-339. [doi: 10.1007/s10926-009-9219-7] [Medline: 19921406]
2. Robroek SJ, van den Berg TI, Plat JF, Burdorf A. The role of obesity and lifestyle behaviours in a productive workforce. Occup Environ Med 2011 Feb;68(2):134-139. [doi:10.1136/oem.2010.055962] [Medline: 20876556]
3. Wierenga D, Engbers LH, Van Empelen P, Duijts S, Hildebrandt VH, Van Mechelen W. What is actually measured in process evaluations for worksite health promotion programs: a systematic review. BMC Public Health 2013 Dec 17;13:1190 [doi: 10.1186/1471-2458-13-1190] [Medline: 24341605]
4. Rongen A, Robroek SJ, van Lenthe FJ, Burdorf A. Workplace health promotion: a meta-analysis of effectiveness. Am J Prev Med 2013 Apr;44(4):406-415. [doi: 10.1016/j.amepre.2012.12.007] [Medline: 23498108]
5. Eurofound; EU-OSHA. Eurofound. 2014. Psychosocial risks in Europe: prevalence and strategies for prevention URL: https://www.eurofound.europa.eu/publications/report/2014/eu-member-states/working-conditions/psychosocial-risks-in-europe-prevalence-and-strategies-for-prevention [accessed 2018-02-17] [WebCite Cache ID 6xJD5VVZi]
6. Chau JY, der Ploeg HP, van Uffelen JG, Wong J, Riphagen I, Healy GN, et al. Are workplace interventions to reduce sitting effective? A systematic review. Prev Med 2010 Nov;51(5):352-356. [doi: 10.1016/j.ypmed.2010.08.012] [Medline: 20801153]
7. van der Klink JJ, Blonk RW, Schene AH, van Dijk FJ. The benefits of interventions for work-related stress. Am J Public Health 2001 Feb;91(2):270-276. [Medline: 11211637]
8. Lamontagne AD, Keegel T, Louie AM, Ostry A, Landsbergis PA. A systematic review of the job-stress intervention evaluation literature, 1990-2005. Int J Occup Environ Health 2007;13(3):268-280. [doi: 10.1179/oeh.2007.13.3.268] [Medline: 17915541]
9. Richardson KM, Rothstein HR. Effects of occupational stress management intervention programs: a meta-analysis. J Occup Health Psychol 2008 Jan;13(1):69-93.[doi: 10.1037/1076-8998.13.1.69] [Medline: 18211170]
10. Bhui KS, Dinos S, Stansfeld SA, White PD. A synthesis of the evidence for managing stress at work: a review of the reviews reporting on anxiety, depression, and absenteeism. J Environ Public Health 2012;2012:515874 [doi: 10.1155/2012/515874] [Medline: 22496705]
11. Hamberg-van Reenen HH, Proper KI, van den Berg M. Worksite mental health interventions: a systematic review of economic evaluations. Occup Environ Med 2012 Nov;69(11):837-845. [doi: 10.1136/oemed-2012-100668] [Medline: 22864248]
12. Cancelliere C, Cassidy JD, Ammendolia C, Côté P. Are workplace health promotion programs effective at improving presenteeism in workers? A systematic review and best evidence synthesis of the literature. BMC Public Health 2011 May 26;11:395 [doi: 10.1186/1471-2458-11-395] [Medline: 21615940]
13. Speklé EM, Hoozemans MJ, Blatter BM, Heinrich J, van der Beek AJ, Knol DL, et al. Effectiveness of a questionnaire based intervention programme on the prevalence of arm, shoulder and neck symptoms, risk factors and sick leave in computer workers: a cluster randomised controlled trial in an occupational setting. BMC Musculoskelet Disord 2010 May 27;11:99 [doi: 10.1186/1471-2474-11-99] [Medline: 20507548]
14. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care

consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
15. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
16. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014 Jun 25;14:646 [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
17. Klasnja P, Consolvo S, Pratt W. How to evaluate technologies for health behavior change in HCI research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI 2011; May 7-12; Vancouver, Canada p. 3063-3072.
18. European Commission. Ec.europa.eu. 2014 Apr 10. Green paper on mobile Health (“mHealth”) URL: https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth [accessed 2018-02-17] [WebCite Cache ID 6xJFZbWCO]
19. Swan M. Sensor Mania! The internet of things, wearable computing, objective metrics, and the quantified self-2.0. J Sens Actuator Netw 2012 Nov 08;1(3):217-253. [doi: 10.3390/ jsan1030217]
20. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [doi: 10.2196/jmir.2583] [Medline: 23598614]
21. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11:97 [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
22. Choe EK, Lee NB, Lee B, Pratt W, Kientz JA. Understanding quantified-selfers' practices in collecting and exploring personal data. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI 2014; April 26- May 01 2014; Toronto, ON, Canada p. 1143-1152.
23. Aarts E, de Ruyter B. New research perspectives on ambient intelligence. J Ambient Intell Smart Environ 2009 Jan;1(1):5-14.
24. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
25. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007 Sep 30;9(3):e26 [doi: 10.2196/jmir.9.3.e26] [Medline: 17942388]
26. Aneni EC, Roberson LL, Maziak W, Agatston AS, Feldman T, Rouseff M, et al. A systematic review of internet-based worksite wellness approaches for cardiovascular disease risk management: outcomes, challenges & opportunities. PLoS One 2014 Jan 08;9(1):e83594 [doi: 10.1371/journal.pone.0083594] [Medline: 24421894]
27. Colkesen EB, Niessen MA, Peek N, Vosbergen S, Kraaijenhagen RA, van Kalken CK, et al. Initiation of health-behaviour change among employees participating in a web-based health risk assessment with tailored feedback. J Occup Med Toxicol 2011 Mar 09;6:5 [doi: 10.1186/1745-6673-6-5] [Medline: 21388524]
28. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012 Nov 21;14(6):e161 [doi: 10.2196/jmir.2171] [Medline: 23171838]
29. Anderson I, Maitland J, Sherwood S, Barkhuus L, Chalmers M, Hall M, et al. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. Mobile Netw Appl 2007 Jun;12(2-3):185-199. [doi: 10.1007/s11036-007-0011-7]

30. Consolvo S, Everitt K, Smith I, Landay JA. Design requirements for technologies that encourage physical activity. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2006 Presented at: CHI 2006; April 22-27, 2006; Montréal, Québec, Canada p. 457-466.
31. Consolvo S, Klasnja P, McDonald DW, Avrahami D, Froehlich J, LeGrand L, et al. Flowers or a robot army?: Encouraging awareness & activity with personal, mobile displays. In: Proceedings of the 10th international conference on Ubiquitous computing.: ACM; 2008 Presented at: UbiComp 2008; September 21-24, 2008; Seoul, Korea p. 54-63. [doi: 10.1145/1409635.1409644]
32. Lin M, Lane ND, Mohammod M, Yang X, Lu H, Cardone G, et al. BeWell+: Multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization. In: Wireless Health. New York, NY, USA: ACM; 2012 Presented at: Wireless Health 2012; October 23-25; San Diego, USA. [doi: 10.1145/2448096.2448106]
33. Bexelius C, Löf M, Sandin S, Trolle Lagerros Y, Forsum E, Litton JE. Measures of physical activity using cell phones: validation using criterion methods. J Med Internet Res 2010 Jan 29;12(1):e2 [doi: 10.2196/jmir.1298] [Medline: 20118036]
34. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009 Jan 13;11(1):e1 [doi: 10.2196/jmir.1100] [Medline: 19141433]
35. Plarre K, Raij A, Hossain SM, Ali AA, Nakajima M, Al'Absi M, et al. Continous inference of psychological stress from sensory measurements collected in the natural environment. In: Proceedings of the 10th International Conference on Information Processing in Sensor Networks.: IEEE; 2011 Presented at: 10th International Conference on Information Processing in Sensor Networks (IPSN); 12-14 April 2011; Chicago, IL, USA p. 97-108.
36. Conroy DE, Yang CH, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
37. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: A content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
38. Nield D. Fortune. 2014 Apr 15. In corporate wellness programs, wearables take a step forward URL: http://fortune.com/2014/04/15/in-corporate-wellness-programs-wearables-take-a-step-forward [WebCite Cache ID 6kgsSWeJe]
39. Gray R. Hrmagazine. 2015 Sep 03. Wearable technology for health and wellbeing URL: http://www.hrmagazine.co.uk/article-details/how-to-use-wearable-technology-for-health-and-wellbeing [accessed 2016-09-21] [WebCite Cache ID 6kgr4yNZ9]
40. van Drongelen A, Boot CR, Hlobil H, Twisk JW, Smid T, van der Beek AJ. Evaluation of an mHealth intervention aiming to improve health-related behavior and sleep and reduce fatigue among airline pilots. Scand J Work Environ Health 2014 Nov;40(6):557-568. [doi: 10.5271/sjweh.3447] [Medline: 25121620]
41. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
42. Stawarz KM, Cox AL. Designing for health behavior change: HCI research alone is not enough. 2015 Presented at: CHI 2015 workshop: Crossing HCI and Health: Advancing Health and Wellness Technology Research in Home and Community Settings; April 18-23, 2015; Seoul, South Korea.
43. West R, Michie S. A Guide to Development and Evaluation of Digital Behavior Change Interventions in Healthcare. London: Silverback Publishing; Jan 21, 2016.

44. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. JMIR Mhealth Uhealth 2016 Sep 09;4(3):e107 [doi: 10.2196/mhealth.5720] [Medline: 27613084]
45. Schelvis RM, Oude Hengel KM, Burdorf A, Blatter BM, Strijk JE, van der Beek AJ. Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs. Scand J Work Environ Health 2015 Sep 01;41(5):491-503 [doi: 10.5271/sjweh.3505] [Medline: 26030719]
46. Vosbergen S, Mahieu GR, Laan EK, Kraaijenhagen RA, Jaspers MW, Peek N. Evaluating a web-based health risk assessment with tailored feedback: what does an expert focus group yield compared to a web-based end-user survey? J Med Internet Res 2014 Jan 02;16(1):e1 [doi: 10.2196/jmir.2517] [Medline: 24384408]
47. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
48. Wixom BH, Todd PA. A theoretical integration of user satisfaction and technology acceptance. Inf Syst Res 2005 Mar;16(1):85-102. [doi: 10.1287/isre.1050.0042]
49. Poole ES. HCI and mobile health interventions: how human-computer interaction can contribute to successful mobile health interventions. Transl Behav Med 2013 Dec;3(4):402-405 [doi: 10.1007/s13142-013-0214-3] [Medline: 24294328]
50. Gulati A, Dubey SK. Critical analysis on usability evaluation techniques. IJEST 2012 Mar;4(3):990-997.
51. Vermeeren AP, Lai-Chong Law E, Roto V, Obrist M, Hoonhout J, Väänänen-Vainio-Mattila K. User experience evaluation methods: current state and development needs. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. 2010 Presented at: NordiCHI 2010; October 16-20, 2010; Reykjavik, Iceland p. 16-20.
52. Nielsen J. Usability Engineering. London: Academic Press; 1993.
53. Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualitative research methods in health technology assessment: a review of the literature. Health Technol Assess 1998;2(16):iii-ix, 1-274 [Medline: 9919458]
54. van Boeijen A, Daalhuizen J, Zijlstra J, van der Schoor R, editors. Delft Design Guide. Amsterdam, Netherlands: BIS Publishers; Apr 01, 2014.
55. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2006 Feb 01;18(1):59-82. [doi: 10.1177/1525822X05279903]
56. Bailey JE, Pearson SW. Development of a tool for measuring and analyzing computer user satisfaction. Management Science 1983 May;29(5):530-545. [doi: 10.1287/mnsc.29.5.530]
57. Coursaris CK, Kim DJ. A meta-analytical review of empirical mobile usability studies. J Usability Stud 2011 May;6(3):117-171.
58. van Lieshout M, Wiezer NM, De Korte E. The digital stress coach. Total control over your mental health, or 'Big Brother is watching you'? In: Kool L, Timmer J, van Est R, editors. Sincere Support - The Rise of the E-Coach. The Hague, Netherlands: Rathenau Instituut; Jun 2015:161-176.
59. Koldijk S, Koot G, Neerincx M, Kraaij W. Privacy and trust in context-aware systems. In: Dimitrova V, Kuflik T, Chin D, Ricci F, Dolog P, Houben GJ, editors. User Modeling, Adaptation, and Personalization. UMAP 2014. Lecture Notes in Computer Science, vol 8538. Cham, Switzerland: Springer; 2014:134-145.
60. Zhang D, Adipat B. Challenges, methodologies, and issues in the usability testing of mobile applications. Int J Hum Comput Interact 2005;18(3):293-308. [doi: 10.1207/s15327590ijhc1803_3]

61. Harrison R, Flood D, Duce D. Usability of mobile applications: literature review and rationale for a new usability model. J Interact Sci 2013 Dec;1:1. [doi: 10.1186/2194-0827-1-1]
62. Nayebi F, Desharnais JM, Abran A. The state of the art of mobile application usability evaluation. In: Electrical & Computer Engineering (CCECE). 2012 Presented at: 25th IEEE Canadian Conference on Electrical & Computer Engineering (CCECE); April 29-May 02 2012; Montreal, QC, Canada. [doi: 10.1109/CCECE.2012.6334930]
63. Tan W, Liu D, Bishu R. Web evaluation: heuristic evaluation vs. user testing. Int J Ind Ergon 2009 Jul;39(4):621-627. [doi: 10.1016/j.ergon.2008.02.012]
64. Jeffries R, Miller JR, Wharton C, Uyeda K. User interface evaluation in the real world: a comparison of four techniques. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1991 Presented at: CHI'91; April 29-May 2, 1991; New Orleans, LA, USA p. 119-124.
65. Hasan L, Morris A, Probets S. A comparison of usability evaluation methods for evaluating e-commerce websites. Behav Inf Technol 2012 Jul;31(7):707-737. [doi: 10.1080/014929X.2011.596996]
66. Fu L, Salvendy G, Turley L. Effectiveness of user testing and heuristic evaluation as a function of performance classification. Behav Inf Technol 2002;21(2):137-143. [doi: 10.1080/02699050110113688]
67. Lathan CE, Sebrechts MM, Newman DJ, Doarn CR. Heuristic evaluation of a web-based interface for internet telemedicine. Telemed J 1999;5(2):177-185. [doi: 10.1089/107830299312140] [Medline: 10908430]
68. Nielsen K, Randall R. Opening the black box: presenting a model for evaluating organizational-level interventions. EJWOP 2013;22(5):601-617. [doi: 10.1080/1359432X.2012.690556]
69. Adams A, Cox AL. Questionnaires, in-depth interviews and focus groups. In: Cairns P, Cox AL, editors. Research Methods for Human Computer Interaction. Cambridge, UK: Cambridge University Press; 2008:17-34.

3.6 Multimedia appendix 1. Codebook

This is a Multimedia Appendix to a full manuscript published in the JMIR mHealth and uHealth. For full copyright and citation information see http://mhealth.jmir.org/0000/0/e0/ doi:10.2196/mhealth.6335.

The codebook is an adapted version of the codebook used by Vosbergen et al (2014) who used the domains and topics described by Wixom and Todd (2005) in their study on evaluating a web-based health risk assessment tool. Definitions are adapted from the framework of Wixom and Todd (2005), Bailey and Pearson (1983) and Vosbergen et al. (2014) and further specified further to the mHealth application that was used in the current study. Some topics that were part of the framework of Wixom and Todd (2005) or part of the codebook of Vosbergen et al (2014) are not described. New topics were added in case the researchers felt it was missing, these were performance (system quality), content, visibility of system status (information quality), adherence (usefulness).

Table 6. Definitions of domains and topics of the codebook used to categorize all remarks of employees and experts

Domain	Topic	Definition
System quality		Users' perceived quality of the app
	Accessibility	The degree to which the system is accessible to its users
	Timelines (responsiveness)	The availability of the system's output at a time suitable for its use
	Flexibility	The capacity of the system to change or adapt in response to new conditions, demands, or circumstances
	Integration	The ability of systems to communicate/ transmit data between systems servicing different functional areas, e.g. to link together different components of the app to act as a coordinated whole
	Efficiency	The rate or speed at which the system enables users to accurately and successfully complete a task
	Tailoring	The ability of the system to take user characteristics into account
	Language	The set of vocabulary, syntax, and grammatical rules used to interact with the system

Domain	Topic	Definition
	Errors / error prevention	The methods and policies governing correction and rerun of incorrect system output
	Performance	Technical performance of the hardware and the software
Information quality		Users' perceived quality of the information given by the app
	Accuracy	Users' perception that the information is correct
	Precision	The variability of the output information from that which it purports to measure
	Reliability	The consistency and dependability of the output information
	Currency	The age of the output information
	Completeness	The degree to which the app provides all information perceived as necessary by the user
	Format	The layout and display of the information throughout the entire app
	Volume	The amount of information conveyed to users
	Content	The content of the information provided
	Visibility of system status	The degree to which the system keeps the users informed about what is going on, through appropriate feedback within reasonable time.
Service quality		Users' perceived quality of the service delivered by the professionals associated with the app
	Relationship with app provider	The method and manner of interactions between users and app provider
	Communication with app provider	The way information is exchanged among users and app provider
	Technical competence of app provider	The skills and expertise of the app provider
	Attitude of app provider	The way users perceive the attitude of the app provider towards users and their health experiences
	Schedule of products or services	The timetable for system output, services, and procedures

Domain	Topic	Definition
	Processing of change requests	The manner, methods, and required time the staff respond to users' requests
	Response time	The time between users' requests for service or action and response to these requests
	Means of input with app provider	The method and medium by which users receive services from app provider and/or the system and the perceived usefulness of this service
Usefulness		General usefulness of the app for its users
	Usefulness	The extent to which the app actually helps to solve users' problems
	Relevancy	The degree of congruence between users' needs and requirements and what the app provides
	Adherence	The extent to which the app stimulates the user to continue to use the app
Ease of use		Degree to which users believe that using the app is effortless
	User friendly	The app is pleasant to use
	Easy to use	The app effectively fills users' needs and is fast and free of errors
	Learnability	The extent to which users are able to easily learn and understand how to operate the system
	Memorability	The extent to which users are able to remember how to use the system
Outcome expectations		Congruence between users' expectations and actual situation with regard to using the app and the feedback provided by the system
	Expectations	Users' expectations of the system
	Understanding of system	The degree of comprehension that a user possesses about the systems or services that are provided
	Confidence in the system	Users' feelings about the reliability of the app and the feedback provided by the system
	Feelings of participation	The degree of involvement and commitment which the user shares with app provider and other app users toward the functioning on system and services

Domain	Topic	Definition
	Feelings of control	Users' perceived power to regulate/ influence the feedback provided by the system
	Degree of training	The amount of specialized instruction and practice that is afforded to the user to increase the user's proficiency in utilizing the system capability that is unavailable.
	Accuracy	Users' perception that the provided feedback is congruent with their expectations about their behavior
	Health & performance effects	Users' (expected) changes in lifestyle, work performance, or other health-related issues as a result of using the App
Organizational factors		Influence of the organization, procedures, and choices on the quality of the app
	Management involvement	The positive or negative degree of interest, enthusiasm, support or participation of any management level above the users own level toward the App or towards the provider of the App
	Organizational competition	Congruence between the assessment and feedback provided by the system and an external health professional/ system (e.g. coach, other app, other system)
	Security of data	The safeguarding of data from misappropriation or unauthorized access, alteration or loss
	Documentation	The recorded description of an information system. This included formal instructions to the user and to program staff about the app
	Timing	The availability of the measurements and feedback of the app at a time suitable for its use
	Communication	The availability of correct information before using the app

References

Bailey JE, Pearson WE (1983). Development of a tool for measuring and analyzing computer user satisfaction. Management Sci, 29(5): 530-545.

Vosbergen S, Mahieu GR, Laan EK, Kraaijenhagen RA, Jaspers MWM, Peek N. (2014). Evaluating a web-based health risk assessment with tailored feedback: what does an expert focus group yield compared to a web-based end-user survey? Journal of medical internet research, 16(1): e1. DOI: 10.2196/jmir.2517.

Wixom BH, Todd PA (2005). A theoretical integration of user satisfaction and technology acceptance. Information systems research, 16(1): 85-102.

3.7 Multimedia appendix 2. Illustrative quotes

This is a Multimedia Appendix to a full manuscript published in the JMIR mHealth and uHealth. For full copyright and citation information see http://mhealth.jmir.org/0000/0/e0/ doi:10.2196/mhealth.6335.

Table 7. Illustrative examples of employees' quotes from interviews and focus groups and experts' quotes within domains and by topic with coded value (positive (+), negative (-), neutral (0) quotes or recommendation (R)).

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
System quality	**Accessibility**	-	-	-
	Timelines (responsiveness)	Most of the time I check the app right before I go to sleep, because I would like to know how I did today (0)	When I receive a message, I take a look at the app. However, I take a look less often now, mostly in the evening or when I am at the toilet (0)	I looked at the app at the end of the day (0)
	Flexibility	I took a look at the shift work app. Actually I don't work in shift and did not use it, but if you travel overseas it tells you about getting rid of your jetlag, but it didn't work that well for me: it seems to start automatically, however, if you go to Belgium it seems to think that I was in another time zone (which was not the case). I think it is not that accurate (-)	I work in nightshifts, I received the advice to go cycling at 4.00 A.M. (-)	The advice has to fit with the working tasks (R)

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
	Integration	I would not try to make all functionalities in one app, but I would use apps that are already measuring that. I would choose different good apps and combine them (R).	I find retrieval of data not really easy, but I would like that (-)	I would appreciate it when the app would contain connecting services with specific wearables. Since I do not believe that a smart phone can measure everything. E.g., I put my working phone off when I quit working and a wearable is easier (R)
	Efficiency	When I swipe the screen, it reacts after 5 seconds (-)	It is not responsive enough, the screen hampers somewhere halfway. It feels like a 15 years ago (-)	-
	Tailoring	There are many different kinds of workers in our company, some need physical activity advice (they lift weights a lot for instance), others have to exercise more (e.g. sitting at desk to much). App can help them to become aware. Goal of such an app is to try to get people think whether they are in balance. Do they have sufficient activity? (R)	To put things on or off, that would be convenient (R).	I believe it is important for apps: an app needs to get to know me, what my goals are, what change techniques do work, pieces of surprise (all of a sudden different tips). Towards artificial intelligence. Currently apps are often one-dimensional, it should be a learning algorithm (R)

3

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
	Language	-	It has a high ' hippy-level'. To fill in 'awesome' as an answer feels awkward. I do not like that. Just put things as they are (-)	English language was not of good quality (-)
	Errors / error prevention	-	Using the app, it's just trial and error (-)	You might give instruction in the beginning and, after a while, when a person understands the app, put them off. For instance a balloon that says "swipe to the right for…" or " did you know….". After a while these instructions disappear, once it is clear that people have found certain components or understand certain functionalities. Now it's just the screens (R)
	Performance	The new version was very much advanced but it was draining my battery. Not very handy going from one to another meeting and not being able to recharge phone (-)	Because of the battery use I do not use the app anymore (-)	Technically, the app feels slow (-)

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
Information quality	**Accuracy**	I did not trust the sleep measurements (-)	If the measurement is good, you also attach more value to the advices (R)	A tip was not correct: to eat repeatedly small bits of food may work for weight management, but it is not healthy from the point of view of dental care, for which it is really, really bad (-)
	Precision	If I move a lot at work, then it's measuring much. Over the weekend, it lies on my desk and measures nothing. It gives an indication and that is good enough (0)	-	-
	Reliability	Do I feel that measurements are true? Sometimes I have moved a lot but then I see no icon. It works only if you have your phone with you. Sometimes I do not have it with me (-)	I notice that I have reached my goal every day, but often this was not true, I have doubts about the accuracy of the app (-)	Movement measures were accurate with me (+)
	Currency	I read all of the articles, background information. They were not refreshed too often (-)	I have read the background information, but I find it a pity that there was no information added. Too bad this information was not being updated (-)	-

3

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
	Completeness	Tips are good. More exercise, better sleep, less stress. You can find a lot of information about that, for instance via the background info, you can also click further (+)	I find it special that it has different aspects together in one app (+)	Then you hope to get more information, but that is not the case. I have searched many times for information (-)
	Format	The motion screen is nice (+)	The design with sliding screens is OK (+)	Plots next to each other, e.g. how sleep relates to resilience, you won't see that. Also the course of resilience over time is not available (-)
	Volume	Annoying about it: the amount of questionnaires you get (+)	I find messages that I constantly get a bit redundant and too much (-)	Quite a lot of reading (-)
	Content	I would like more information about food, what you should eat. Shift workers have to eat very fast at times (and therefore, the choices are not always healthy). I would like tips about food that is healthy and that you can eat fast (R)	I would like to see more about calories in the app, I would like to add a food coach (R)	I think of an app that shows the effects of your behavior, for example to show visually: what you have done now leads to this effect (R)
	Visibility of system status	-	-	-

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
Service quality	**Relationship with app provider**	Technical support contact: I got personal response (+)	-	-
	Communication with app provider	Technical support contact: I got response but not what they had done. I would have liked to have received a message that I had to remove the app for a while (R)	There are many people who have had contact with the Brightr team (0)	-
	Technical competence of app provider	You can see that they constantly take further steps in developing the app	-	-
	Attitude of app provider	-	-	-
	Schedule of products or services	Technical support contact: I received a response within a day (+)	-	-
	Processing of change requests	-	-	-
	Response time	-	-	-
	Means of input with app provider	Right now I lost the password (after installing) and cannot find it anymore (-)	I also asked the coach, I got a good answer (+)	-
Usefulness	**Usefulness**	With the app, I can better recognize coming stressful times (+)	It triggers to do things better in your behavior. The fact that I saw that I pretty quickly	I find it useful to gain insight that I move too little, especially when I

3

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
			reached my physical activity goals was good, to see that it was not a problem for me (+)	am working at home (+)
	Relevancy	The best parts, for me, are the shift work part (I work morning, evening, night shift). Since I try to follow the advices about maintaining a healthy lifestyle and working with shift hours. It helped me to keep down the stress in my body. I felt that I could focus better on the task during the daily (nightly) work (+)	The mental resilience part is doing absolutely nothing for me. I often think for what reason am I doing this? If you are doing well, it has no added value (-)	Mental resilience also triggered ... well, it yielded only frustration, I did not receive any tips (-)
	Adherence	I stopped using it at some time. I understand that there has to be a learning period, but there was too minimal progress in improvement of app. I prefer to have for instance 20 questions if that provides good feedback instead of 1 question and not relevant feedback (-)	A community would challenge people (R)	Because there are too many problems with the app, people will throw it away quickly (-)

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
Ease of use	**User friendly**	Usability: good, clear, but I am a technical guy. Sometimes the app is unstable and I have to restart it. Sometimes it is too slow (+)	It is not really user friendly (-)	Usability is not good enough to make people use the app spontaneously, it is not clear what it aims to do (-)
	Easy to use	Usability? It is unclear how you need to do some settings. For example, when your goals change. Difficult to set (-)	-	-
	Learnability	You have to work your way through it. That is also a disadvantage (-)	To me it's all still a bit of an unexplored area. I have a phone to call people (0)	I found out only after a week that there was more than just physical activity. I swiped once accidentally and there they were: all sorts of modules! (0)
	Memorability	Installing and operating was pretty clear, it was also not complicated to discover and remember how the app worked, that was all very logical (+)	-	-
Outcome expectations	**Expectations**	Why I use such apps: it is basically extra information. If you want to improve running, you need that extra	I know by myself when I was not physically active. I would especially like to see a trend and the app	-

3

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
		information, and an app can offer that. I hoped that Brightr would offer that as well, but it is rather basic (-)	pointing that out. Since it is impossible to be successful every day (R)	
	Understanding of system	Then you have another screen with biorhythms, I don't know what I have to do with that, that is totally unclear to me (-)	-	-
	Confidence in the system	Relationship between what the app indicates and what I actually sleep was not entirely clear (-)	If the feedback is correct, then this information is nice to know. However, if the feedback is not correct, then it is no longer useful, that is a reason not to use it anymore (-)	Implementation is weak in some parts. For example mental resilience scores are not accurate. It measures whether you experience stress, but there is no link to the causes of your stress. Also, certain tips were bad in the mental resilience part (-)
	Feelings of participation	I expected to be able to see 'xx% has answered this', that you can see how you rank in comparison to your colleagues (R)	To follow each other, give each other compliments. I would like that in the Brightr app (R)	Social: compare with peers and so on, that would help to keep the app accessible (R)
	Feelings of control	I found it useful to be able to adjust my goals, that encouraged me to	Especially that you can set things by yourself I find important (R)	If you can turn of the parts that you don't like, then one does not change what might be

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
		really get more exercise (+)		necessary from a health point of view (R)
	Degree of training	-	-	You would have to be able to turn on instructions at start of use. And turn it off once someone knows how the app works (R)
	Accuracy	I noticed on the pop-ups that cycling is not measured accurately, same with walking. For example the pop up said: "you have been 30 minutes physically active', but in reality I cycled for an hour (-)	Those graphs, that's not quite the image I have myself (-)	-
	Health & performance effects	There are many different kind of workers at our company, some need physical activity advice (e.g. they lift weights a lot at work), others have to exercise more (e.g. sitting at desk to much). App can help them to become aware. Goal of such an app is to try to get people think whether they are in balance. Do they	An app can improve your health (+)	The goal of the app is to raise awareness (0)

3

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
		have sufficient activity. I think it is possible that an app could help to reach goals Reaching some goals must be possible (+)		
Organizational factors	**Management involvement**	The best way to introduce an app for employees: that definitely has to go through management, from higher level management to lower management. Not via email, people get so much email that people miss such things. Via management is also a signal that they encourage you and allow you to use it. They should use it also themselves, as an example. It is important to know that the management supports it (R)	I would not introduce the app through the management. I also don't know if my management is aware of the existence of this app. It would be interesting to know which managers use the app. If the management is actively using this app, that would help, this is a big challenge (R)	It matters how you introduce the app: as a corporate app or as a fun gift (0)
	Organizational competition	It is useful to monitor for a company whether your employees are doing well. For example a health check. You want to know whether	An app is a good tool within a larger program (R)	App should be a part of a bigger program, in terms of intervention. It is supportive within an intervention (R)

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
		there are improvements with your employees, you can measure that and, if necessary, plan actions (R)		
	Security of data	I realize that with the app some things are logged somewhere central. Because you can compare yourself to the rest of the organization. I mean, how would you make comparisons otherwise? But it is not that secret for me to worry about it (0)	Privacy, I have nothing to hide (0)	Within an organization, it is always the question what the organization plans to do with the data. You should be very clear on this (R)
	Documentation	-	A simple manual, with photos or something, in which you can search quickly. A help function in the app itself is also an option. To be able to start quickly (R)	Tips, instructions or background information consisting of lots of text will not be read (R)
	Timing	It should be more predictable when new articles are added. I mean not every day, but it should be predictable. You don't know when it's updated. It is OK if it is not	It strikes me that this app especially disturbs you when you are deeply concentrated at work (-)	-

Domain	Topic	Illustrative quotes interviews employees	Illustrative quotes focus groups employees	Illustrative quotes focus group experts
		updated very often. But if it is not updated regularly then it should be more hidden (R)		
	Communication	For instance, you need to inform people with posters about the existence of the app. It is also important to pay more attention to the benefits of this app. Not everyone would like to install the application because they do not know specific benefits for them. The benefits should be clarified (R)	I was introduced to the app via e-mail (0)	Embed the app in team sessions, embed it in a wider program (R)

4

Effects of a feedback signal in a computer mouse: laboratory experiment[1]

[1] Published as: De Korte, E.M., De Kraker, H., Bongers, P.M., & Van Lingen, P. (2008). Effects of a feedback signal in a computer mouse on movement behaviour, muscle load, productivity, comfort and user friendliness. *Ergonomics, 51*(11), 1757-1775.

E.M. de Korte[a,b]
H. de Kraker[a]
P.M. Bongers[a,c]
P. van Lingen[a]

[a] TNO, Hoofddorp, The Netherlands
[b] Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands
[c] Research Centre Body@Work, Free University of Amsterdam, Amsterdam, The Netherlands

Abstract
To study the effects of a tactile feedback signal in a computer mouse on reduction of hovering behaviour and consequently on changes in muscle load, productivity, comfort and user friendliness, a comparative, experimental study with repeated measures was conducted. Fifteen subjects performed five trials with different mouse actions and a standardised task, once with a mouse with the feedback signal and once with a mouse without the feedback signal. Holding the hand just above the mouse caused higher muscle loading than clicking and scrolling. Holding the hand on the mouse caused higher muscle loading than resting the hand on the desk. The feedback signal effectively decreased hovering behaviour. It also led to a more dynamic activation pattern of the extensor muscles of the forearm. The overall opinion of the feedback signal for future use was rated as somewhat variable. No effects on discomfort or productivity were found. The use of a mouse with a tactile vibrating feedback signal seems promising for preventing arm complaints, although more research is needed to establish the clinical relevance.

Keywords: feedback; computer mouse; movement behaviour; muscle load; productivity; comfort.

4.1 Introduction

Of the European working population, 19% works with a computer constantly and this number is still increasing (Paoli and Merllié 2001). In office workers, neck and upper limb symptoms are common (de Kraker and Blatter 2005). Recently, the potential harmful effects of computer work were studied in several longitudinal studies (Marcus *et al*. 2002, Andersen et al. 2003, Jensen 2003, Korhonen *et al*. 2003, Kryger *et al*. 2003, Brandt *et al*. 2004, Juul-Kristensen *et al*. 2004, Lassen *et al*. 2004, IJmker *et al*. 2007). Results from these studies show that mouse usage longer than 10-20 h per week might be a risk factor for hand/arm symptoms. The effects on neck/shoulder symptoms are less consistent. In addition, an increase in computer use may increase the number of errors and the fatigue level of the computer user resulting in reduction of productivity (Hagberg *et al*. 2002).

A possible explanation of these unfavourable health effects of mouse use is the Cinderella hypothesis, which postulates that with prolonged low-static muscle activity the low-threshold (type I) motor units are always first recruited. With low-static muscle activity these low-threshold motor units will be active almost continuously without periods for recovery. This will lead to microscopic damage to the muscle fibres. Because only a small part of the muscle is active, the neuromuscular feedback system that indicates fatigue and gives signals to interrupt muscle activity is not working properly.

As a result, the muscle remains activated while certain fibres are already damaged, through which damage may increase further (Hägg 1991). Over the past few years a reasonable number of studies have been published, showing that some motor units in the trapezoid muscle (Forsman *et al*. 1999, Kadefors *et al*. 1999, Westgaard and de Luca 1999, Thorn *et al*. 2002) and the wrist extensors (Forsman *et al*. 2002) remain continuously active when a broad range of tasks are being performed. Along with the Cinderella hypothesis, poor blood circulation in the muscle is often regarded as a cause of damage as it leads to an energy deficit in the muscle cells (Keller *et al*. 1998).

Poor circulation can be the result of a rise in pressure in the muscle caused by muscle activity. If pressure is elevated for an extended period, even relatively low values can damage muscle cells (Hargens *et al*. 1981). The mechanisms described indicate that both the level of activation and the uninterrupted duration of activation are important (Mathiassen *et al*. 1995, 2002, Jensen *et al*. 1999, van Dieën *et al*. 2003).

According to these theories, postural changes, rest breaks and micro-rest breaks may positively influence the pattern of muscle activation, in a way that ensures computer workers stay fit during the day and will be more productive. van den Heuvel *et al*. (2003) showed that indeed frequent rest breaks and alternation of postures and working tasks lead to better comfort, better wellbeing and, most likely, higher productivity. Also, Thompson (1990) and Henning (1997) showed productivity increases (18-25%) after introducing active work rest breaks. This resulted in working less overtime and therefore lower costs (Thompson 1990). In addition, Galinsky *et al*. (2000) and McLean *et al*. (2001) showed productivity increases with rest breaks. Balci and Aghazadeh (2003) also showed that the introduction of micro breaks interrupted the static workload of computer work, which resulted in higher productivity, increased working pace, more comfort and a decrease in fatigue and risks for symptoms of neck and upper limbs. According to these studies, it seems that introduction of (micro) pauses is a meaningful strategy for increasing productivity, comfort and health.

However, aspects that have not received much attention thus far are the human behavioural factors to actually take rest breaks and avoid unhealthy behaviour. Thé *et al*. (2003) found that behaviour plays a crucial role in actually taking rest breaks. Results from that study show that computer workers did not use the opportunities to work healthily and productively, resulting in decreased productivity.

Stimulating healthy work rest behaviour of computer users to support health, productivity and comfort by ways of an intelligent feedback system, which adapts to the user needs, seems promising. Already, there are examples of such systems on the market, for example, software to stimulate taking rest breaks in time by means of visual feedback. The evidence for health-promoting effects of this software is mixed (Van den Heuvel *et al*. 2003, Brewer *et al*. 2006, Slijper *et al*. 2007). A major drawback of the visual feedback used in this rest break software is the interference with the daily task performance. It is clear that an intelligent feedback system should be able to stimulate healthy work rest behaviour without intervening in the work tasks through so-called non-intrusive feedback.

A good way to study the effects of this type of feedback is to propose micro-rest breaks during mouse use. The use of the computer mouse is often characterised by frequent and prolonged inactive periods. Often, the user will rest his/her hand on or just above the mouse, during which the muscles remain activated, for instance, whilst viewing the computer screen or doing something else. This is called hovering behaviour, of which an example is shown in Figure 1.

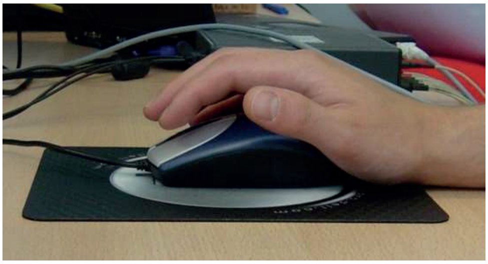

Figure 1 Example of hovering behaviour: holding the hand just above the mouse

Hovering behaviour results in a maintained extended wrist posture, which is known to be a risk factor for developing repetitive strain injury (RSI) complaints, particularly when this awkward posture is sustained for a long time (Jensen *et al*. 1998). If computer users should relax the hand instead, a behavioural change may result in less static load and more micro breaks for the upper extremity.

The purpose of the present study is to determine the effects of a tactile, vibrating feedback signal in a computer mouse on hovering behaviour, productivity, comfort and muscle load. In this study, a mouse will be used that gives a tactile vibration signal to remind the user to let go of the mouse and relax the hand on the desk. van de Ven and Ruijsendaal (2005) compared four different non-intrusive feedback signals on hovering behaviour. One of the signals tested was a continuous tactile vibrating signal, which appeared to be able to decrease hovering behaviour. In order not to interrupt the user in the work process it is important that the signal is noticeable, but not distractive.

The main research questions to be answered in this study are: What is the effect of a mouse with a tactile, vibrating feedback signal compared to a regular mouse on:

› hovering behaviour: duration and frequency of hand position on and just above the mouse?
› muscle activity: muscle activation and relaxation in the shoulder and forearm muscles?
› productivity: task performance and self-reported productivity?
› discomfort: locally perceived discomfort per body region?

- comfort and user friendliness: subjective experiences of using a mouse with a feedback signal?

It is hypothesised that usage of a mouse with a tactile, vibrating feedback signal will decrease hover time. This reduced hover time will result in decreased static muscular load and increased muscular relaxation of the forearm extensor muscles. Therefore, perceived exertion will decrease. Furthermore, productivity is expected to be the same or slightly increased when using the mouse with a tactile, vibrating feedback signal, compared to a regular mouse.

4.2 Methods

4.2.1 Subjects

A total of 15 symptom-free subjects participated in this study, three men and 12 women. Subject characteristics were (means ± SD): age 24 ± 2 years; height 174 ± 8 cm; and weight 70 ± 16 kg. Participants did not have a history of musculoskeletal complaints in neck, shoulders or hands/wrists. They were experienced computer users. Before the experiment, the participants signed informed consent.

4.2.2 Procedures

The subjects were seated on a height-adjustable chair at a desk. The measurements consisted of two parts. The first part consisted of five trials of mouse-operating actions to compare muscle activation during these actions and the second part consisted of a standardised computer task to study the effect of the feedback signal. During all measurements, the subject used their hand with which they normally use the mouse.

During the experiments, the subjects used a two-button mouse with scroll wheel and a vibrating tactile feedback signal (Hoverstop-mouse©; Hoverstop BV, Amsterdam, Netherlands). This mouse is comparable to a regular mouse, except it is equipped with a sensor capable of detecting the hand being in close proximity to the mouse, and with a small vibrating motor. Mouse action (using buttons or scroll wheel) is monitored. The motor starts to vibrate if the hand is present on or just above the mouse but the buttons and/or scroll wheel are not operated for more than 10 s. The vibration signal gives the user feedback in order to let go of the mouse. In this case, the subjects were instructed to place the hand on the desk.

The motor vibrates at a 40% level of the motor power, with a maximum duration of 4 s. Within the 4 s, the vibration will continue until the mouse is clicked, the scroll wheel is rolled or the hand is removed. Every time the subject clicks or scrolls, the timer-counter is reset to zero, so during ordinary, active use of the mouse a signal is never generated. The vibration does not react to cursor movements. Therefore, 'stirring' the mouse will not affect the vibration.

It is also possible to turn the motor off, so that there is no feedback signal; the sensor is, however, still capable of detecting the presence of the hand. With the feedback signal turned off, the mouse will be referred to as a regular mouse; with the signal turned on the mouse will be referred to as the experimental mouse.

Part 1: Mouse actions
The subjects were asked to perform five trials with different mouse actions: rest the hand on the desk; rest the hand completely on the mouse; hold the hand just above the mouse (Figure 1); click the left button of the mouse; roll the scroll wheel. Clicking and scrolling were performed with a frequency of 1 Hz, with the aid of a metronome. These actions all reflect actions regularly observed during normal computer work. These five trials were presented in a randomised order and performed for 1.5 min each. During the five trials the feedback signal was turned off.

Part 2: Standardised tasks
A standardised task was designed in MS Word using the toolbox and Visual Basic to study the effects of a feedback signal during mouse operation. Each subject was asked to read a text that was visible in a textbox of 4 cm high, with a scrollbar on the right side. The box contained approximately 110 words, which would take approximately 1 min to read, after which the subject had to scroll down for the next piece of text. The subjects were allowed to use the mouse scroll wheel or drag the scroll bar by using the mouse button. They were not allowed to use the keyboard. After the whole text was read, the subjects had to answer eight multiple-choice questions. Total time given to read the text and answer the questions was 8 min. The standardised task was performed twice, once with the regular mouse, once with the experimental mouse. Texts and mice were randomly offered.

4.2.3 Muscle activity

Muscle activity was measured using electromyography (EMG). Electrodes were placed on the prepared skin of the m. trapezius pars descendants (bilaterally), m. deltoïdeus pars anterior, m. deltoïdeus pars acromialis, m. extensor carpi radialis longus, m. extensor digitorum communis and m. flexor digitorum superficialis, all at the dominant side.

Bipolar Ag/AgCl (Blue Sensor, N-00-S; Medicotest, Denmark) surface electrodes with a recording distance of 20 mm were used. A reference electrode was placed on C7 processus spinosus. A maximal voluntary contraction (MVC) of each muscle was performed three times with a duration of 3 s while using manual resistance (Kendall *et al*. 1983) with at least 30 s rest in between. The trial with the highest average of myo-electric activity was used as the reference value for normalisation. The EMG signals were amplified 20 times (Porti 17S; TMS, Enschede, The Netherlands), band-pass filtered (10-400 Hz) and A-D converted at 1000 Hz.

EMG signals were recorded for 1.5 min for each of the five mouse action trials and for 8 min for each trial of the standardised task. The first 15 s and the last 15 seconds of the EMG signals were subtracted, so 1 min of the five mouse action trials and 7.5 min of the standardised tasks were used for analysis. EMG data were digitally high-pass filtered at 30 HZ cut-off with a finite impulse response filter (FIR), order 100, to reduce electrocardiogram (ECG) contamination. All EMG signals were rectified, filtered (fourth order Butterworth low-pass 5 Hz) and normalised.

Static (P10), median (P50) and peak (P90) load levels of EMG were computed by using amplitude probability distribution functions (APDF). The APDF method of the EMG signal offers a profile of the variation in muscle activity amplitude during the analysed time period in terms of the probability level of a given fraction of time not being exceeded. The calculated P10, P50 and P90 EMG activity levels are estimates of the EMG activity levels in terms of percentage MVC, below which the muscle activity remained for 10%, 50% and 90% respectively of the recording time (Jonsson 1978).

Furthermore, to study the distribution of micro-pauses in the EMG pattern, a gap analysis was performed (number of gaps per min and mean gap length). This method is an analysis of the temporal pattern of the muscle activity by quantifying the numbers and duration of short pauses (gaps) in the EMG signal. The number of gaps was defined as the number of periods with an EMG level below 0.5% MVC for at least 0.2 s (Veiersted *et al*. 1990).

4.2.4 Hovering behaviour

Hovering behaviour was determined using computer use registration software (RSI-Master; http://www.rsimaster.com/index.html). This registration software is able to register the data from the sensor and the vibration motor in the mouse. Hovering behaviour was looked at by total duration of hand position on or above the mouse (active mouse use included) and frequency of detection of hand presence. Since the standardised tasks are very similar to each other, it is plausible that the hand presence during active mouse use will be constant and differences in durations can be attributed to differences in hovering behaviour without using the mouse.

4.2.5 *Productivity*

More and more studies on musculoskeletal symptoms and computer work take some kind of productivity measure into account. However, office work is too complex for measuring productivity directly, since a great variety of tasks are performed. Hagberg *et al*. (2002) argued for the use of self-reported productivity. The present study measured productivity both objectively by assessing the number of questions answered and the amount of correct answers within the time allowed for the standardised tasks and subjectively by rating four items of self-reported productivity on a 7-point scale in a questionnaire that was presented after the experiments (amount of work, working tempo, number of errors, duration of task).

4.2.6 *Discomfort*

As a measure of discomfort, subjects were asked to rate their perceived discomfort in 13 different body regions of the upper body before and after each trial of standardised tasks. Subjective ratings of localised musculoskeletal discomfort were obtained of neck and both sides of shoulders, upper arms, elbows, forearms, wrists and hand/fingers. The Dutch validated method for body part discomfort was used (Lokaal Ervaren Ongemak; van der Grinten and Smitt 1992), an ordinal scale ranging from no discomfort (0) to extreme discomfort (10). This scale is derived from the original Borg scale.

4.2.7 *Comfort and user friendliness*

After the experiment, a questionnaire was presented to the subjects containing subjective questions and measurement scales about productivity (see above), comfort and user friendliness. Comfort was rated on a 7-point scale. Comfort items included overall comfort of the experimental mouse, comfort of the feedback signal and duration of the feedback signal (7-point scales). Also, subjects had to give their opinion on comfort items such as functionality, appearance and colour.

User friendliness included items on use and learning experiences (7-point scale), distraction (how distracting is the signal?), irritation (how irritating is the signal?), noticeability (how noticeable is the signal?), motivation (does the signal motivate the user to remove the hand?) and logic (is it a logical signal for removing the hand?) (visual analogue scale (0-10)). Finally, some general questions on user friendliness were asked.

4.2.8 *Statistics*

Differences in EMG analysis of the five trials of mouse actions were tested using repeated measures multivariate ANOVA ($p < 0.05$). The three within subject factors were task (five levels), muscle (seven levels) and muscle load (three levels). Interactions between these factors were not analysed in this context. Contrasts, however, were used to determine differences between levels of the within subject factor task, corrected for the other factors. Because three subjects did not perform the clicking and scrolling tasks, they were excluded from this analysis.

For the standardised tasks, differences in hovering behaviour and EMG were tested with a paired t-test ($p < 0.05$) and differences in objective productivity and discomfort were tested using Wilcoxon signed ranks tests.

4.3 Results

4.3.1 Part 1: Mouse actions

Muscle activity

As expected, significantly lower static, median and peak load levels of muscle activation of the m. extensor digitorum and m. extensor carpi radialis were found when placing the hand on the desk compared to placing the hand on the mouse, holding the hand above the mouse, clicking or scrolling (Figure 2).

A notable result is the unfavourable, relatively high muscle activity when holding the hand just above the mouse. In addition, it appears that even when subjects are asked to relax the hand on the mouse, this leads to significantly higher activity than relaxing the hand on the desk. For example, median activity of the m. extensor digitorum was 7.7% MVC with the hand above the mouse, compared to 0.4% MVC with the hand on the desk, 1.5% MVC with hand on mouse, 3.8% MVC clicking and 3.6% MVC scrolling. As expected, no differences were found between clicking and scrolling in m. extensor digitorum. In addition, no differences were found between hand above mouse, clicking and scrolling in m. extensor carpi radialis.

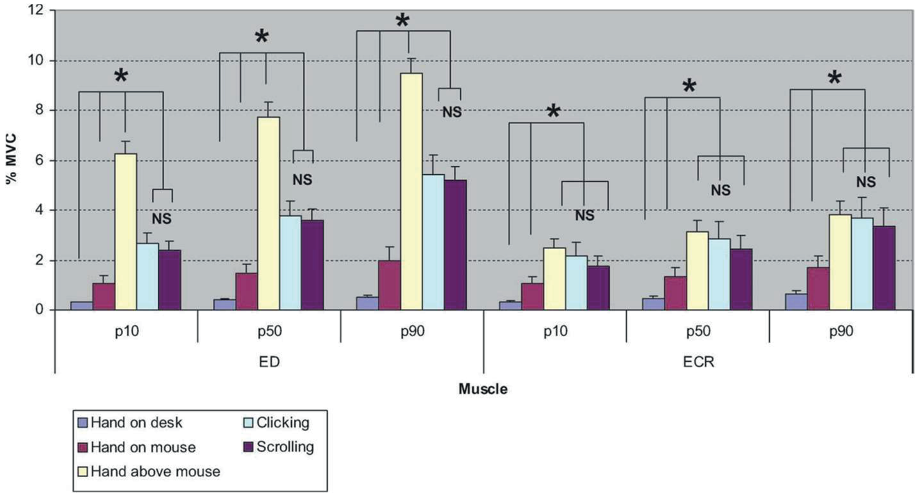

Figure 2 Electromyographic activity (P10-, P50-, P90-values, in % maximal voluntary contraction (MVC)) of m. extensor digitorum (ED) and m. extensor carpi radialis (ECR), all at the dominant side for five trials of mouse actions: hand on desk; hand on mouse; hand above mouse; clicking; scrolling. NS = not significant. *p<0.05.

Significantly more gaps per min (more micro-pauses in the muscle activation pattern) were found in the m. extensor digitorum when the hand was placed on the desk (11.8 ± 14.8), compared to holding the hand above the mouse (0 ± 0), clicking (0 ± 0) or scrolling (0.2 ± 0.6) (means ± SD). No differences were found between hand on desk and hand on mouse. In addition, gap length (s) in the m. extensor digitorum was longer (longer micro-pauses in the muscle activation pattern) when the hand was placed on the desk (13.8 ± 18.1) compared to holding the hand on the mouse (3.0 ± 8.6) above the mouse (0 ± 0), clicking (0 ± 0) or scrolling (0.0 ± 0.1) (means ± SD).

Comparable results for gap length were found for the m. extensor carpi radialis. Gap length was longer when the hand was placed on the desk (20.4 ± 25.6) compared to holding the hand above the mouse (0.1 ± 0.4), clicking (0.3 ± 0.7) or scrolling (0.3 ± 0.9) (means ± SD). No differences were found between hand on desk and hand on mouse. In addition, no differences in gap frequency were found in the extensor muscles of the forearm.

These five trials did not show any electromyographic effects in the m. flexor carpi radialis and the neck and shoulder muscles.

4.3.2 Part 2: Standardised tasks

Hovering behaviour

The frequency of hand detection on and/or just above the mouse was significantly higher with the experimental mouse compared to the regular mouse (Table 1). The total duration of hovering behaviour was significantly less with the experimental mouse (49% of total time) compared to the regular mouse (75% of total time).

Table 1 Frequency and duration (mean, SD) of hand detected on or just above the mouse for regular and experimental mouse during standardised task

	Regular mouse		Experimental mouse		
	Mean	SD	Mean	SD	*p*
Frequency hand on/above mouse	21.8	8.9	7.4	8.2	0.002
Duration hand on/above mouse (min)	3.9	1.5	6.0	1.4	0.003

Note: The *p*-value gives the difference between the regular and experimental mouse.

Since both variables were significantly different, the correlation between total number of times and total duration was also examined (Figure 3). Both correlations appeared to be weak. Two deviated points mainly caused the weak correlation for the experimental mouse. Since the total number of points is small, this influence is considerable. As can be seen from Figure 3, a clear trend is likely to arise if more measurements were

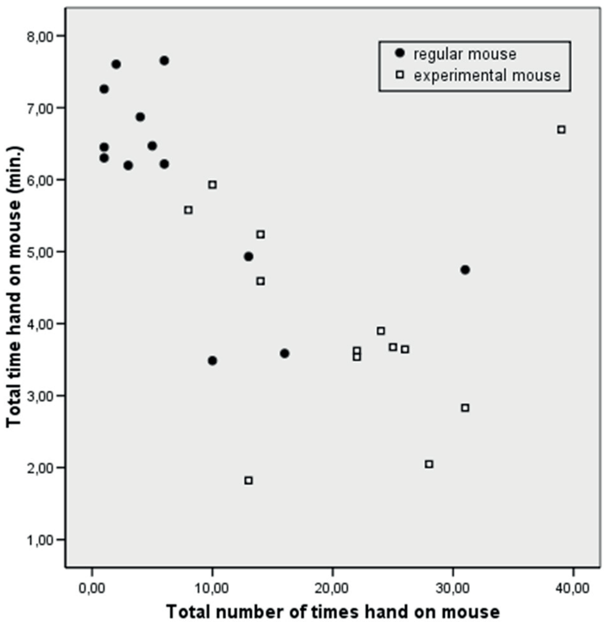

Figure 3 Relationship between total duration hand on mouse and total number of times hand on mouse for regular mouse and experimental mouse (n = 14)

added. The data points indicating the total duration of hovering with the regular mouse are at first sight more clustered and do not seem to be related to the number of times the hand grabs the mouse.

In a previous study of van de Ven and Ruijsendaal (2005), task difficulty appeared to influence hovering behaviour. Although tasks of approximately the same level were offered, texts may still be of influence. To see if the texts influenced hovering behaviour, hover-time per mouse and per text was examined separately. As can be seen in Figure 4, the direction of the difference in hover-time between the regular mouse and the experimental mouse is similar for both texts: both texts show a decrease in hover-time with the experimental mouse. It is, however, quite safe to conclude that the experimental mouse does indeed decrease total hover-time. This assumption is supported by the non-significant result of the t-test, in which text was used as the independent variable of hover-time.

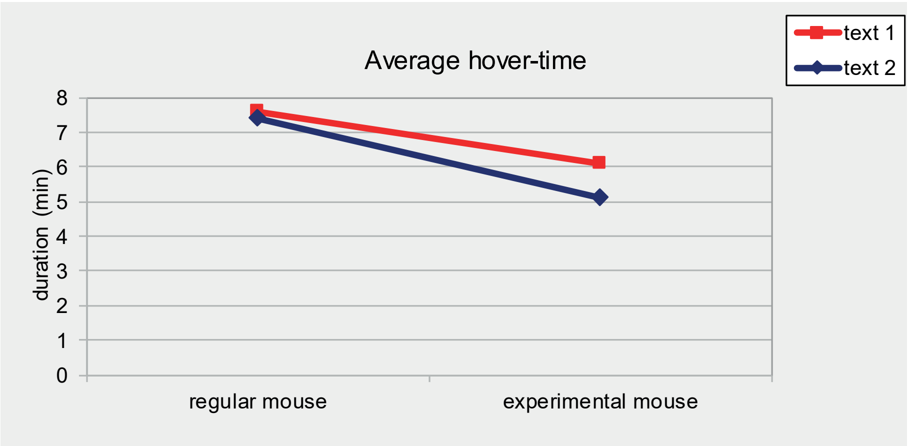

Figure 4 Average hover-time for text 1 and text 2 for both the regular mouse and experimental mouse

Muscle activity

Significantly lower median (P50) and peak (P90) levels of muscle activation of the m. extensor digitorum and lower peak levels of the m. extensor carpi radialis were found with the regular mouse compared to the experimental mouse (Figure 5). No differences were found for the static load levels of these muscles (P10).

The gap frequency of the m. deltoïdeus anterior was higher with the experimental mouse (13.2 ± 9.9) than the regular mouse (10.4 ± 8.7). The gap length of the m. extensor digitorum was longer with the experimental mouse (1.5 ± 1.3) compared to the regular mouse (0.6 ± 0.9) (means ± SD).

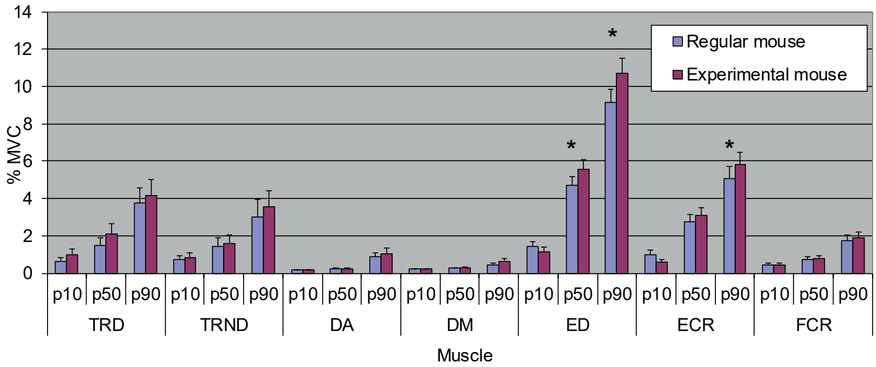

Figure 5 Electromyographic activity (P10-, P50- and P90-values, in % maximal voluntary contraction (MVC)) of m. trapezius pars descendens dominant side (TRD), m. trapezius pars descendens non dominant side (TRND), m. deltoïdeus pars anterior (DA), m. deltoïdeus pars medior (DM), m. extensor digitorum (ED), m. extensor carpi radialis (ECR) and m. flexor carpi radialis (FCR), all at the dominant side for regular mouse and experimental mouse. *$p<0.05$

Productivity

No objective differences were found between the performance on the tasks with the experimental mouse and the regular mouse. Neither was difference found in the number of correct answers (Table 2).

Table 2 Objective productivity (total answered questions and total correct answered questions) during the standardised tasks with the regular mouse and the experimental mouse (mean, SD)

	Regular mouse		Experimental mouse		
	Mean	SD	Mean	SD	*p*
Total answered questions	5.5	2.5	5.7	2.0	0.167
Total correct answered questions	4.5	2.4	4.8	2.1	0.077

Note: The *p*-value gives the difference between the regular and experimental mouse.

Self-reported productivity showed that 57.1% of the subjects judged the amount of work they could perform with the experimental mouse to be the same as with the regular mouse (Figure 6). For working speed and number of mistakes, similar results were found: 64.3% said they can work as fast with the experimental mouse and 57.1% estimated they make as many mistakes with the experimental mouse as with the regular mouse. However, 50.0% judged the time necessary to complete the task to be either ‘slightly more’, ‘more’ or ‘much more’ with the experimental mouse.

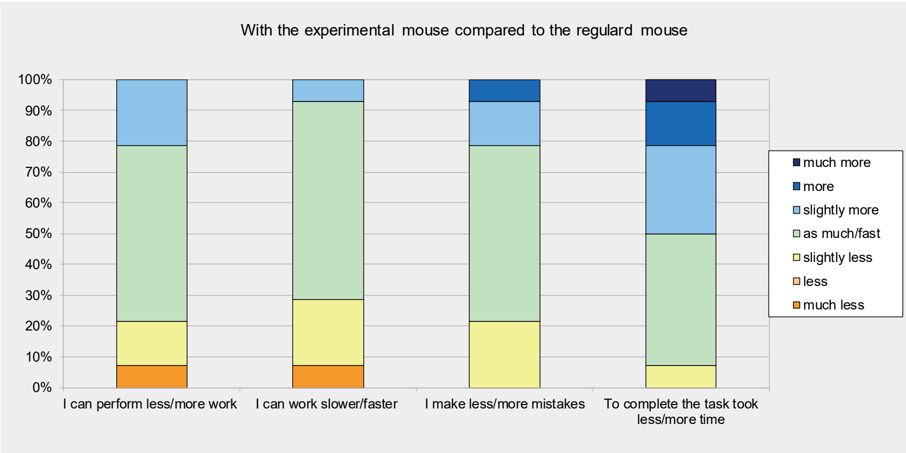

Figure 6 Subjective productivity during the two tasks (n = 14)

Discomfort

No significant differences were found between the scores in locally perceived discomfort in the 13 body regions before and after the tasks and between the regular mouse and the experimental mouse.

Comfort and user friendliness

The self-reported overall comfort of the experimental mouse reported in the questionnaire varied considerably between subjects: 21.4% of the subjects (n = 3) rated the experimental mouse uncomfortable or a little bit uncomfortable, 28.6% (n = 4) rated it neither comfortable nor uncomfortable, while 50.0% (n = 7) rated it comfortable or somewhat comfortable.

The vibration signal itself was rated uncomfortable by one subject (7.1%) and somewhat uncomfortable by 50% of the subjects. The duration of the signal was rated uncomfortable or somewhat uncomfortable by 35.8% of the subjects (n = 5) and the same percentage rated the duration of the signal as a little bit comfortable or comfortable.

In Figure 7, the subjects' opinion on comfort items such as functionality, appearance and colour are presented.

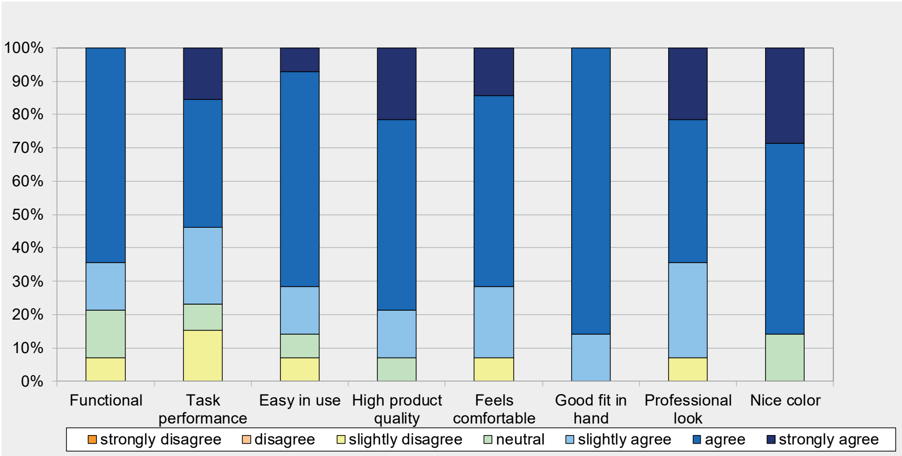

Figure 7 Opinions on comfort items

On a visual analogue scale (0-10) distraction, irritation, noticeability, motivation and logic of the feedback signal were rated (Table 3). The high noticeability of the signal corresponded with low irritability.

Table 3 Ratings of distraction, irritation, noticeability, motivation and logic of the feedback signal (mean ± SD) on a visual analogue scale from 0 (not at all) -10 (extreme)

Distraction	Irritation	Noticeability	Motivation	Logic
5.8 ± 2.7	3.1 ± 3.2	8.0 ± 2.2	7.4 ± 2.4	7.3 ± 2.5

Several questions on user friendliness of working with the experimental mouse were asked after the experiments. Learning to use the experimental mouse was rated a little bit difficult by one subject and working with the experimental mouse was rated a little bit difficult by two subjects. The other subjects rated these items as neutral or little to very easy. The basic idea of getting a warning if unhealthy behaviour occurs is appreciated by 57% of the subjects. The vibration signal and working with the experimental mouse did increase the awareness of working in the right posture for 11 subjects, at times for two subjects and one subject reported that the signal did not increase the awareness of the working posture. When the subjects were asked if they would like to use the experimental mouse in the future, opinions were mixed: 35.7% answered ‘yes, often’; 28.6% answered ‘yes, sometimes’; and 35.7% answered ‘no, never’. All in all, 53.8% of the subjects preferred a mouse with feedback signal to a mouse without a signal.

4.4 Discussion

4.4.1 Mouse actions

As expected, differences in extensor muscle activity were found between the mouse action trials. A noticeable result was the significant higher muscle activity during holding the hand above the mouse, compared to clicking and scrolling. In addition, a significant difference was found between resting the hand on the desk compared to holding the hand on the mouse. A plausible explanation for both these findings might be that computer users subconsciously try to prevent their fingers from clicking the buttons accidentally, even when they do not actively use the mouse.

In this study, hand postures, clicking and scrolling were simulated. The muscle activation pattern of these tasks could be different in real practice. For example, in the simulated task with the hand placed on the mouse, subjects are possibly more relaxed compared to a situation in practice, where computer workers alternate between active and passive periods of working and anticipate working tasks.

p-Values of the EMG signals and gap frequencies of extensor muscles in the forearm and trapezius muscles found for holding the hand above the mouse, clicking and scrolling are comparable to those reported in other studies on mouse or computer tasks (Laursen *et al*. 2002, Blangsted *et al*. 2004). Laursen *et al*. (2002) also found almost zero gap frequency in the forearm extensors during mouse tasks. In the present study, a large contrast in gap frequency of the m. extensor digitorum was found with the hand placed on the desk (11.8 per min). This finding could be explained by the relaxed posture, which is confirmed by the results of the *p*-values. It may be concluded that hovering behaviour leads to sustained muscle activation of the extensor muscles, with a negative effect on the distribution of micro-pauses in the EMG signals. Although the Cinderella hypothesis (Hägg 1991) gives a plausible explanation for static loading of the muscle fibres, the relevance of this hypothesis for developing complaints in office work is still not clear (Visser and van Dieën 2006). Also, other hypotheses discussed in the review of Visser and van Dieën (2006), such as the hypothesis on impaired blood flow, fail to explain completely the pathophysiology of work-related upper extremity disorders. Nonetheless, it seems beneficial to avoid sustained muscle activation to prevent the development of complaints (Visser and van Dieën 2006).

4.4.2 Standardised tasks

Hovering behaviour

The use of a feedback signal leads to decreased duration of the hand on or just above the mouse. It was found that subjects react to the signal by placing the hand on the desk and, as a result of reacting to the feedback signals, placing the hand more often on and off the mouse. The mouse tasks performance becomes a less static and more dynamic activity.

The conclusion that the experimental mouse decreases hover behaviour was also found in an earlier study of van de Ven and Ruijsendaal (2005). The difference in hover duration between the regular mouse and the experimental mouse, in percentage of total time, was even larger in the current study (26%) than found by van de Ven and Ruijsendaal (2005), who found a difference of 16%.

The concept of feedback to workers has shown to be effective in other studies. For example, Eklöf *et al*. (2004) found positive effects of different feedback methods. They used oral feedback to individual workers, to the supervisor and the whole workgroup. All three methods showed a positive effect on workplace design and work technique. Visschers *et al*. (2004) compared four warnings with a no-warning situation and the effect on changing working postures: a warning displayed on a computer screen (with the text ‘Caution!’); a warning hanging on the wall; an educational brochure; and a neutral interruption (neutral image) on the computer screen. The results showed that the computer warning led to significantly more correct posture adjustments. Gerard *et al*. (2002) studied the short- and long-term effects of enhanced auditory feedback on typing force, EMG and comfort while typing. The introduction of enhanced auditory feedback caused a 10-20% reduction in the 90th percentile typing force, finger flexor EMG and finger extensor EMG. It also appeared that full adaptation to the signal appeared within 3 min of exposure to the auditory feedback. This adaptation continued roughly at the same level during the following 2-week period. A subset of the subjects participated in an experiment after 4 months and they were able to further continue to reduce their typing force and EMG. After 1 week of intermittent auditory feedback, subjects were able to type with approximately the same force and EMG regardless of the presence or absence of the feedback. This finding suggests that subjects were able to use the enhanced feedback to alter their typing behaviour and were able to carry that learning over to when they typed without feedback.

Madeleine *et al*. (2006) studied the effect of electromyographic and mechanographic (MMG) biofeedback on upper trapezius muscle activity during standardised computer work. The results showed that audio or visual biofeedback from an electromyographic or MMG signal were effective in diminishing upper trapezius static muscle load.

A recent study of van den Heuvel *et al*. (2007) showed that the effect of unfavourable work-related exposure (i.e. high job demands and prolonged visual display unit work) on increased neck and upper limb symptoms was mediated by a high-risk work style, indicating that these working conditions mainly increase symptoms when the worker also adopts a high-risk working behaviour. A high-risk work style implies, for instance, taking shorter or fewer breaks or even skipping breaks, working through pain, anticipating the possible negative reactions of colleagues and making high demands on one's own performances at work.

These results confirm the findings of a limited number of other studies examining the relationship between work style and neck and upper limb symptoms (Haufler *et al*. 2000, Nicholas *et al*. 2005). The concept of work style has been developed by Feuerstein and it is conceptualised as a learned and reinforced strategy for coping with job demands that may affect musculoskeletal health and not as a personality trait; thus, with consistent feedback this behaviour may be influenced. The present study also shows that a feedback signal coming from a computer mouse has a very promising effect on changing working techniques and/or working postures.

However, some methodological comments have to be made in respect of the present study. Of 15 subjects who participated, one subject performed the task with the feedback signal turned on without hovering for at least 10 s. Therefore, this subject did not get the feedback signal and could not make a judgement about the experimental mouse, its feedback signal and the possible effects on productivity, comfort and health. A possible explanation for this finding is the variety of work style and work rest strategies between individuals. For instance, during the experiment some subjects crossed arms during reading, while others pressed the mouse button continuously during reading and scrolling. Sudhakaran and Mirka (2005) studied the role of personality types and break-taking behaviour. The results of their study showed a wide range in work–rest strategies. Some chose regularly scheduled breaks, others seemed to identify a specific pain threshold at which they would take a break, while still others adopted a strategy of taking a small number of longer breaks. Although personality type did not have a significant effect on break-taking behaviour, the individual break-taking behaviour patterns seemed to be stable over time, indicating that there are probably additional individual characteristics that may be driving the response. Nevertheless, in spite of varieties in work style and work–rest strategies, a clear effect was found in the present study. A factor that also might have influenced the results is the standardised task used in this study. The task was designed to study influence of feedback on mouse-hovering behaviour. The aim was to reflect general working practice as much as possible. However, different types of computer tasks influence the way input devices, e.g. a mouse, are used. For example, mouse use during browsing differs from mouse use during word processing (de Korte *et al*. 2003). It is expected that a feedback signal to prevent hovering behaviour is especially effective in computer tasks with intensive mouse use alternated with reading tasks and/or looking at the computer screen, such as browsing and searching databases. Further research is necessary to confirm this assumption.

Muscle activity

The results showed an increased level of median and peak muscular activity in the extensor muscles with the experimental mouse. No effects on static activity were found. At the same time, an increased gap length was found in the m. extensor digitorum, indicating less continuous activity. Interruptions in electromyographic muscle activity may prevent muscular pain (Veiersted *et al*. 1990), although the meaning of these interruptions for the development of complaints in office work remains unclear (Visser and van Dieën 2006).

A very plausible explanation for the higher median and peak levels (APDF) in combination with less continuous muscular activity (more/longer gaps) with the experimental mouse is the increased frequency of hand movement above the mouse and decreased duration of hand detection on or above the mouse, which causes more dynamic movements of the arm. This means that computer users appear to change their mouse operation style from static into a more dynamic one. It could be argued that a more dynamic mouse-operating strategy is preferable to static mouse operation. Westad *et al*. (2003) studied the effect of brief increases in contraction amplitude of the trapezius muscle on motor unit recruitment and de-recruitment. The results indicate that force (i.e. EMG) variation in either direction promotes motor unit de-recruitment during sustained contractions of the trapezius muscle. They conclude that from a behavioural standpoint, if silent periods are important in the control strategy of postural motor units, it seems that this mechanism relies on force-variation to facilitate that. Westgaard and de Luca (1999) concluded that this so-called substitution mechanism of motor units protects motor units in postural muscles from excessive fatigue when there is a demand for sustained low-level muscle activity. Also, Sundelin (1993) found variation in shoulder muscle activity to be preventive for muscle fatigue. These preventive effects of varied muscle activity patterns might also be applicable on the extensor muscles of the forearm.

Interestingly, an increased gap frequency of the m. deltoideus anterior was also found. This effect was not found during the five trials of mouse actions, but when performing a task it might be the case that the feedback signal also has a positive side effect on unfavourable mouse positions. In practice, unfavourable mouse positions are often seen (Karlqvist *et al*. 1994, Harvey and Peper 1997, Dennerlein and Johnson 2006). The mouse is positioned away from the centre of the body, through which the shoulder is exorotated and anteflected. M. deltoideus anterior is one of the muscles that activates in that case. The muscle load decreases when the mouse is placed close to the centre of the body, which is exactly what happens as subjects respond to the feedback signal and place the hand in front of the body.

Productivity

In order to draw conclusions about the effect of decreased hovering behaviour on productivity, musculoskeletal health and comfort, it is necessary to affirm that hovering behaviour in the experimental condition did decrease. It can be concluded that hover duration in the experimental condition was significantly shorter and thus it can be concluded that the experimental set-up was successful.

Objective productivity, expressed as the number of questions answered and the number of correct answers, was not different between the regular and experimental mouse. This might be explained by the short duration of the task and should be investigated further during tasks of longer duration. This is supported by the findings of Liao and Drury (2000), who studied interrelationships of posture, comfort and performance. They found that performance was not as closely linked to discomfort and posture as would have been expected, although some trends were found. They argued that performance measures might not be vulnerable or sensitive to low levels of perceived discomfort and fatigue and with task durations shorter than 2 h.

Subjective productivity showed that the majority of the subjects rated the experimental mouse 'as productive as' the regular mouse. One variable for subjective productivity shows a deviation: 50% of the subjects report they need more time to complete the task. This is contradictory to the opinion that the subjects report they can perform as much work, can work as fast and make as few mistakes with the experimental mouse. The discrepancy between the objective and subjective measured results might be caused by the short habituation to the signal; it is possible that the subjects did not get quite used to the signal within the short duration of the experiment, which may have influenced the subjective ratings.

More meaningful production effects should be measured in a field study, before relevant conclusions can be drawn.

Discomfort

No effects on the body part discomfort scale were found. This is most likely due to the combination of short duration of the experiments with a low-intensity task. To find differences on this scale, a longer duration or a more intensive task has to be done.

Comfort and user friendliness

Self-reported comfort of the experimental mouse varied considerably among subjects. Comfort predominantly seems to be rated as neutral to comfortable. Also, user friendliness was rated reasonably well. Overall, learning to use a mouse with a feedback signal and working with the mouse was not that difficult. The same holds for the irritation and noticeability results. Subjects reported that the mouse helped to remind them of unfavourable behaviour and it did increase awareness. All in all, 50% of the subjects preferred a mouse with feedback signal to a mouse without a signal.

4.5 Conclusions

Holding the hand above the mouse causes higher muscle loading than clicking and scrolling. In addition, holding the hand on the mouse causes higher muscle loading than resting the hand on the desk. Furthermore, a feedback signal can decrease hovering behaviour. This coincides with a change in mouse-operating behaviour to more dynamic movements of the forearm and hand, due to increased movements. More dynamic mouse-operating behaviour leads to a more dynamic activation pattern of the extensor muscles of the forearm and a positive effect on the distribution of micro-pauses. This may contribute to prevention of musculoskeletal complaints and staying fit during the day. However, there is a strong need to perform studies that could link these findings to symptoms to establish the clinical relevance.

Productivity measures seemed to remain equal, although subjective ratings did not support these findings. Therefore, an in-depth analysis in a field study with measurements over a longer period of time is recommended. The same holds for locally perceived discomfort measures.

Although the overall opinion of the feedback signal for future use was rated somewhat variable, it appeared to meet the requirements of being noticeable, but not distractive. To gain more insight into the optimal noticeability–irritability relationship, future research is required.

Future research priorities are effects of feedback on productivity and discomfort in a field study, preferably a randomised controlled trial. Such a field study could also give more insight into learning and task effects. The use of a mouse with a tactile vibrating feedback signal seems promising for preventing arm complaints.

Acknowledgements

We thank the Dutch Ministry of Economic Affairs and Hoverstop BV for supporting this study.

4.6 References

Andersen, J.H., *et al*., 2003. Computer use and carpal tunnel syndrome: a 1-year follow-up study. *The Journal of the American Medical Association,* 289 (22), 2963-2969.

Balci, R. and Aghazadeh, F., 2003. The effect of work-rest schedules and type of task on the discomfort and performance of VDT users. *Ergonomics,* 46 (5), 455-465.

Blangsted, A.K., Søgaard, H.C., and Sjøgaard, G., 2004. The effect of physical and psychosocial loads on the trapezius muscle activity during computer keying tasks and rest periods. *European Journal of Applied Physiology,* 91, 253-258.

Brandt, L.P.A., *et al*., 2004. Neck and shoulder symptoms and disorders among Danish computer workers. *Scandinavian Journal of Work Environment and Health,* 30 (5), 399-409.

Brewer, S., *et al*., 2006. Workplace interventions to prevent musculoskeletal and visual symptoms and disorders among computer users: A systematic review. *Journal of Occupational Rehabilitation,* 16, 325-358.

De Korte, E.M., *et al*., 2003. Use of input devices within different types of computer tasks. In: C. Stephanidis, ed. *HCI International 2003 Adjunct Proceedings*. Heraklion: Crete University Press, 41-42.

De Kraker, H. and Blatter, B., 2005. Prevalence of RSI-complaints and the occurrence of risk factors in 15 European countries (in Dutch). *Tijdschrift voor gezondheidswetenschappen*, 1, 8-15.

Dennerlein, J.T. and Johnson, P.W., 2006. Changes in upper extremity biomechanics across different mouse positions in a computer workstation. *Ergonomics*, 49 (14), 1456-1469.

Eklöf, M., *et al*., 2004. Feedback of workplace data to individual workers, workgroups or supervisors as a way to stimulate working environment activity: a cluster randomized controlled study. *International Archives of Occupational and Environmental Health*, 77, 505-514.

Forsman, M., *et al*., 1999. Motor-unit recruitment in the trapezius muscle during arm movements and in VDU precision work. *International Journal of Industrial Ergonomics,* 24, 619-630.

Forsman, M., *et al*., 2002. Motor-unit recruitment during long-term isometric and wrist motion contractions: a study concerning muscular pain development in computer operators. *International Journal of Industrial Ergonomics,* 30, 237-250.

Galinsky, T.L., *et al*., 2000. A field study of supplementary rest breaks for data-entry operators. *Ergonomics,* 43 (5), 622-638.

Gerard, M.J., *et al*., 2002. Short and long term effects of enhanced auditory feedback on typing force, EMG, and comfort while typing. *Applied Ergonomics,* 33, 129-138.

Hagberg, M., Tornqvist, E.W., and Toomingas, A., 2002. Self-reported reduced productivity due to musculoskeletal symptoms: associations with workplace and individual factors among white-collar computer users. *Journal of Occupational Rehabilitation,* 12 (3), 151-162.

Hägg, G.M., 1991. Static workloads and occupational myalgia - a new explanation model. In: P.A. Anderson, D.J. Hobart, and J.V. Danoff, eds. *Electromyographical kinesiology*. Amsterdam: Elsevier Science, 141-143.

Hargens, A.R., Schmidt, D.A., and Evans, K.L., 1981. Quantitation of skeletal muscle necrosis in a model compartment syndrome. *Journal of Bone and Joint Surgery, American volume,* 63A, 631-636.

Harvey, R. and Peper, E., 1997. Surface electromyography and mouse use position. *Ergonomics,* 40, 781-789.

Haufler, A.J., Feuerstein, M., and Huang, G.D., 2000. Job stress, upper extremity pain and functional limitations in symptomatic computer users. *American Journal of Industrial Medicine,* 38 (5), 507-515.

Henning, R., 1997. Frequent short rest breaks from computer work: effects on productivity and wellbeing at two field sites. *Ergonomics,* 40, 78-91.

IJmker, S., *et al*., 2007. Should office workers spend fewer hours at their computer? A systematic review of the literature. *Occupational and Environmental Medicine,* 64, 211-222.

Jensen, C., 2003. Development of neck and hand/wrist symptoms in relation to duration of computer use. *Scandinavian Journal of Work Environment and Health,* 29 (3), 197-205.

Jensen, C., *et al*., 1998. Job demands, muscle activity and musculoskeletal symptoms in relation to work with the computer mouse. *Scandinavian Journal of Work Environment and Health,* 24 (5), 418-424.

Jensen, C., Laursen, B., and Sjøgaard, G., 1999. Shoulder and neck. In: S. Kumar, ed. *Biomechanics in ergonomics*. London: Taylor and Francis, 201-220.

Jonsson, N., 1978. Kinesiology with special reference to electromyographic kinesiology. *Contemporary Clinical Neurophysiology,* 34, 417-428.

Juul-Kristensen, B., *et al*., 2004. Computer users' risk factors for developing shoulder, elbow and back symptoms. *Scandinavian Journal of Work Environment and Health,* 30 (5), 390-398.

Kadefors, R., *et al*., 1999. Recruitment of low threshold motor-units in the trapezius muscle in different static arm positions. *Ergonomics,* 42 (2), 359-375.

Karlqvist, L., Hagberg, M., and Selin, K., 1994. Variation in upper limb posture and movement during word processing with and without mouse use. *Ergonomics,* 37, 1261-1267.

Keller, K., Corbett, J., and Nichols, D., 1998. Repetitive strain injury in computer keyboard users: Pathomechanics and treatment principles in individual and group intervention. *Journal of Hand Therapy,* 11, 9-26.

Kendall, F.P., *et al*., 1983. Muscles testing and function. Baltimore: Williams & Wilkins.

Korhonen, T., *et al*., 2003. Work-related and individual predictors for incident neck pain among office employees working with video display units. *Occupational and Environmental Medicine,* 60, 475-482.

Kryger, A.I., *et al*., 2003. Does computer use pose an occupational hazard for forearm pain; from the NUDATA study. *Occupational and Environmental Medicine,* 60 (11), e14.

Lassen, C.F., *et al*., 2004. Elbow and wrist/hand symptoms among 6,943 computer operators: a 1-year follow-up study (the NUDATA study). *American Journal of Industrial Medicine,* 46 (5), 521-533.

Laursen, B., *et al*., 2002. Effect of mental and physical demands on muscular activity during the use of a computer mouse and a keyboard. *Scandinavian Journal of Work Environment and Health,* 28 (4), 215-221.

Liao, M.H. and Drury, C.G., 2000. Posture, discomfort and performance in a VDT task.

Ergonomics, 43 (3), 345-359.

McLean, L., *et al*., 2001. Computer terminal work and the benefit of microbreaks. *Applied Ergonomics,* 32, 225-237.

Madeleine, P., *et al*., 2006. Effects of electromyographic and mechanographic biofeedback on trapezius muscle activity during standardizes computer work. *Ergonomics,* 49 (10), 921-933.

Marcus, M., *et al.*, 2002. A prospective study of computer users: II. Postural risk factors for musculoskeletal symptoms and disorders. *American Journal of Industrial Medicine,* 41, 236-249.

Mathiassen, S.E., Burdorf, A., and van der Beek, A.J., 2002. Statistical power and measurement allocation in ergonomic intervention studies assessing upper trapezius EMG amplitude. A case study of assembly work. *Journal of Electromyography and Kinesiology,* 12, 45-57.

Mathiassen, S.E., Winkel, J., and Ha¨ gg, G.M., 1995. Normalization of surface EMG amplitude from the upper trapezius muscle in ergonomic studies. A review. *Journal of Electromyography and Kinesiology,* 5 (4), 197-226.

Nicholas, R.A., Feuerstein, M., and Suchday, S., 2005. Workstyle and upper-extremity symptoms: A biobehavioural perspective. *Journal of Occupational and Environmental Medicine,* 47 (4), 352-361.

Paoli, P. and Merllié, D., 2001. Third European survey on working conditions 2000. European Foundation for the Improvement of Living and Working Conditions. Luxembourg: Office for Official Publications of the European Communities.

Slijper, H.P., *et al*., 2007. The effects of pause software on the temporal characteristics of computer use. *Ergonomics,* 50 (2), 178-191.

Sudhakaran, S. and Mirka, G.A., 2005. A laboratory investigation of personality type and break- taking behavior. *International Journal of Industrial Ergonomics,* 35 (3), 237-246.

Sundelin, G., 1993. Patterns of electromyographic shoulder muscle fatigue during MTM-pace repetitive arm work with and without pauses. *International Archives of Occupational and Environmental Health,* 64, 485-493.

Thé, K.H., *et al*., 2003. Behavioural aspects of developing or sustaining musculoskeletal disorders. *In: Proceedings of the XVth Triennial Congress International Ergonomics Association and the 7th Joint Conference of Ergonomics Society of Korea, Japan Ergonomics Society,* 24-29 August, 2003. Seoul, South Korea [CD-ROM]. The Ergonomics Society of Korea. 24-29 August 2003. Seoul, South Korea.

Thompson, D.A., 1990. Effect of exercise breaks on musculoskeletal strain among data-entry workers: a case study. In: S.L. Sauter, M. Dainoff, and M. Smith, eds. *Promoting health and productivity in the computerized office: models of successful ergonomic interventions*. London: Taylor & Francis.

Thorn, S., *et al*., 2002. Low-threshold motor unit activity during a 1-h static contraction in the trapezius muscles. *International Journal of Industrial Ergonomics,* 30, 225-236.

Van den Heuvel, S.G., *et al*., 2003. Effects of software programs stimulating regular breaks and exercises on work-related neck and upper-limb disorders. *Scandinavian Journal of Environmental Health,* 29 (2), 106-116.

Van den Heuvel, S.G., et al., 2007. Workstyle and overcommitment in relation to neck and upper limb symptoms. *International Journal of Behavioral Medicine,* 14 (1), 12-20.

Van der Grinten, M.P. and Smitt, P., 1992. Development of a practical method for measuring body part discomfort. In: S. Kumar, ed. *Advances in industrial ergonomics and safety IV*. London: Taylor & Francis.

Van de Ven, J.G.M. and Ruijsendaal, M., 2005. *Studying the effect of non-intrusive feedback on mouse hovering behaviour.* TNO report no. TNO-DV3 2005-C 019 TNO: Soesterberg, The Netherlands (Confidential).

Van Dieën, J.H., Visser, B., and Hermans, V., 2003. The contribution of task-related biomechanical constraints to the development of work-related myalgia. In: H. Johansson, U. Windhorst, M. Djupsjöbacka, and M. Passatore, eds. *Chronic work-related myalgia. Neuromuscular mechanisms behind work-related chronic muscle syndromes*. Gävle: Gävle University Press.

Veiersted, K.B., Westgaard, R.H., and Andersen, P., 1990. Pattern of muscle activity during stereotyped work and its relation to muscle pain. *International Archives of Occupational Environmental Health,* 62, 31-41.

Visschers, V.H.M., *et al*., 2004. The effects of warnings and an educational brochure on computer working posture: a test of the C-HIP model in the context of RSI-relevant behaviour. *Ergonomics,* 47 (14), 1484-1498.

Visser, B. and van Dieën, J.H., 2006. Pathophysiology of upper extremity muscle disorders. *Journal of Electromyography and Kinesiology,* 16, 1-16.

Westad, C., Westgaard, R.H., and de Luca, C.J., 2003. Motor unit recruitment and derecruitment induced by brief increase in contraction amplitude of the human trapezius muscle. *Journal of Physiology,* 552 (2), 645-656.

Westgaard, R.H. and de Luca, C.J., 1999. Motor unit substitution in long-duration contractions of the human Trapezius muscle. *Journal of Neurophysiology*, 82 (1), 501-504.

5

Effects of a feedback signal in a computer mouse: short-term RCT in the field[1]

[1] Published as: De Kraker, H., De Korte, E.M., Van Mil, F.L.K., Rijs, B.P., Bongers, P.M. (2008). The effect of a feedback signal in a computer mouse on hovering behaviour, productivity, comfort and usability in a field study. *Ergonomics, 51*(2), 140-155.

Heleen de Kraker[a]
Elsbeth M. de Korte[a]
Ferry L.K. van Mil[a]
Bartjan P. Rijs[a]
Paulien M. Bongers[a,b]

[a] TNO Quality of Life, Hoofddorp, The Netherlands
[b] Research Centre Body@Work, Free University of Amsterdam, Amsterdam, The Netherlands

Abstract

The aim of this study was to determine the effect of a tactile feedback signal on hovering behaviour, productivity, usability and comfort after 1 week of using an experimental mouse. In a randomized controlled trial, a regular computer mouse was compared to a new developed mouse with a tactile, vibrating feedback signal to prevent unnecessary hovering above the computer mouse. According to this study, participants do decrease their hovering behaviour when using a mouse with tactile feedback. Furthermore, the mouse with tactile feedback did not influence productivity. Usability was rated somewhat mixed. The use of a mouse with a tactile vibrating feedback signal seems promising for preventing neck, shoulder and arm complaints. Further research is needed to study long-term effects on (prevention of) neck, shoulder and arm complaints and development of learning effects.

Keywords: Prevention of work related upper limb symptoms; Productivity; Feedback; Behaviour; Computer mouse.

5.1 Introduction

5.1.1 Motivation

It is estimated that 15% of the Dutch working population experience limitations due to work-related upper limb symptoms (WRULS). The total costs in The Netherlands due to WRULS are estimated to be €2.1 billion per year. The main part (around €962 million) of these costs is caused by absence from work. The costs due to the loss of work productivity are estimated to be €808 million per year (Blatter *et al*. 2006).

5.1.2 Literature

At the end of 2003, TNO studied the working conditions in The Netherlands (Van den Bossche and Smulders 2004). According to this study, 74% of Dutch employees worked at least 1 h per day with a computer, while Heinrich and Blatter (2005) found that 43% of frequent computer users work on average 6-8 h per day with their computer.

The long working hours, together with static stress in the muscles of the extremities, repetitive movements and awkward joint angles, lead to an increased risk for physical symptoms to the neck, shoulder, arm, wrist, hand and fingers (Rempel *et al*. 1992). All of these risk factors are present simultaneously during computer work.

With continued computer work, the computer mouse is the most commonly used input device other than the keyboard (Cook and Kothiyal 1998, Wahlström *et al*. 2000). The statement ‘the mouse is more prone to causing injury than a keyboard’ (Pascarelli and

Kella 1993), which was ill-founded at that time, has nowadays been confirmed by a number of longitudinal studies. Brandt *et al*. (2004) and Lassen *et al*. (2004) show more elevated risks for upper limb disorders when duration of mouse use was studied compared to the duration of keyboard use.

Kryger *et al*. (2003) performed a study in which they related self-reported hours of mouse use per week to forearm complaints. Forearm complaints were defined as self-reported pain in the forearm within the past 7 d combined with quite a lot of pain/discomfort during the past 12 months. For people using the mouse ≥30 h per week, the risk of developing complaints appeared to be 8.4 times higher than for the reference group of people who used the mouse for 0-9 h per week. A recent review (Ijmker *et al*. 2007) shows an overview of studies about the relationship between duration of mouse use and the incidence of hand-arm and neck-shoulder complaints. From this review, a mean reference odds ratio for continued mouse use (≥30 h per week) can be estimated to be approximately 4. The results of this review only show an association between hand-arm symptoms and mouse use (and not for keyboard use and computer use) and therefore confirm the suggestion that duration of mouse use is an important factor in prevention of WRULS.

Another experimental study shows that micro breaks of several seconds can lead to less local discomfort in the upper extremities among computer workers (McLean *et al*. 2001). By implementing micro breaks, the total duration of mouse use and local discomfort in the upper extremities can thus be useful in slowing down or preventing the development of WRULS. These micro breaks should, however, not result in lower productivity in the primary work tasks.

5.1.3 Hovering behaviour

Often, the mouse is used unnecessarily during computer work, in other words, holding the hand on or above the mouse without using it. This behaviour is called 'hovering behaviour'. A negative side effect of hovering is the maintained extended wrist posture on or above the mouse (de Korte *et al*. 2005). If computer users could relax their arm and hand instead, this change in behaviour may result in less static load and more micro breaks for the upper extremity, allowing computer users to work in a healthier and more productive way (Henning 1997, Galinsky *et al*. 2000, Thé *et al*. 2003). Computer users should therefore have attention drawn to their 'hovering behaviour' so as to change it. One way to stimulate this change in behaviour is to offer a computer mouse that gives a feedback signal when hovering behaviour occurs too long without actively using the mouse.

In two laboratory studies, this theoretical mechanism was studied. Van de Ven and Ruijsendaal (2005) studied the effects of four different non-intrusive feedback signals on hovering behaviour. They concluded that a continuous tactile vibration signal decreased hovering behaviour. This signal was also preferred by users as the most suitable signal for a computer task. De Korte *et al*. (2005) concluded that hovering behaviour decreased by 36% when a mouse with a tactile feedback mechanism was used during a standardized task. Furthermore, a more dynamic muscle activation pattern was found. The mouse with tactile feedback did not show any effects on productivity. Usability was rated somewhat inconsistently.

5.1.4 Aim

The aim of the present study is to determine the effects of feedback on hovering behaviour, productivity, usability and comfort in a field study. Particularly, the aims are to see if the results of the laboratory studies will hold true in the field and to gain a better insight into how computer users get accustomed to the feedback, task effects and possible aversion to a new computer mouse during the initial phase of working with a feedback mechanism.

The main research questions of this study are:

› What is the effect of a mouse with a tactile, vibrating feedback signal compared to a regular mouse on:
  - hovering behaviour: duration and frequency of hand position on and above the mouse?
  - productivity: task performance and self-reported productivity?
  - discomfort: locally perceived discomfort per body region and subjective experience of using a mouse with a feedback signal?
  - usability and comfort: subjective experience of using a mouse with a feedback signal?

Since earlier experimental studies showed promising results, it is hypothesized that the use of a mouse with a tactile, vibration feedback signal will decrease hover time and increase relaxation of the forearm. Due to the decreased perceived exertion that is likely to result, productivity is expected to be the same.

5.2 Method

5.2.1 Participants

A total of 91 employees of a Dutch call centre participated in this study, 25 men and 66 women. All participants worked at least 3 d per week. At the onset of the study, participants gave informed consent.

5.2.2 Design and materials

In a randomized controlled trial, a regular computer mouse was compared to a new developed mouse with a tactile, vibrating feedback signal (Hoverstop mouse®; Hoverstop BV, Amsterdam, The Netherlands). This mouse is comparable with a regular mouse, except it is equipped with a small vibrating motor (comparable to vibration signals nowadays used in mobile phones). This motor starts to vibrate when a user keeps his or her hand on or above the mouse for more than 10 s without actively using the mouse (clicking or scrolling). The vibration signal gives feedback to the user in order to let go of the mouse and place the hand on the desk, resulting in a decreased duration of 'unnecessary' muscle tension. The motor vibrates at a 40% level of the motor power, with a maximum duration of 4 s. The Hoverstop mouse with the feedback signal will be referred to as the experimental mouse. A mouse with the same features as the experimental mouse but without the tactile, vibrating feedback signal will be referred to as the regular mouse. The mouse that the participants used before the experiment started will be referred to as the standard mouse.

5.2.3 Randomization

All participants used the regular mouse during the first week of the study to get used to the basic (external) features of the computer mouse. After the first week, all participants were assigned to the intervention group (46 participants) or the control group (45 participants). The procedure was as follows. All participants (working at least 3 d per week in both weeks) were alphabetically ordered and numbered 1 to 91. The odd numbers were assigned to the intervention group, the even numbers to the control group. The participants assigned to the intervention group were asked to work with the experimental mouse during the second week of the study. Participants in the control group kept working with the regular mouse. Five participants dropped out during the study, resulting in final groups of 45 and 41 participants respectively, see figure 1.

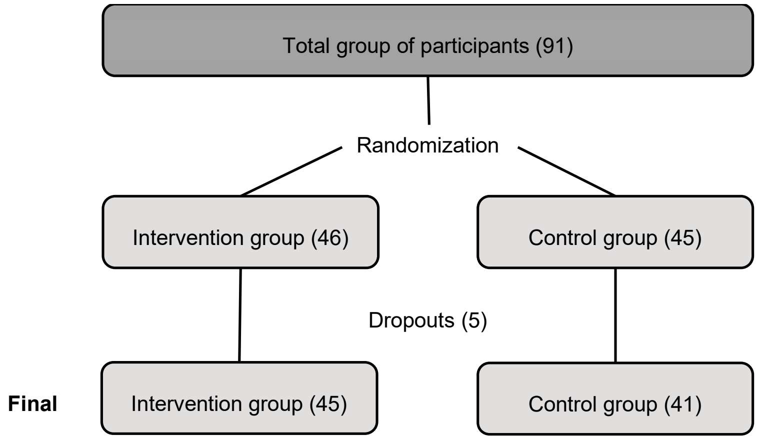

Figure 1 Illustration of the design of the study

5.2.4 Measurements

General

At the onset and at the end of the second week, all participants were asked to fill out a questionnaire. In the first questionnaire, general questions (such as age and gender), questions on working hours and questions on physical complaints were asked. At the end of the study, all participants were given a second questionnaire. The intervention group filled out extra questions specific about the experimental mouse and its effect on productivity, body part discomfort, comfort of using the experimental mouse, usability and level of distraction.

5

Hovering behaviour

A selection of 18 persons was observed during the first and second week to determine possible differences in hovering behaviour. In a laboratory study (De Korte *et al*. 2005), it was suggested that participants did not get fully used to the vibration feedback signal within a short duration. Therefore, all participants in the current study had been working with the experimental mouse for at least 3 d before they were observed. Observations were done using the software program The Observer (V5, Noldus Information Technology, Wageningen, The Netherlands) and were limited to the hand that operated the computer mouse. The position of the hand with regard to the mouse on and above the mouse, next to the mouse and ‘other’ (hands on the participants lap, holding a coffee cup) were observed. Participants were also asked in the questionnaire if they perceived a change in their own hovering behaviour.

Productivity

Reports on productivity of the call centre itself of the first and second week of the study were used to objectively determine the effect on productivity. Most important variables were 'talk time' (the mouse was particularly used during conversations to seek answers to customers' questions), 'mean handling time' (the mean time needed to answer a customer's call, seek a specialist's advice where needed and log the call into the system) and a third measure of productivity was calculated by dividing the total number of calls (incoming as well as outgoing) by the duration an agent was logged on. Since productivity in this profession is also dependent on customers' questions, productivity was also assessed subjectively by rating three items of self- reported productivity on a 5-point scale in the questionnaire. These three items are an estimation of the number of mistakes made during their work, work speed and number of tasks performed.

Discomfort

As a measure of discomfort, participants were asked to rate their perceived discomfort in 13 different body regions of the upper body at the start and end of their working day. These subjective ratings of localized musculoskeletal discomfort were obtained for the neck and both sides of shoulders, upper arms, elbows, fore-arms, wrists and hand/fingers. The Dutch validated method for body part discomfort was used (Lokaal Ervaren Ongemak (LEO), Van der Grinten and Smitt 1992), an ordinal scale ranging from no discomfort (0) to extreme discomfort (10). This scale is derived from the original Borg scale.

Comfort and usability

Comfort was rated on a 5-point scale. Comfort items included overall comfort of the experimental mouse, comfort of the feedback signal and duration of the feedback signal. Also, participants gave their opinion on 14 comfort items such as functionality, appearance and colour. Usability included items on use and learning experiences (5-point scale), distraction of the vibration signal, distraction of the desired behaviour, irritation of working with the experimental mouse (all visual analogue scales (0-10)) and which mouse was preferred in the future. Finally, some general questions on usability were asked. For the control group, the questions were only on the external features and general usability of the regular mouse (e.g. 'this computer mouse has a professional appearance', 'this computer mouse is of high quality', 'this computer mouse is easy to use' to be answered on a 5-point scale).

5.2.5 *Statistics*

Paired sample t-tests were performed to test the differences between the first and the second week in the variables collected during the observations. This was done for the intervention group as well as for the control group (α = 5%, two tailed). Variables tested were relative duration of hand on or above the mouse and hand next to the mouse, mean duration per occurrence that the hand was on or above the mouse and the rate of occurrences that the hand was on or above and next to the mouse.

Differences in the objective productivity data of the call centre between the intervention and the control group were tested using an independent sample t-test (α = 5%, two sided), differences in the three subjective productivity items and discomfort were tested using chi- square tests. Independent sample t-tests (α = 5%, two sided) were used to test differences in the discomfort items between the intervention and the control group and differences in discomfort within groups between the first and second week were tested with paired t-tests (α = 5%, two sided).

In the questionnaire, the degree of overall comfort was measured by ten items. A reliability analysis (Cronbach's alpha) has been performed to test if all items are related to each other. A regression analysis was subsequently performed to study the direct effects of four independent variables on comfort. To test the difference in preference of which mouse they would like to use in the future between the intervention group and the control group, a Mann-Whitney test was used (α = 5%, two sided).

5.3 Results

5.3.1 *General*

Participant characteristics of the intervention group, control group and total group are given in table 1. No significant differences for one of the variables were found; the randomization was apparently successful.

Table 1 Distribution of self-reported mean computer use (days per week/hours per day) and the use of input devices, for intervention, control and total group*

	Intervention group (n = 45)	Control group (n = 41)	Total (n = 86)
Age	38 (8.6)	37 (11)	38 (9.8)
Gender (% men)	28.9	22.0	25.6
Working days/week (%)			
3 d/week	20.0	36.6	27.9
4 d/week	33.3	26.8	30.2
5 d/week	37.8	36.6	37.2
>5 d/week	8.9	0.0	4.7
Computer us at work (%)			
2-4 h	2.2	0.0	1.2
4-6 h	0.0	7.3	3.5
>6 h	97.8	92.7	95.3
Use of input devices (%)			
Keyboard	43.4	40.4	42.0
Mouse	52.4	55.3	53.7
Other	4.2	4.3	4.2

* No significant differences between the two groups were found.

5.3.2 *Response*

A total of 37 participants of the intervention group (82%) and 23 participants of the control group (56%) responded to the second questionnaire. Response to the first questionnaire was 100% for both groups.

5.3.3 *Hovering behaviour*

Objective results

The total number of observed participants was 18, with a duration of 20 min. For 13 participants, observations were done during the first as well as the second week (six participants of the intervention group and seven of the control group). An overview of the objective results of hovering behaviour is presented in table 2. As hypothesized, the total duration of hovering behaviour, expressed as the relative duration during the observation, decreased in the intervention group in the second week compared to the first week, from 62 to 47% (mean difference of 14%, $p < 0.04$). In the control group, the mean difference was 5%, from 47 to 42% ($p < 0.6$).

Table 2 Objective hovering behaviour from observations

	Intervention group				Control group			
	First week	Second week	Mean difference	p	First week	Second week	Mean difference	p
Duration hand on or above mouse (%)	62	47	-14	037	47	42	-5	0.523
Duration hand next to mouse (%)	2	6	4	0.106	0	1	1	0.045
Mean duration per occurrence (s)	27	14	-13	058	17	14	-3	0.288
Rate hand on or above mouse (number per min)	1.5	2.2	0.7	0.25	1.6	2.0	0.4	0.078

The duration of periods that the hand was next to the mouse, in the second week (6%) compared to the first week (2%) was not significantly different in the intervention group ($p < 0.2$), while a significant increase (from 0% to 1%) was found for the control group ($p < 0.05$).

Considering the mean duration per occurrence that the hand was above the mouse, this duration decreased from 27 to 14 s in the intervention group. However, this difference was borderline significant: $p < 0.06$. In the control group, the mean duration of hand above the mouse did not change significantly either between the first and the second week (from 17 to 14 s, $p < 0.3$). Finally, the rate (number per min) of the occurrences that the hand was above the mouse significantly increased in the intervention group ($p < 0.03$) and did not change in the control group ($p < 0.08$). The rate of occurrences that the hand was next to the mouse did not change significantly in the intervention group ($p < 0.2$) or in the control group (t-test not applicable).

5

Subjective results

Three questions in the questionnaire of the intervention group were about hovering behaviour and the vibration signal. These questions were only asked to the intervention group, since the control group did not experience the vibration signal. Figure 2 shows the distribution of the self-reported hovering behaviour.

According to the responses of 37 volunteers who participated in the intervention group and returned a completed second questionnaire, 78.4% indicated that they often or always took their hand off the mouse when it began to vibrate.

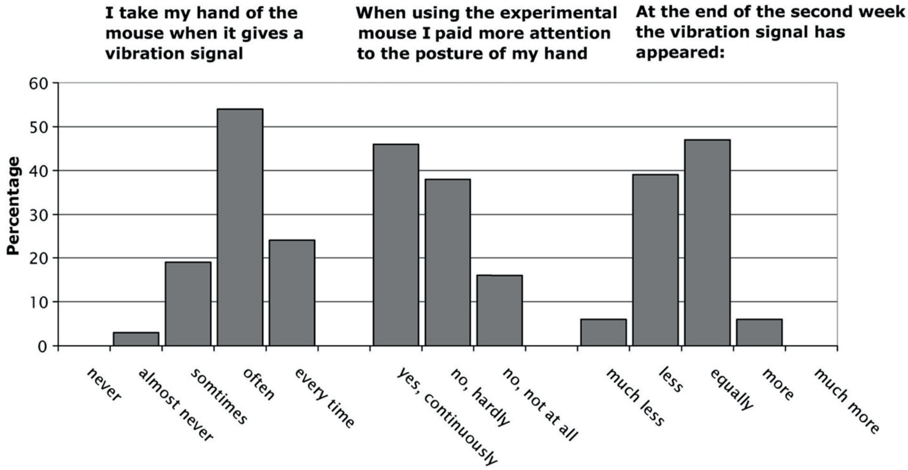

Figure 2 Distribution of self-reported hovering behaviour of the intervention group, n = 37

Of all participants, 45.9% indicated that they paid more attention to their hand posture during computer work and 91.7% of the participants reported that the vibrating signal had appeared equally often or less often at the end of the second week. Almost 45% indicated a decreasing appearance of the vibration signal at the end of the second week.

This means that participants take their hand more often off the mouse spontaneously and thus that the mouse functioned as was anticipated. This is in accordance with the earlier- mentioned objective results of the intervention group, in which it was found that the rate (number per min) of the occurrences that the hand was on the mouse significantly increased, while the total duration decreased.

5.3.4 Productivity

Objective results

The reports of the productivity of the call centre showed that the second week was slightly busier than the first week (18.1 calls/login duration and 20.3 calls/login duration respectively). For both the intervention group and the control group, the difference in calls per login duration was significant between the first and second week (p = 0.000 and p = 0.000). The differences between the intervention and control group in the degree of increase or decrease in the second week relative to the first week were tested. An overview of the objective productivity is given in table 3.

Table 3 Objective productivity of the intervention and control group during the first and second week

	Intervention group		Control group		Total group		
	First week	Second week	First week	Second week	First week	Second week	*p**
Calls per login time (call/h)	18.6	20.5	17.3	19.5	18.1	20.3	0.114
Talk time (s)	96	99	96	99	96	99	0.543
Mean handling time (s)	132.0	137.0	132.0	136.0	132.0	137.0	0.947

* *p*-value gives the difference between the intervention and control group in the degree of increase between the first and second week

The increase in talk time did not significantly differ between the intervention and control group ($p = 0.543$), neither did the increase in 'mean handling time' ($p = 0.947$). Consistent with these results, the increase in the number of calls per login duration was also not significantly different between the intervention and control group ($p = 0.114$). Overall, the experimental mouse appeared to have no unfavourable effects on objective productivity.

Subjective results

Self-reported productivity (figure 3) showed that 70.0% of the participants in the intervention group mentioned that they make as many or less mistakes with the experimental mouse compared to the standard mouse. In contrast, 87.5% of the control group mentioned that they make as many or less mistakes with the regular mouse compared with the standard mouse. This difference was not significant ($p = 0.492$). For working speed, 45.9% in the intervention group mentioned that they worked slower with the experimental mouse. However, in the control group, 33.3% reported that they worked slower with the regular mouse. The difference between the intervention and control group was not significant ($p = 0.421$). For the number of tasks the participants could do with the experimental mouse, none of the respondents in either group reported that they could perform more tasks. In the intervention group, 34.1% mentioned that they could perform less tasks, in the control group 8.3% reported that they could perform less tasks. This difference was also not significant ($p = 0.087$).

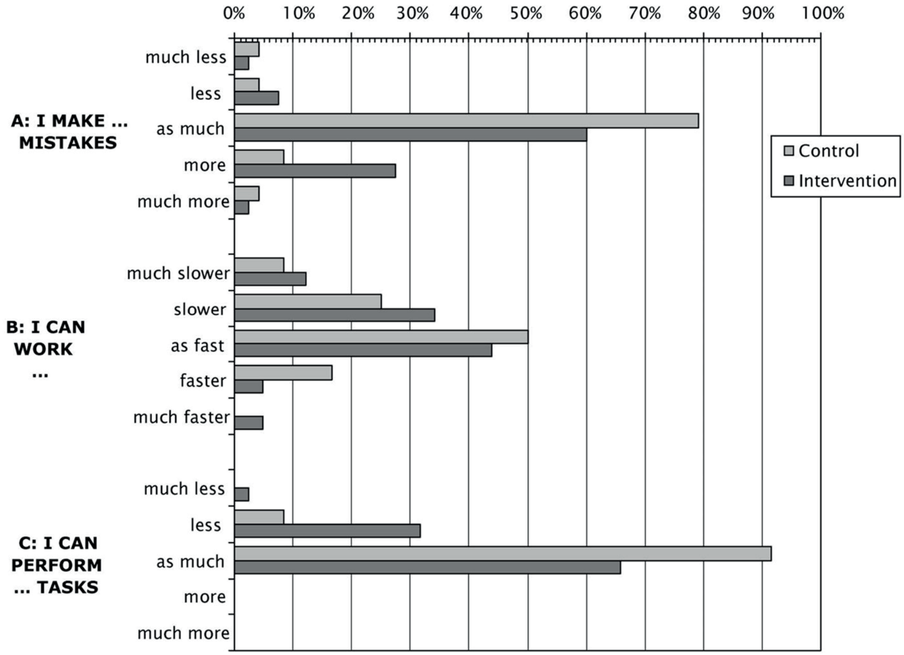

Figure 3 Subjective results on productivity for the intervention and the control group. Subjects answered whether they made less/more mistakes (part A), whether they could work slower/faster (part B) and whether they could perform less/more tasks (part C) when working with the experimental mouse (intervention group) or regular mouse (control group) compared to working with the standard mouse

5.3.5 Discomfort

Discomfort questionnaire

Self-reported discomfort in the arm and shoulder reported in the questionnaire showed that 58.3% of the participants mentioned they perceived the same level of discomfort or less discomfort with the experimental mouse. For fitness, 78.4% indicated there was no difference in fitness at the end of the day between working with the experimental mouse and working with the standard mouse (figure 4).

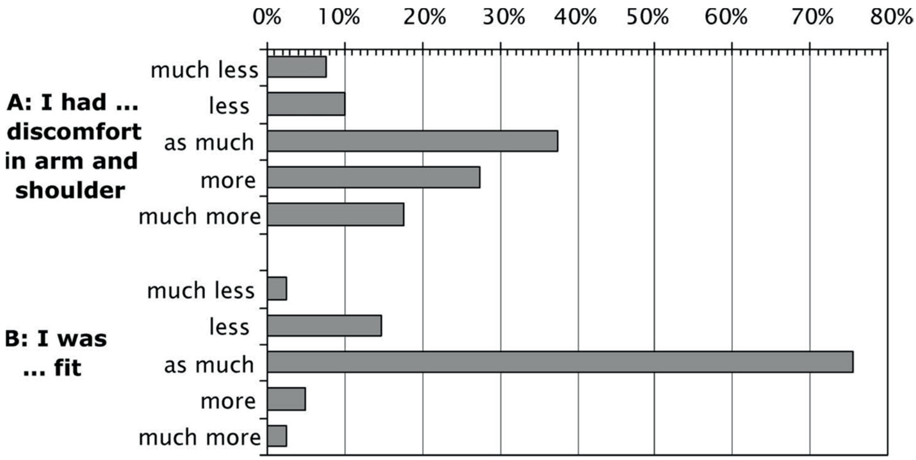

Figure 4 Subjective results on discomfort in the intervention group while working with the experimental mouse compared to working with the standard mouse

Local discomfort

The subjective ratings of localized discomfort per body region (LEO) did not show any significant differences between the control group and the intervention group.

5.3.6 Usability and comfort

The self-reported overall comfort of the experimental mouse and the regular mouse reported in the questionnaire varied considerably between participants. Examples of these items are the quality of the mouse, functionality and how well it fits the hand. All items were scored on a 5-point scale. Overall comfort was calculated by averaging all 14 comfort items (the vibration signal was explicitly left out in this analysis). The mean overall comfort score was 3.47 (SD 0.78; n = 61). There is no significant difference between the intervention and the control group in rating the overall comfort of the mouse.

Cronbach's alpha scores were calculated and appeared to be high for both the intervention group (α = 0.850) and the control group (α = 0.908). A multiple regression analysis has subsequently been performed to examine the dependence of overall comfort on group (experimental or control), on age, on the level of education and on WRULS. From this analysis only the degree of WRULS has a significant influence on the level of comfort. This means that those persons with neck and upper limb symptoms have a more positive attitude towards the overall comfort of the experimental mouse. According to this analysis, the level of comfort is not significantly dependent on the group and, thus, the presence of the vibrating feedback signal.

In addition to the comfort items, the questionnaire also contained some questions with regard to usability. Both groups were asked if they were satisfied with working with the new mouse. The mean score on a visual analogue scale (0 = totally not satisfied and 10 = totally satisfied) was 5.9 (SD 2.2; n = 24) for the control group, the mean score of the intervention group was 4.2 (SD 3.4; n = 37). The distribution of the scores appeared to be much larger in the intervention group. In fact, it seemed as if participants tend to judge the experimental mouse either very well or quite bad, with a relatively low number of participants with a neutral opinion.

In addition, the participants of the intervention group were asked to give their opinion about the difficulty of learning to work with the experimental mouse, the level of distraction of the vibrating signal, the level of distraction of the desired behaviour and the level of irritation of working with the experimental mouse. The mean scores are shown in table 4. To all scores, the following applies: '0' = I totally do not agree; '10' = I totally agree.

Table 4 The means, medians and standard deviations of self-reported scores on usability*

n = 37	Mean	Median	SD
It is difficult to learn to work with the experimental mouse	3.9	3.5	3.4
The vibration signal is distracting	6.4	7.6	3.6
The desired behaviour is distracting	5.8	7.2	3.2
Working with the experimental mouse is irritating	6.4	6.9	3.3

* Scores are based on a visual analogue scale (0 = I totally do not agree; 10 = I totally agree).

Finally, all participants were asked if they would like to continue to work with the experimental mouse in the future. In the intervention group, 29.7% of the participants preferred using the experimental mouse. In the control group, however, this percentage was 2.5 times higher; 77.3% preferred to continue to work with the regular mouse. The difference between these groups was significant ($p < 0.001$). Although the percentage of participants who prefer to use the experimental mouse in the future, after just 1 week, is satisfactory, the difference with the control group is quite large.

5.4 Discussion

The aim of the present study was to determine the effects of feedback on hovering behaviour, productivity, usability and comfort in a field study; in particular, to see if the results of previous laboratory studies (De Korte *et al*. 2005, van de Ven and Ruijsendaal 2005) will hold true in the field. Another aim was to gain better insight into how computer users get accustomed to the feedback, task effects and possible aversion to a new computer mouse during the initial phase of working with a feedback mechanism. According to the current study, the use of a vibrating feedback signal leads to decreased duration of the hand on or above the mouse. It also shows that participants react to the signal by moving their hand away from the mouse. Participants mentioned that the vibration signal decreased at the end of the second week. It was cautiously concluded that the reaction to the vibration feedback signal changed in time into a healthier work behaviour. Participants changed their mouse operating behaviour from static to more dynamic behaviour.

5.4.1 Hovering behaviour

The total number of observed participants in both weeks was 13, although a total of 20 participants was intended. Reasons for losing a number of observations were absence from work (sickness) and changes in work schedules in the second week. However, the observations were performed on the same weekday in the first and second week and therefore were considered highly reliable.

Based on the results of two laboratory studies (De Korte *et al*. 2005, Van de Ven and Ruijsendaal 2005), it was hypothesized that hover time would decrease and relaxation of the forearm (in the current study defined as holding the hand next to the mouse) would increase when using the experimental mouse. Total hover time did indeed decrease in the intervention group in the second week. During the observations, the total duration of hovering behaviour decreased in the intervention group in the second week compared to the first week (p = 0.037, mean difference 14%). This result was consistent with the results from the questionnaires. In the questionnaires, 78.4% of the participants mentioned they had removed their hand ‘often’ or ‘always’ away from the mouse when it began to vibrate. The subjective data are thus in line with the objective data. The data are also comparable to the studies of De Korte *et al*. (2005) and Van de Ven and Ruijsendaal (2005). This means that objective and subjective data on hovering behaviour indicate that hover time decreases when the experimental mouse is used.

De Korte *et al*. (2005) found that the rate (number per min) of occurrences that the hand was on the mouse increased when using the experimental mouse. The observations showed that the rate of occurrences that the hand was on or above the mouse significantly increased in the intervention group and did not change in the control group. In the questionnaire, 44.5% of the participants from the intervention group mentioned that the vibration signal appeared less often at the end of the second week. The results above imply that, at the end of the second week, participants in the intervention group took their hands off the mouse before the vibration feedback signal actually appeared. Therefore, they decreased the total duration that they held the mouse unnecessarily and showed less hovering behaviour. A change from reacting to the vibration signal to an automatic healthier work behaviour seems to develop.

The decrease in hover time (14%) appeared to be slightly lower in this field study compared to earlier laboratory studies in which the difference in total hover duration between the regular and experimental mouse was 16% and 26%, respectively (De Korte *et al*. 2005, Van de Ven and Ruijsendaal 2005). A possible explanation for this difference is that participants are more willing to focus on their hovering behaviour during a laboratory experiment (and the tasks that are performed during the experiments) than during their own primary work tasks. Also, the standardized task that the participants performed in the laboratory studies is not comparable to the call centre task in the field study. However, the difference in total hover duration found in the intervention group of the field is significant. This is a relevant decrease. For example, when extrapolating to a complete working day with a total duration of 4 h of mouse use, this would decrease hovering behaviour with a mean of 36 min per day. The duration of time that the hand was next to the mouse did not increase significantly in the intervention group, while it did increase significantly in the control group. An important explanation for the fact that this small increase in the control group (mean increase less than 1%) is still significant, is that this behaviour, 'hand is next to the mouse', did not occur at all during the first week. This increase is therefore statistically significant, but not considered relevant.

The decrease in duration per occurrence of hand on or above the mouse (13 s) for the intervention group was not found to be significant ($p = 0.058$). Since the p-value is close to 0.05, in combination with the low number of participants and the finding that this duration decreased for all six participants, it is cautiously concluded that the decrease in duration per occurrence of hand above the mouse will be significant if a greater number of participants are tested.

Finally, the rate (number per min) of occurrences that the hand was on or above the mouse increased significantly in the intervention group. This finding corresponds with an earlier laboratory study, in which it was concluded that computer users tend to change their mouse operation behaviour from static into more dynamic (De Korte *et al*. 2005).

5.4.2 Productivity

From the literature, it appears that there are indications that short breaks have no negative (and in some studies a positive) effect on productivity. However, most studies have been performed under laboratory conditions, with a simulated task and with students or temporary personnel. However, the current study is a field study in which participants performed their regular daily tasks. The conclusion that no differences were found in objective productivity between the intervention and control group can therefore be considered as a positive result.

The results for self-reported productivity showed that there are subgroups in the intervention group (30.0%) as well as in the control group (12.5%) who believed they made more mistakes with respectively the experimental mouse and the regular mouse. It is not possible to attribute this to the vibration signal, since the answers to the questions on self-reported productivity were not significantly different between the intervention and control group. A possible reason is the special layer (soft top tables) on their desktop. The purpose of this layer is noise reduction, but a lot of participants mentioned that the mouse did not slide well on their table and did not always react accurately.

5.4.3 Discomfort

The results of the measurements of the localized discomfort per body region showed no significant effects. This is most likely due to the low intensity of computer work. This result corresponded with the results found in an earlier laboratory study (De Korte *et al*. 2005). Swanson *et al*. (1997) mentioned that levels of discomfort and fatigue for keyboard work were quite low. This also confirms that computer work is low in intensity and therefore it is hard to measure intensity changes within computer workers. Swanson *et al*. (1997) and Haward (1998) also found that there is minimal impact of changing the input device on self-reported body part discomfort and fatigue.

5.4.4 Usability and comfort

The measurement of comfort of the experimental and the regular mouse showed that there is no significant difference in rating the overall comfort of the mouse between the two groups. This suggests that further differences found between the intervention group and the control group could be attributed to the vibrating signal.

Furthermore, the dependence on the vibration signal, age, level of education and the presence of neck, shoulder and arm symptoms on overall comfort was tested. Presence of neck, shoulder and arm symptoms was defined as the self-reported number of regions with work-related discomfort for at least 4 weeks during the past 6 months. This analysis showed that only the number of regions with reported symptoms was of significant influence on comfort, meaning that participants with more regions with discomfort judged the mouse as more comfortable. The other variables did not influence the overall comfort rating. An important conclusion is that the vibration signal does not significantly influence the rating of comfort.

Furthermore, the intervention group answered four extra questions about usability. The results of these questions showed that there is variation between the participants. There is no clear consensus about liking or disliking the mouse. The answers to the question about continuing to work with the experimental or the regular mouse in the future showed that without the vibrating signal (control group) a more substantial group of the participants chose the new mouse. With the vibration signal (intervention group) a smaller group of participants preferred to continue to work with the experimental mouse. It seems that participants without neck, shoulder and arm symptoms do not judge the vibration signal as useful, while participants with (starting) symptoms judge the vibration signal as more useful and therefore favourable. The vibration signal is meant to build in more relaxation to prevent computer workers from developing neck, shoulder and arm symptoms. It seems that participants without symptoms feel less need to take action to reduce risks for developing neck, shoulder and arm symptoms. This corresponds with the suggestion that Haward (1998) made that participants without musculoskeletal disorders had no reason to modify their work behaviour to cope with any discomfort and pain. Also, Heinrich *et al*. (2005) found that employees with neck, shoulder and arm complaints have more need for preventive measures than employees without complaints.

5.4.5 Recommendation

In the present study, call centre agents were studied to determine the effect of a vibration feedback signal on hovering behaviour. In future studies, it is recommended that the influence of different job types and job tasks is examined. Furthermore, the long-term effects on (prevention of) neck, shoulder and arm symptoms of using a mouse with a feedback signal to stimulate healthy behaviour should be studied in a randomized control trial. The study of long-term development of learning effects is also recommended.

Acknowledgements

We thank the Dutch Ministry of Economic Affairs, Hoverstop BV and ABN AMRO for supporting this study.

5.5 References

BLATTER, B.M., HOUTMAN, I., VAN DEN BOSSCHE, S., KRAAN, K. and VAN DEN HEUVEL, S., 2006, Gezondheidsschade en kosten door RSI en psychosociale arbeidsbelasting. Werkdocumenten nummer 358 [In Dutch]. (Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid).

BRANDT, L.P.A., ANDERSEN, J.H., LASSEN, C.F., KRYGER, A., OVERGAARD, E., VILSTRUP, I. and MIKKELSEN, S., 2004, Neck and shoulder symptoms and disorders among Danish computer workers. *Scandinavian Journal of Work and Environmental Health,* **30,** 399-409.

COOK, C.J. and KOTHIYAL, K., 1998, Influence of mouse position on muscular activity in the neck, shoulder and arm in computer users. *Applied Ergonomics,* **29,** 439-443.

DE KORTE, E.M., DE KRAKER, H., BONGERS, P. and VAN LINGEN, P., 2005, Effects of a feedback signal in a computer mouse on behaviour, muscle load, productivity, comfort and user friendliness. *Ergonomics* (submitted).

GALINSKY, T.L., SWANSON, N.G., SAUTER, S.L., HURREL, J.J. and SCHLEIFER, L.M., 2000, A field study of supplementary rest breaks for data-entry operators. *Ergonomics,* **43,** 622-638.

HAWARD, B., 1998, An evaluation of a trackball as an ergonomics intervention. In *Contemporary Ergonomics,* M.A. Hanson (Ed.), pp. 135-139 (London: Taylor and Francis).

HEINRICH, J., BLATTER, B.M., VAN DEN BOSSCHE, S.N.J. and SMULDERS, P.G.W., 2005, RSI-maatregelen in de Nederlandse beroepsbevolking [in Dutch]. *Tijdschrift voor Bedrijfs- en Verzekeringsgeneeskunde,* **13,** 69-71.

HEINDRICH, J. and BLATTER, B.M., 2005, RSI-klachten in de Nederlandse beroepsbevolking. Trends, risicofactoren en verklaringen [in Dutch]. *Tijdschrift voor Gezondheidswetenschappen,* **83,** 1-16.

HENNING, R., 1997, Frequent short rest breaks from computer work: effects on productivity and well being at two field sites. *Ergonomics,* **40,** 78-91.

IJMKER, S., HUYSMANS, M., BLATTER, B.M., VAN DER BEEK, A.J., VAN MECHELEN, W. and BONGERS, P.M., 2007, Should office workers spend fewer hours at their computer? A systematic review of the literature. *Occupational and Environmental Medicine,* **64,** 211-222.

KRYGER, A.I., ANDERSEN, J.H., LASSEN, C.F., BRANDT, L.P., VILSTRUP, I., OVERGAARD, E., THOMSEN, J.F. and MIKKELSEN, S., 2003, Does computer use pose an occupational hazard for forearm pain; from the NUDATA study. *Occupational and Environmental Medicine,* **60,** article no. e14.

LASSEN, C.F., MIKKELSEN, S., KRYGER, A.I., BRANDT, L.P.A., OVERGAARD, E., THOMSEN, J.F., VILSTRUP, I. and ANDERSEN, J.H., 2004, Elbow and wrist/hand symptoms among 6,943 computer operators: a 1-year follow-up study (the NUDATA study). *American Journal of Industrial Medicine,* **46,** 521-533.

MCLEAN, L., TINGLEY, M., SCOTT, R.N. and RICKARDS, J., 2001, Computer terminal work and the benefit of microbreaks. *Applied Ergonomics,* **32,** 225-237.

PASCARELLI, E. and KELLA, J., 1993, Soft-tissue injuries related to use of the computer keyboard. *Journal of Occupational Medicine,* **35,** 522-532.

REMPEL, D.M., HARRISON, R.J. and BARNHART, S., 1992, Work-related cumulative trauma disorders of the upper extremity. *Journal of the American Medical Association,* **267,** 838-842.

SWANSON, N.G., GALINSKY, T.L., COLE, L.L., PAN, C.S. and SAUTER, S.L., 1997, The impact of keyboard design on comfort and productivity in a text-entry task. *Applied Ergonomics,* **28,** 9-17.

THÉ, K.H., HEUVEL, S.M., VAN DEN WIEZER, N.M. and BONGERS, P.M., 2003, Behavioural aspects of developing or sustaining neck and upper limb disorders. In Ergonomics in the Digital Age.

Proceedings of the XVth Triennial Congress of the International Ergonomics Association, 24-29 August 2003, Seoul, Korea.

VAN DEN BOSSCHE, S.N.J. and SMULDERS, P.G.W., 2004, De Nationale Enquête Arbeidsomstandigheden 2003: Methodologie en globale resultaten (Hoofddorp: TNO Arbeid).

VAN DER GRINTEN, M.P. and SMITT, P., 1992, Development of a practical method for measuring body part discomfort. In Ad*vances in Industrial Ergonomics and Safety IV,* S. Kumar (Ed.), pp. 311-318 (London: Taylor & Francis).

VAN DE VEN, J.G.M. and RUIJSENDAAL, M., 2005, *Studying the Effect of Non-Intrusive Feedback on Mouse Hovering Behaviour*. Report no. TNO-DV3 2005-C 019 Confidential (Soesterberg: TNO Defence, Security and Safety).

WAHLSTROM, J., SVENSSON, J., HAGBERG, M. and JOHNSON, P.W., 2000, Differences between work methods and gender in computer mouse use. *Scandinavian Journal of Work, Environment and Health,* **26,** 390-397.

6

Effects of four types of non-obtrusive feedback: laboratory experiment[1]

[1] Published as: De Korte, E.M., Huysmans, M.A., De Jong, A.M., Van de Ven, J.G.M. & Ruijsendaal,M.. (2012). Effects of four types of non-obtrusive feedback on computer behaviour, task performance and comfort. *Applied Ergonomics, 43*, 243-353.

Elsbeth M. de Korte[a,b]
Maaike A. Huysmans[c,d]
Annelise M. de Jong[b]
Josine G.M. van de Ven[e]
Mark Ruijsendaal[f]

[a] TNO Work & Employment, P.O. Box 718, 2130 AS Hoofddorp, The Netherlands
[b] Faculty of Industrial Design Engineering, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands
[c] EMGO Institute for Health and Care Research, VU University Medical Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands
[d] Body@Work, Research Center on Physical Activity, Work and Health TNO-VU/VUmc, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands
[e] TNO Defense, Safety and Security, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands
[f] TNO Information Society, P.O. Box 5050, 2600 GB Delft, The Netherlands

Abstract

This study investigated the effects of non-obtrusive feedback on continuous lifted hand/finger behaviour, task performance and comfort. In an experiment with 24 participants the effects of two visual and two tactile feedback signals were compared to a no-feedback condition in a computer task. Results from the objective measures showed that all types of feedback were equally effective to reduce lifted hand/finger behaviour (effectiveness) compared to absence of feedback, while task performance was not affected (efficiency). In contrast to objective measures, subjective user experience was significantly different for the four types of feedback signals. Continuous tactile feedback appeared to be the best signal; not only the effectiveness and efficiency were rated reasonable, it also scored best on perceived match between signal and required action. This study shows the importance of including user experiences when investigating usability of feedback signals. Non-obtrusive feedback embedded in products and environments may successfully be used to support office workers to adopt healthy, productive and comfortable working behaviour.

Keywords: Human computer interaction; Feedback; Behaviour; Comfort; Task performance; Usability.

6.1 Introduction

A field of research getting much attention in Human Computer Interaction is Ambient Intelligence (AmI), which refers to electronic systems embedded in our everyday environments and are sensitive and responsive to people in a seamless, unobtrusive, and often invisible way (Aarts and De Ruyter, 2009; Cook and Song, 2009) The emphasis of AmI is on greater user-friendliness, more efficient services support, user-empowerment and support for human interactions (Ducatel et al., 2001).

In the field of office ergonomics, AmI can play an important role to solve the question how to motivate and stimulate office workers to adopt healthy, productive and comfortable work styles (Kuijt-Evers and Steen, 2008) by offering new possibilities to this end. In the context of AmI, Aarts and De Ruyter (2009) identified the issues of 'suggestion' (reminding people to perform certain behaviour at opportune moments) and 'self-monitoring' (allows people to monitor themselves and to inform them about how they could modify their behaviour). The challenge is to develop new approaches for the office environment that can monitor compliance and trigger persuasive interventions. Preferably, these approaches should be intuitive, with technology non-obtrusively integrated in everyday objects and environments.

Examples of existing intelligent products for the office environment are the Multi-modal Mouse with tactile feedback by means of a pin through a hole in de left mouse button and forcefeedback using an electromagnet (Akamatsu et al., 1995) and the FeelIt Mouse (Immersion Corporation, San Jose, CA), a haptic force-feedback computer mouse (Dennerlein and Yang, 2001), all aimed at optimizing task performance.

Examples of products aimed at changing unhealthy behaviour of office workers are rest break software (Slijper et al., 2007; Van den Heuvel et al., 2003) and the Hoverstop Mouse© (Hoverstop BV, Amsterdam, Netherlands) (De Korte et al., 2008; De Kraker et al., 2008). Rest break software introduces rest breaks and sometimes exercises when computer use has been too intensive or prolonged. The Hoverstop Mouse is designed to reduce sustained muscle tension during mouse use by giving tactile feedback.

Whether an intelligent product will be successful depends, among other things, on the feedback method. In order to interact with it in a natural way, it is important to develop an optimal feedback method to indicate situations where the user should change his or her working behaviour. To ensure that a feedback signal has a minimally disturbing effect on the attention to the primary task, the feedback has to pass almost unnoticed by the user in the attentional background, enabling the user to communicate at an unconscious level. (Fairclough, 2009; Oakley et al., 2000; Raisamo et al., 2004). Unfortunately, little is known about the efficacy of implicit, non-obtrusive feedback signals in order to provoke changes in user states (Fairclough, 2009).

Furthermore, the user has to know how and why certain feedback is presented (correct mental model, the meaning of the signal has to be understood) and preferably has to be able to personalize the feedback (Keyson, 2007; Van der Veer and Puerta Melguizo, 2002). In addition, it is important that the feedback signal is easy to learn, that there is a minimal chance of misinterpretation and that the signal is not disruptive for the environment (especially in open office spaces).

In the present study, we want to examine the application of AmI in the office setting and investigate the use of different types of non-obtrusive feedback signals in order to change unhealthy behaviour of office workers. During mouse use, office workers often maintain a prolonged lifted hand or finger posture, even when they are not actively using the mouse, e.g. during stationary activities like gazing at the computer screen (Lee, 2006). During this unnecessary, sustained lifted hand and finger behaviour, muscles remain activated for a long period of time, which is hypothesized to be a risk factor for developing repetitive strain injury (RSI) complaints, particularly when this awkward posture is sustained for a long period of time (Jensen et al., 1998; Lee, 2006). When computer users are provided with feedback when this unnecessary lifted hand and finger behaviour takes place, they can choose to relax the hand instead, and thus are likely to lower their risk to develop RSI. As mentioned above, the type of feedback may determine to a large extent how successful the AmI is in bringing about the intended behaviour.

In an expert session, with six experts in the fields of cognitive psychology, environmental and industrial hygiene, usability, product design, ergonomics and human computer interaction, four feedback signals were selected for setting up an experiment: two tactile signals (a continuous vibration signal and an on-off-on vibration signal in a computer mouse) and two visual signals (a peripheral signal and full screen transparent signal visible on the computer screen).

Both auditory and thermal feedback are considered unsuitable to be used in an office environment. Auditory feedback involves the risk of disturbing employees nearby, especially in open office spaces (Akamatsu et al., 1995). Thermal feedback seems unsuitable because the response of the signal may be too slow and a fast return to the starting temperature is not possible.

Visual feedback was decided to be included in the experiment, because it may be used for more complex messages. The most common type of non-obtrusive visual feedback is the use of *peripheral displays* (Matthews et al., 2007), with which visual feedback is placed in a *specific part* of the screen not central to a person's current task. It requires minimal attention and cognitive effort and allows a person to be aware of extra information without being overburdened by it (Matthews et al., 2007). A peripheral display can appear as a small part of the screen near the border of the monitor or as an extra monitor placed in the peripheral view of the user. An example of a peripheral display is the system tray in the Windows task bar. It contains abstract icons about the programs active on your machine that you do not use for your primary task, for example the connection to the network.

However, display characteristics, like brightness and resolution, may also be informative channels when looking at a screen. These characteristics are usually *full screen*, but allow you to continue your work. Even though we have not found any information on this type of feedback in the literature, we liked to include full screen visual feedback in our experiment. The advantage of this type of feedback is that users do not have to focus their attention on a specific place on the screen because the information is everywhere on the screen.

Tactile feedback was included in the experiment, because it appears to be a promising type of feedback (Akamatsu et al., 1995; Hopp et al., 2005; Jones and Sarter, 2008; Van Erp and Verschoor, 2004; Vitense et al., 2003). It has a tremendous potential to contribute to effective interaction when the visual and or auditory modalities are compromised, engaged or overwhelmed (Vitense et al., 2003). Computer work is characterized by a strong dependence on the visual channel, which may cause visual fatigue, or may necessitate directing too much attention to the computer screen (Akamatsu et al., 1995). Several potential advantages of the tactile modality over the visual and auditory modality have been identified: for example, the tactile channel is not heavily used, tactile cues can be readily detected, a limited vocabulary can be transmitted, the cues are not highly obtrusive, tactile information can be omnidirectional, and tactile information can be perceived along with additional visual and auditory information (Hopp et al., 2005). For example, it is possible to develop a tactile vocabulary to communicate simple and more abstract messages by varying the rhythm of a tactile signal (Jones and Sarter, 2008). However, a tactile signal that consists of one pattern always indicating the same action may be easier to perceive unconsciously, than a tactile signal that changes rhythms to indicate different actions. Results from the experiments of Hopp et al. (2005) showed that tactile cues allowed participants to effectively direct attention where needed, without disrupting ongoing information processing. However, Raisamo et al. (2004) pointed out that it is important that the tactile signal matches the task and the signal strength matches the environment in which it is used.

The research questions of this paper are:

1. What are the effects of non-obtrusive feedback signals on lifted hand and finger behaviour?
2. How are these four non-obtrusive feedback signals perceived by the users and which feedback signal is preferred?

Following ISO 9241 (ISO, 1998), the results of this study will be discussed in terms of usability, which consists of 3 factors: effectiveness, efficiency and satisfaction. First of all, it is presumed that the use of feedback results in the intended behavioural changes, i.e. the signal will be noticed and participants will be able to react to it with the required actions, in this case taking the hand off the mouse (effectiveness). Furthermore, it is presumed that the total time spent on a task and the number of errors made are equal to a situation without a signal (efficiency). Finally, comfort, which is influenced by physical factors and experience (De Looze et al., 2003; Vink et al., in press), and acceptability of the feedback signals need to be considered as well (satisfaction).

It is expected that the on-off-on signal will be better noticed and thereby will be more effective than the continuous feedback signal. However, it may be rated more uncomfortable than the continuous signal, because it is likely to be more distracting. In case users immediately take action when noticing the continuous signal, there will be no effect between the two signals and both will be just as effective.

It is expected that the full screen transparent signal will be most effective due to its appearance and noticeability. However, it may be rated inefficient due to interference with task performance. The peripheral signal is expected to be both effective and efficient.

6.2 Material and methods

In an experiment, four different feedback signals (two tactile and two visual signals) and a control condition without a feedback signal were evaluated by 24 participants. They performed a standardized computer task with two levels of task difficulty.

6.2.1 General design

Twenty-four subjects (6 female, 18 male) participated in this experiment. Their ages ranged from 19 to 37 years (mean 24 years, SD 5 years). All participants were required to have received higher general secondary education, no musculoskeletal complaints of the upper extremity and to have used the computer during the last 2 years and for more than 2 h a day. They were monetarily compensated (40,- Euro) for their participation. Before the start of the experiment, the participants signed an informed consent.

6.2.2 Material

Feedback signals
Two tactile signals and two visual signals were included in the experiment.

- *Tactile feedback*. The first tactile signal was developed by Hoverstop BV, as it is used in the commercially available Hoverstop mouse (Hoverstop mouse©, Hoverstop BV, Amsterdam, The Netherlands), originally designed to reduce sustained lifted hand and finger behaviour during active mouse use.
  This two-button mouse with scroll wheel was comparable to a regular mouse, except that it was equipped with a capacitive sensor capable of detecting the hand being in close proximity to the mouse through changes in the electric environment, and with a small vibrating actuator. Mouse action (using buttons or scroll wheel) was monitored. The feedback signal started when the buttons and/or scroll wheel were not operated for more than 10 s while the hand was still present on or above the mouse and consisted of a continuous vibration for 4 s at a 40% level of the motor power. It notified the user to let go of the mouse and to relax the hand on the desk. The delay of 10 s was based on an informal test with a few users and was chosen form a range between 6 and 12 s. The users did not perceive this signal as annoying (De Korte et al., 2008; De Kraker et al., 2008; Meijer et al., 2009).
  The vibration continued (with a maximum of 4 s) until the buttons of the mouse were clicked, the scroll wheel was rolled or the hand was removed from the mouse. Every time the participant clicked or scrolled, the timer-counter was reset to zero; therefore, during ordinary, active use of the mouse no-feedback signal

was generated. The feedback signal did not stop due to cursor movements and, therefore, 'stirring' would not stop the vibration. This signal will be referred to as the continuous signal.

The second tactile feedback signal was a modification of the Hoverstop signal. Again, the feedback signal started when the buttons and/or scroll wheel were not operated for more than 10 s with the hand still present on or above the mouse, but this time the vibration switched on for 1.5 s, then off for 1 s and on again for 1.5 s. Within the total duration of 4 s, the signal continued to go on and off until the hand was removed or buttons or scroll wheel were used. The idea behind this signal was, that a user who is locked up in his or her work may reduce the focus on the work at first but might not take the desired action. The second phase, after the pause, could then trigger the right action. This signal will be referred to as the on-off-on signal. The mouse for the on-off-on feedback signal was identical to the Hoverstop mouse (Fig. 1) and developed for this experiment by Engineering Spirit Ltd.

Figure 1 On the left the Hoverstop Mouse (continuous signal) and on the right the mouse with the on-off-on signal, both are identical on the outside. The mouse on the right was also used for the no-feedback condition, but then the vibro-motor was switched off

- *No-feedback signal condition*. With the modified Hoverstop mouse it was also possible to shut down the vibration motor, while still detecting the presence of the hand. This setting was used for the no-feedback signal condition. For the visual feedback conditions, the vibro-motor was switched off as well, but the capacitive sensor kept monitoring the presence of the hand on or above the mouse and kept triggering a software program that could set-off the visual feedback signals.

- *Visual feedback*. The two visual signals were a peripheral signal and a full screen transparent signal. Each signal was designed based upon expert opinion. The visual feedback was displayed on the monitor 10 s after the hand was detected on or above the mouse while scroll wheel and buttons were not used, with a maximum of 4 s.
  The first visual signal, the peripheral signal, consisted of 4 boxes which popped up in the corners of the screen. In this way, the signal could be noticed without

moving the eyes wherever the focus on the screen. The inner box was orange with a black frame. In this way the colour would always stand out regardless of the background (Fig. 2). This signal will be referred to as the peripheral signal.

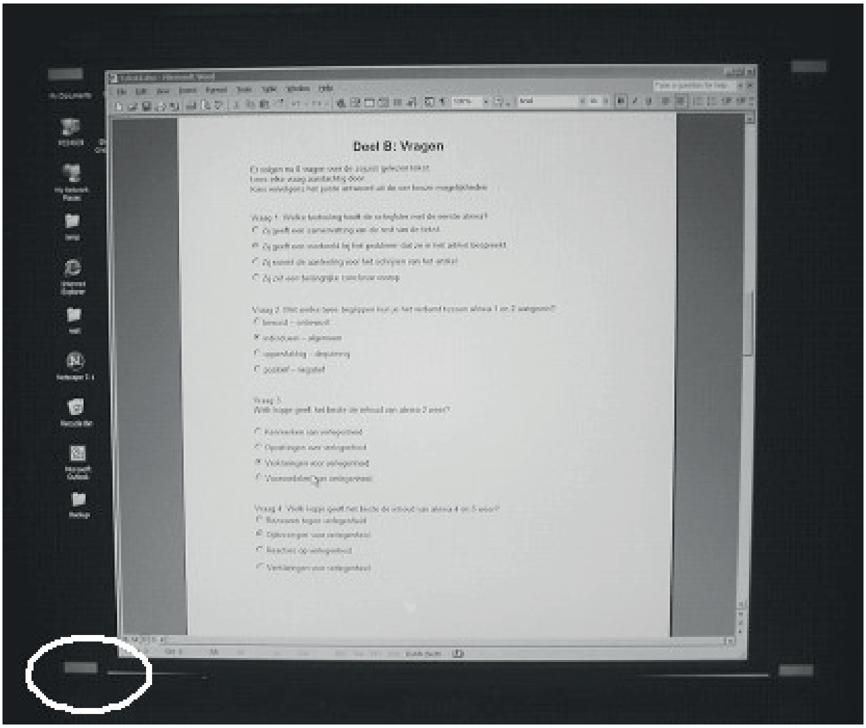

Figure 2 Peripheral signal: orange boxes in the four corners of the computer screen (brightness of the picture is altered and the box in the left lower corner is circled to show the effect more clearly to the reader)

The second visual signal consisted of a full screen transparent signal that was reflected on top of the text. Alternating bands of grey and white started in the middle and moved outwards. Fig. 3 shows a schematic drawing of the transparent signal (below) as well as what it looked like on the computer screen (above). This signal will be referred to as the transparent signal.

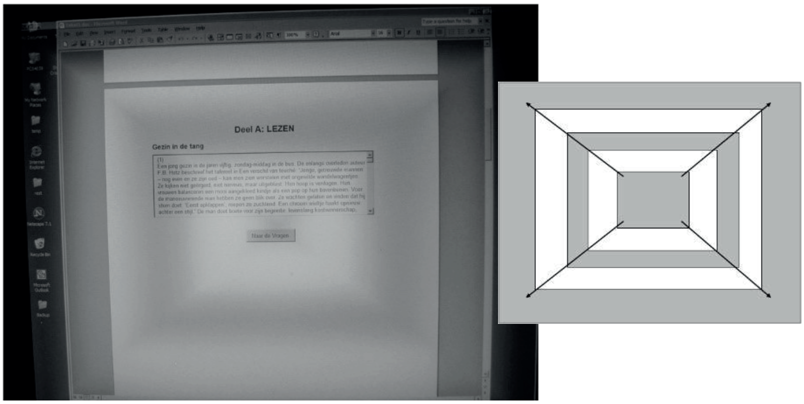

Figure 3 On the left the transparent signal reflected on the text (darker than participants saw on the screen in reality) and on the right a schematic drawing of the transparent signal)

Monitor

The experiment was performed on a normal computer. However, in order to be able to provide the transparent full screen signal, a special monitor was used (Fig. 4). The monitor was used throughout the experiment.

The monitor consisted of a black box containing two flat screens and a two-way mirror. A flat screen monitor was placed behind the two-way mirror. The information on this flat screen was visible for the participant in front of the mirror, like a normal monitor. A second monitor was placed above the mirror, so that images on this monitor were reflected in the mirror. When both monitors provided input, one behind the mirror and the other in the mirror (reflection), the participant saw both images at the same time, they blended.

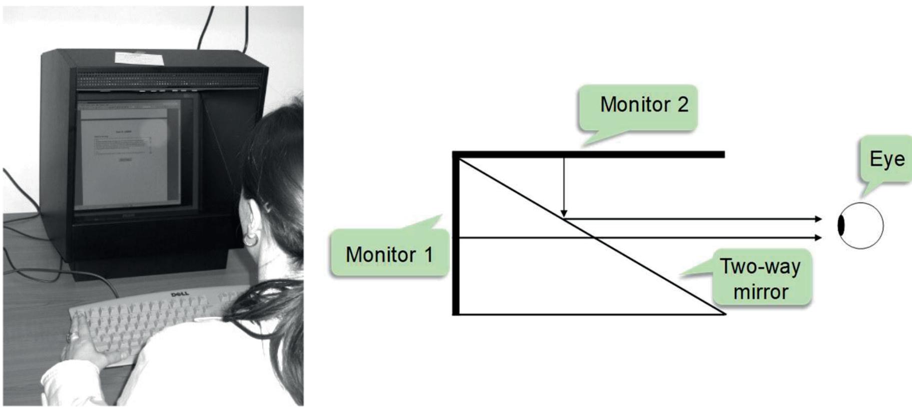

Figure 4 On the left the monitor used during experiment from the user point of view and on the right a schematic drawing of the monitor

Software

The experiment was developed in MS Word using the toolbox and Visual Basic. Participants interacted with the software using the mouse. No other processes were run on the computer during the experiment, besides the logging of the data, the Hoverstop mouse software and (when necessary) the programs for the visual signals. The Hoverstop mouse software was used to control the settings of the hover mouse. The data were logged using the program RSI-Master (http://www.rsimaster.com/index.html). The programs for the visual signals, one for each signal, were developed in Borland C++, by TNO.

Texts

For the standardized task 10 Dutch texts for the Dutch secondary education exam (2001-2004) were used (www. oefenexamens.nl). To manipulate task difficulty five easy and five difficult texts were selected.

The five easy texts were at intermediate general secondary education level and consisted of approximately 800 words. The five difficult texts were at higher general secondary education level and consisted of approximately 1400 words. Each text ended with 8 multiple-choice questions about the text. The text and questions could be read on the computer screen.

6.2.3 Task

Each participant was asked to read a text that was visible in a textbox of 4 cm high, with a scrollbar on the right. The box contained approximately 110 words, 9 lines, of text. Average readers reach approximately 200 words per minute (wpm). With this box size it was estimated that it would take approximately 1 min until participants had to scroll for the next piece of text. After having read the text participants had to answer 8 multiple-choice questions about the text. A question contained on average 55 words, which means that a feedback signal may occur once per question on average when a participant shows lifted hand and finger behaviour. Two levels of task difficulty were presented, an easy task level and a more difficult task level. We expect a stronger effect on the task with the higher difficulty level than the task with the lower level. Participants will take more time to complete the difficult texts, due to the number of words to read and more difficult questions to answer.

6.2.4 Protocol

The experiment took approximately 4 h, depending on the individual reading speed of the participants. The standardized task was carried out 10 times, first two times without any signal - the no-feedback signal condition - and then 8 times with a signal.

Participants first received an instruction on paper about the experiment. It explained in short the task and the type of feedback signals they could expect and how to react to these signals. When a signal was detected, they had to remove the hand from the mouse and place it on the desk. This instruction was available for the participant throughout the experiment. The experimenter told participants about the sequence of the different stages in the experiment and also when they were allowed to take a break. After this, participants signed an informed consent, filled in a questionnaire on general characteristics of the participants (e.g. age and education level), and started with the first condition.

In total, there were five conditions, one without feedback and four conditions with a feedback signal. The no-feedback signal condition was always given first and used as a control. The four conditions with feedback signals were systematically varied, which resulted in 24 different orders. A pair difficult text-easy text was fixed, in total five pairs existed. The difficult text was always given first.

In each condition, the document was opened, and users had to push the ‘start’ button to get to the text. This action logged the starting time of the task. The textbox with the text became visible and the participant read the entire text. After reading the text the participant proceeded to the questions using a button ‘to the questions’. Then 8 multiple-choice questions were displayed, and the participant used the mouse to select the correct answer. Participants were allowed to go back to the text to read a specific part again in order to answer the questions. At the end of the questions another button was placed, the ‘end’ button. By pushing this button the end-time was logged, the answers saved, the document closed and the next document was opened. Participants were now ready to start reading the second text (easy), which followed the same steps as described above. After this condition, the participants received a paper questionnaire containing questions on the difficulty level of the text and on how they perceived the noticeability, obtrusiveness and the match between the signal and required action of the specific feedback signal. Obviously these last questions were not asked after the no-feedback condition.

At the end of the experiment, participants were asked to complete a questionnaire in which they were asked to rank personal preference, efficiency and effectiveness of the five conditions.

6.2.5 *Dependent variables*

The following dependent variables were measured:

- Effectiveness (RSI-Master):
  - %time-hand on mouse: percentage time of the total task time the hand was placed on the mouse;
  - %time-buzzer: percentage time of the total task time the buzzer was on (before the hand was removed from the mouse);
- Efficiency (MS Word Toolbox/Visual Basic):
  - Total time on task;
  - Number of errors in answering the multiple-choice questions;
- Satisfaction (questionnaires):
  - Rank personal preference for feedback signals from 1. (like to use) - 5. (do not like to use);
  - Rank efficiency (amount of time spent on task and/or performance) of feedback signals from 1. (most efficient signal) - 5. (least efficient signal);
  - Rank effectiveness (decrease duration of lifted hand/fingers) of feedback signals from 1. (most effective signal) - 5. (least effective signal);
  - Rating noticeability, obtrusiveness and match between signal and required action on a ten-point scale from 1. (not at all) - 10. (extreme);
- Task difficulty:
  - Rating difficulty of the text on a ten-point scale from 1. (difficult) - 10. (easy).

6.2.6 *Statistics*

Following a within-subjects design, two independent variables were manipulated: type of feedback signal and task difficulty. Type of signal consisted of 5 levels: ‘no-feedback signal’ (mouse functions normal and gives no-signal), ‘continuous signal’ (the Hoverstop mouse was used for this signal), ‘on-off-on signal’, ‘transparent signal’ and ‘peripheral signal’ (modified mouse). Task difficulty consisted of 2 levels: ‘easy’ and ‘difficult’.

A repeated measure ANOVA was used to analyze recorded data and Tukey HSD was used for post hoc testing. Friedman ANOVA was used for questionnaire data and Wilcoxon was used for post hoc testing. An alpha level of 0.01 was used to indicate statistical significance, unless specified differently.

6.3 Results

6.3.1 Feedback signal

Effectiveness

A significant effect of type of feedback was found on %time-hand on mouse ($F(4,92) = 10.57$, $p < 0.001$). Tukey HSD test showed a significant difference between the no-signal conditions and the other 4 conditions. There was no significant difference between the 4 conditions with a signal. In the no-signal condition 64% of the time the hand was placed on the mouse. In the other conditions approximately 48% of the time the hand was placed on the mouse (Fig. 5).

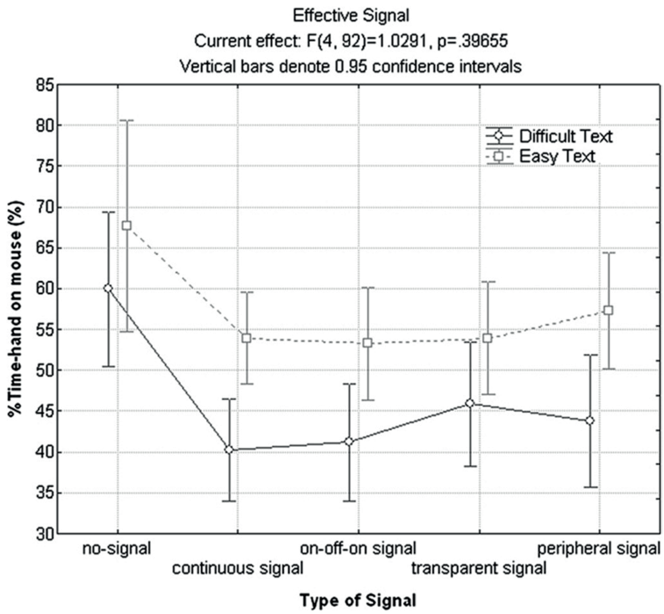

Figure 5 Effectiveness of the signals, by %hand-time on mouse, as function of type of signal

The effect of type of feedback signal on %time-buzzer was significant ($F(4,92) = 46.48$, $p < 0.001$) (Fig. 6). Tukey HSD test showed that there was a significant difference between the no-signal condition and the other 4 feedback conditions. There was no significant difference between the 4 conditions with a signal. In the no-signal condition 5.7% of the time the buzzer was on. (Note that the buzzer was not actually on, because the vibro-motor was shut off). For the other 4 conditions the %time-buzzer was on for approximately 1% of the total time. No interaction effect was found. An interaction effect between type of signal and task difficulty was found ($F(4,92) = 4.83$, $p < 0.001$).

For the continuous signal and the peripheral signal, the difference between easy and difficult tasks was smaller than for the other signals (Fig. 6).

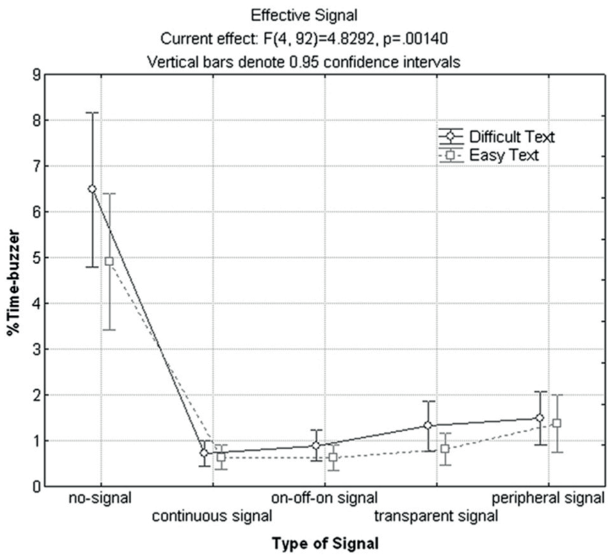

Figure 6 Effectiveness of the signals, by %time-buzzer, as function of type of signal

Efficiency

There was no effect of type of feedback signal on the total time on task ($F(4,92) = 1.35$, $p > 0.01$). The average time spent on reading a text was 799 s (approximately 13 min) (see Fig. 7). No interaction effect between type of signal and task difficulty was found.

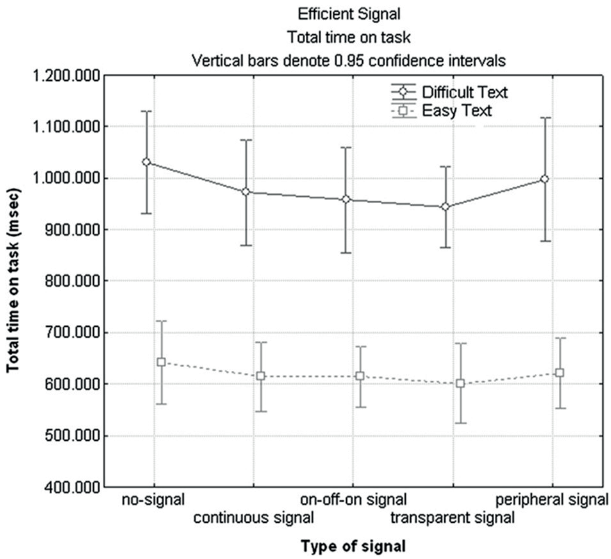

Figure 7 Efficient signal, by total time on task, as a function of type of signal

Six questions per text were included in the analysis. Questions 4 and 8 from text 7 had to be excluded because of interpretation problems, therefore questions 4 and 8 of each text were excluded from the analysis.

Type of feedback signal had no effect on the number of errors made (answering multiple-choice questions) ($F(4, 92) = 1.10$, $p > 0.01$) (Fig. 8). No interaction effect between type of signal and task difficulty was found.

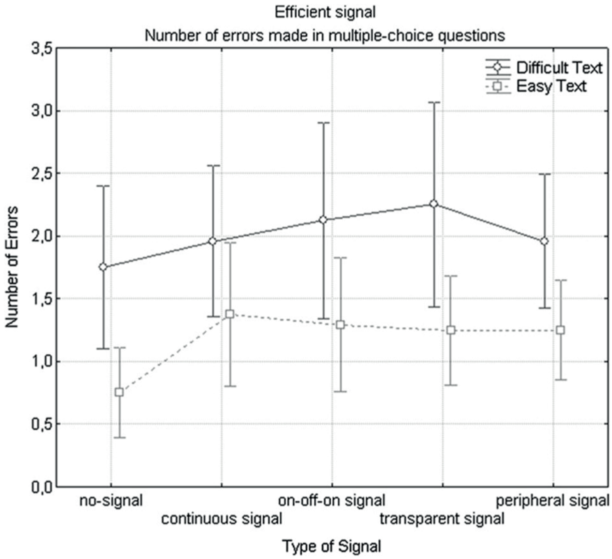

Figure 8 Efficient signal, by number of errors made, as a function of type of signal

Satisfaction

Significant effects were found for the comfort and acceptability of the feedback signals. A statistically significant effect was found for feedback signals when participants ranked the signals based on their personal preference ($X(4, N = 24) = 30.74$, $p < 0.001$). From Fig. 9 it is clear that participants preferred the no-signal condition (1.8). The continuous signal (2.8), on-off-on signal (3.0) and peripheral signal (3.1) were ranked equal. The transparent visual signal was least preferred (4.3).

Users preferred the no-feedback signal condition, their current situation. However, when asked to work with a signal, they preferred the tactile conditions or a peripheral signal.

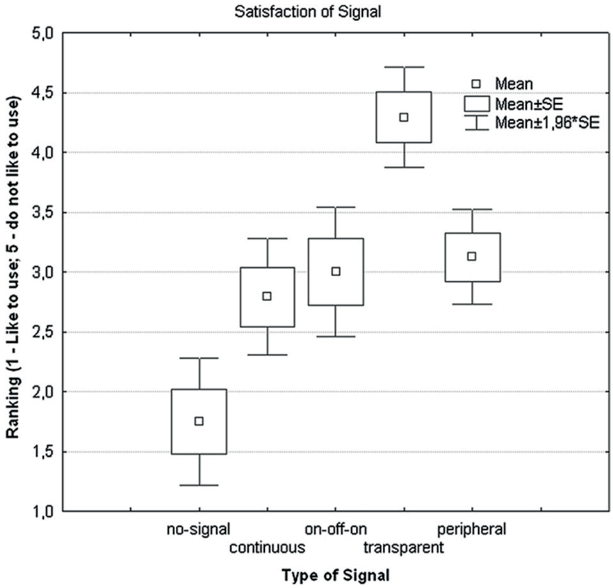

Figure 9 Personal preference as a function of type of signal

A statistically significant effect was found when participants ranked the efficiency of the signals from 'most efficient (1)' to 'least efficient (5)' (X(4, N = 24) = 14.96, p < 0.01). The no-signal condition received the lowest rank, 2.2, meaning most efficient for the participants. The continuous signal (2.7), the on-off-on signal (3.1) and the peripheral signal (3.2) were medium efficient. Least efficient for the participants was the transparent signal (3.9). Furthermore, a statistically significant effect was found for ranking the signals on effectiveness ('most effective (1)' to 'least effective (5)') (X(4, N = 24) = 57.19, p < 0.001). From Fig. 10 it can be derived that both tactile conditions and the transparent condition were equally effective as rated by the participants. The no-signal condition was rated least effective (4.5). The peripheral condition was just below that (3.9).

In addition to ranking the signals on personal preference, efficiency and effectiveness, participants were asked to rate noticeability, obtrusiveness and the match between signal and required action from 'Not at all' (1) to 'Extreme' (10).

Statistically significant effects were found for noticeability (X(3, N = 23) = 38.62, p < 0.001) obtrusiveness (X(3, N = 23) = 18.61, p < 0.001) and match between signal and required action (X(3, N = 23) = 40.79, p < 0.001).

From Table 1, it can be seen that the noticeability was rated high for both tactile signals and significantly less for the peripheral signal. In addition, obtrusiveness was rated lowest for the peripheral and continuous signal. Finally, participants rated the tactile signals to be better matched with the required action.

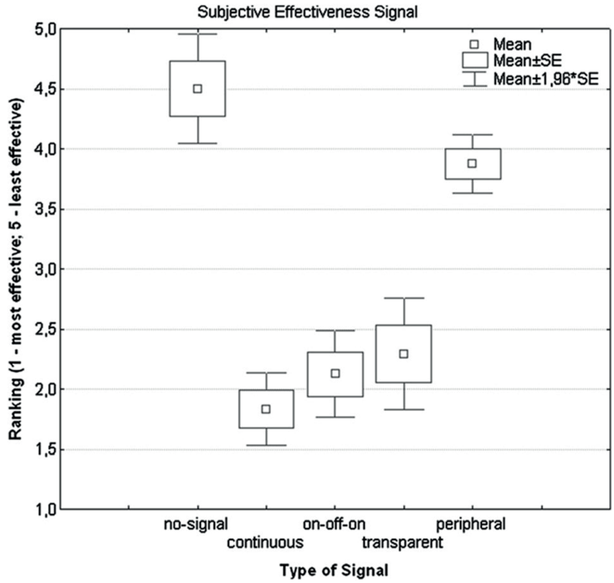

Figure 10 Subjective effectiveness as a function of type of signal

Table 1 Ratings of noticeability, obtrusiveness and the match between signal and required action (mean (SD)) from ‘Not at all’ (1) to ‘Extreme’ (10). * = $p < 0.05$

	Signal								
	A		B		C		D		
	Continuous		On-off-on		Transpar-ent		Peripheral		Significance
Noticeability	9.0	(1.3)	9.2	(1.3)	9.3	(1.0)	6.2	(2.3)	*D<A,B,C
Obtrusiveness	5.0	(2.3)	5.8	(2.8)	6.8	(2.7)	3.9	(2.1)	*A<B,C; D<B,C
Match signal-re-quired action	8.3	(1.3)	7.8	(2.0)	4.7	(2.1)	4.6	(1.9)	*C,D<A,B

6.3.2 *Task difficulty*

To check for learning effects during the experiment, all data were analyzed based on the presentation order for %time-hand on mouse. If any learning effect took place, this variable would have a lower value for later conditions. A significant effect of presentation order on %time-hand on mouse was found ($F(9,207) = 11.96$, $p < 0.001$). A Tukey HSD was carried out and showed that this effect was based on the difference between the first two no-signal presentations (always presented first) and the remaining 8 (with signal) presentations. There is a learning effect between the no-signal (control) condition and the experimental conditions with a signal, but not a learning effect between the experimental conditions.

The effect of task difficulty on %time-hand on mouse was significant ($F(1,23) = 85.16$, $p < 0.001$) (Fig. 5). The %time-hand on mouse was lower for the difficult task condition (46.2%) than for the easy task conditions (57.2%). No interaction effect between type of signal and task difficulty was found for all of these tests. The effect of task difficulty was significant on %time-buzzer ($F(1,23) = 16.79$, $p < 0.001$) (Fig. 6). On average the %time-buzzer was lower for the easy task condition (1.67%) than for the difficult task condition (2.18%).

Furthermore, the effect of task difficulty on the total time on task was significant ($F(1,23) = 376.04$, $p < 0.001$) (Fig. 7). Participants needed on average 618 s (10 min) to read the texts in the easy condition, and 980 s (16 min) to read the texts in the difficult condition. No interaction effect between type of signal and task difficulty was found.

After each condition (after the second text), participants rated the text from 1 (difficult) to 10 (easy). A Friedman ANOVA was used for the subjective ranking on task difficulty. The effect of task difficulty was statistically significant ($X(1, N = 24) = 23.00$, $p < 0.001$) indicating that users did experience the difficult tasks as more difficult than the easy tasks. The average rate for the difficult tasks was 5.4 and for the easy tasks 7.0.

We also compared all the rates for the difficult tasks with each other, using a Friedman ANOVA. No effect was found ($X(4, N = 24) = 2.68$, $p > 0.01$) indicating that overall, all difficult conditions were considered equally difficult. A similar test was done for the easy tasks. Again no effect was found ($X(4, N = 24) = 3.26$, $p > 0.01$) indicating that overall all easy conditions were considered equally simple.

6.4 Discussion

This study investigated the effects of non-obtrusive visual and tactile feedback, suitable for an office environment, as a reminder to reduce continuous lifted hand/finger behaviour. The feedback was designed to be minimally disruptive during task performance and only lasted for 4 s.

From the objectively recorded measures, it appeared that feedback in general is effective to reduce continuous lifted hand/finger behaviour. One could say that the intended behaviour with feedback was easy to learn, since participants immediately showed a behavioural change in the feedback conditions after the no-feedback condition. Also, we may conclude that the feedback signals were intuitive, as participants showed the behavioural change directly without training. However, long term effects remain unclear and extinction effects were not investigated at present. In future research, it would be of interest to know if mouse users continue to use their learned skills after a period of using a feedback signal as a training device and then removing it. This would provide recommendations on the use of a computer mouse aiming to stimulate healthy computer behaviour with feedback signals, either as a training device or as a permanent preventive tool for computer work which can be used without extensive practice.

Providing feedback resulted in the percentage hand above mouse time to decrease from 64% to 48%. Using these figures to generalize to a theoretical working day of 5 h intensive mouse use, it would mean a reduction of lifted hand/finger behaviour from approximately 3.2 h to 2.4 h, a reduction of 48 min. These results seem promising in order to prevent arm complaints and are in line with the results of previous studies (De Korte et al., 2008; De Kraker et al., 2008). However, Meijer et al. (2009) found no effects of the use of a mouse with a continuous feedback signal on the prevalence and incidence of upper extremity musculoskeletal symptoms in their randomized controlled trial with 8 months of follow-up, except that it lowered disability scores in participants who already had symptoms.

The measures on task performance indicated that providing a feedback signal did not have a negative effect on task performance as compared with the reference condition, neither time to complete the task nor the number of errors were affected. This is in line with other studies on feedback. For instance, Sigurdsson and Austin (2008) found that real-time visual feedback to improve postures at computer workstations did not affect task performance and also Hopp et al. (2005), who studied the use of attention-directing tactile cues to aid management of multiple tasks, found that the time to complete the task was unaffected.

The objectively recorded measures showed no differences for efficiency, effectiveness and learnability for the different types of feedback. Therefore, one could conclude that there are no objective reasons to choose one feedback signal over another. However, when user experience is taken into consideration, a different conclusion has to be drawn. Here, notable differences between types of feedback are seen, relating to comfort and acceptability. Participants rated the no-feedback signal to be the most and the transparent feedback signal to be the least efficient and preferable for use. The explanation for these findings may be twofold.

The first possible explanation is methodological in nature. The test periods for the feedback signals may have been too short to get fully used to the type of feedback. According to De Looze et al. (2003) and Vink et al. (in press) it is important to give end users the opportunity to test a product for a longer period of time because short term comfort is not always the same as long term comfort. It remains to be seen if experiencing a feedback signal for about 30 min in our experiment was enough to become used to the signal. This may have influenced the results on satisfaction.

Another explanation may be the influence of the feedback signals on the workload. Oakley et al. (2000) also reported on differences between behavioural measures and subjective experience. They suggest that workload plays an important role: cognitive resources are required for a task and there is a finite amount of these. As a task becomes more difficult (for example, attention has to be paid to the feedback), the same level of performance can only be achieved by investing in more cognitive resources. Users may then perform the tasks equally well and fast and yet find them more frustrating and requiring more effort to complete.

Also Hopp et al. (2005) suggested that the arousing characteristics of feedback may interfere with other modes of performance, particularly in more complex tasks, and therefore could be experienced as particularly disruptive. However, this hypothesis was not supported by their data. In their experiment the multiple-task environment was rated equally difficult by the participants with or without tactile feedback, and task performance was also not affected.

Based on the subjective ratings, it is questionable if we succeeded to design non-obtrusive, yet effective feedback signals. As is shown in the results, the peripheral signal is rated least obtrusive, but was also rated to be the least effective of the feedback signals for changing behaviour. The continuous tactile feedback signal seems to be the best option of the four feedback signals; not only effectiveness and efficiency were rated reasonable, it also scored best on the perceived match between the feedback signal and the required action, which is considered to be an important factor for usability (Norman, 2002).

The balance between effectiveness and non-obtrusiveness may be optimized by further testing properties of the four types of feedback signals, such as timing, duration and intensity. For example, it would be interesting to study the effects of the continuous tactile signal with a lower percentage of the motor power (in the experiment we used 40%), since a lower intensity may be less obtrusive. According to Jones and Sarter (2008), duration of vibrotactile stimuli holds the most promise for encoding information in tactile displays, next to locus. They also suggest that, when the tactile signal is functioning as a simple alert rather than a complex message, people would prefer the duration of the tactile pulses to be between 50 and 200 ms, as stimuli of longer durations are perceived as annoying. Comparing this with the duration of the tactile feedback signals used in the present study, i.e. 4 s, there seems to be room for improvement.

Another way to optimize the balance between effectiveness and unobtrusiveness is to provide the user with more control. According to Keyson (2007), understanding and sense of being in control is one of the three key factors that are important for user experience. The other two factors are sense of emotional appeal and engagement, and expected and perceived functional performance. With this in mind, it may advisable to give people the opportunity to set their own feedback properties, such as type, duration or intensity, in order to fit the users' context and tasks.

Furthermore, it would be of interest to study the effects of multimodal feedback, as promising effects can be seen from the literature (Vitense et al., 2003). The idea behind multimodal interfaces is that, like in humanehuman communication, effective communication between human and machine is likely to take place when different sensory input channels are used in combination (Oviatt, 2003; Sebe, 2009; Vitense et al., 2003). They may have many advantages, such as adding alternative communication methods to different situations and environments, giving the user more flexibility and control (Sebe, 2009). For that reason it may improve subjective experiences of using feedback to influence behaviour.

Finally, an unexpected effect will be considered. Unfortunately, we did not randomize the difficulty of the tasks, necessary to draw conclusions on the relationship between task difficulty and computer behaviour. By not randomizing, we could have been changing behaviour in the difficult task which might have positively affected the behaviour in the easy task, which was run last. However, the results contradict this possibility. The tendency to hold the hand and fingers on or above the mouse was lower during the difficult task than during the easy tasks. Therefore, we assume that lifted hand/finger behaviour is dependent on task difficulty. An explanation for this finding might be, that with difficult tasks mouse behaviour is more active in nature (more clicks and/or scrolling) compared to easy tasks, preventing the appearance of a feedback signal. In this experiment, for instance, participants could have scrolled and/or clicked more often to better comprehend the text and/or the questions. Further research is necessary to find out how lifted hand/finger behaviour and task difficulty are related with each other.

6.5 Conclusion

With this study we have shown that intelligent products, in this case a computer mouse with feedback, may have benefits to motivate office workers to adopt healthy, comfortable and productive work styles. The results suggest that it is possible to provide effective and efficient feedback with a non-obtrusive signal, to monitor compliance and to trigger persuasive interventions. However, there is a clear dissociation between objective measures and subjective experiences. Also, task difficulty seems of influence on specific behaviour. This study shows the importance of studying user experiences, as they determine the acceptability of a product, and to take into account the context and the tasks in which a product is used.

Acknowledgements

The authors thank the Dutch Ministry of Economic Affairs and Hoverstop BV for supporting this study.

6.6 References

Aarts, E., De Ruyter, B., 2009. New research perspectives on ambient intelligence. Journal of Ambient Intelligence and Smart Environments 1, 5-14.

Akamatsu, M., MacKenzie, I.S., Hasbroucq, T., 1995. A comparison of tactile, auditory and visual feedback in a pointing task using a mouse-type device. Ergonomics 38, 816-827.

Cook, D.J., Song, W., 2009. Ambient intelligence and wearable computing: sensors on the body, the home, and beyond. Journal of Ambient Intelligence and Smart Environments 1, 83-86.

De Korte, E.M., De Kraker, H., Bongers, P.M., Van Lingen, P., 2008. Effects of a feed-back signal in a computer mouse on movement behavior, muscle load, productivity, comfort and user friendliness. Ergonomics 51, 1757-1775.

De Kraker, H., De Korte, E.M., Van Mil, F.L.K., Rijs, B.P., Bongers, P.M., 2008. The effect of a feedback signal in a computer mouse on hovering behaviour, productivity, comfort and usability in a field study. Ergonomics 51, 140-155.

De Looze, M.P., Kuijt-Evers, L.F.M., Van Dieën, J.H., 2003. Sitting comfort and discomfort and the relationships with objective measures. Ergonomics 46, 985-997.

Dennerlein, J.T., Yang, M.C., 2001. Haptic force-feedback devices for the office computer: perfomance and musculoskeletal loading issues. Human Factors 43, 278-286.

Ducatel, K., Bogdanowicz, M., Scapolo, F., Leijten, J., Burgelman, J.C., 2001. Scenarios for Ambient Intelligence in 2010. IPTS, Seville. ISTAG 2001 Final Report.

Fairclough, S.H., 2009. Fundamentals op physiological computing. Interacting with Computers 21, 133-145.

Hopp, P.J., Smith, C.A.P., Clegg, B.A., Heggestad, E.D., 2005. Interruption management: the use of attention-directing tactile cues. Human Factors 47, 1-11.

ISO, 1998. ISO 9241-11:1998 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability. International Organization for Standardization, Geneva, Switzerland.

Jensen, C., Borg, V., Finsen, L., Hansen, K., Juul-Kristensen, B., Christensen, H., 1998. Job demands, muscle activity and musculoskeletal symptoms in relation to work with the computer mouse. Scandinavian Journal of Work Environment and Health 24, 418-424.

Jones, L.A., Sarter, N.B., 2008. Tactile displays: guidance for their design and application. Human Factors 50, 90-111.

Keyson, D., 2007. The experience of intelligent products. In: Schifferstein, H.N.J., Hekkert, P. (Eds.), Product Experience. Elsevier Science, San Diego, pp. 515-530. Kuijt-Evers, L., Steen, M., 2008. Early user involvement in designing intelligent products and environments. NES 2008 Abstracts. In: Helgadóttir, B. (Ed.), Ergonomics Is a Lifestyle. Gutenberg, Reykjavik, p. 98.

Lee, D.L., 2006. Prevalence of Lifted Finger Behavior and Postures during Two-button Computer Mouse Use. In: Human Factors and Ergonomics Society Annual Meeting Proceedings. Industrial Ergonomics. Human Factors and Ergonomics Society, Santa Monica. 1361-1365.

Matthews, T., Rattenbury, T., Carter, S., 2007. Defining, designing and evaluating peripheral displays: an analysis using activity theory. Human-Computer Interaction 22, 221-261.

Meijer, E.M., Sluiter, J.K., Frings-Dresen, M.H.W., 2009. Effectiveness of a feedback signal in a computer mouse on upper extremity musculoskeletal symptoms: a randomised controlled trial with an 8-month follow-up. Occupational and Environmental Medicine 66, 305-311.

Norman, D.A., 2002. The Design of Everyday Things. Basic Books, New York. Oakley, I., Adams, A., Brewster, S.A., Gray, P.D., 2000. Putting the feel in look and feel. Proc. CHI Letters 2 (1), 415-422.

Oviatt, S.L., 2003. Multimodal interfaces. In: Jacko, J.A., Sears, A. (Eds.), The Human Computer Interaction Handbook. Fundamentals, Evolving Technologies and Emerging Applications. Lawrence Erlbaum Associates, Publishers, London, pp. 413-434.

Raisamo, J., Raisamo, R., Kangas, K., June 5e7, 2004. Interactive graph manipulation tools enhanced with low-frequency tactile and force feedback. In: Buss, M., Fritschi, M. (Eds.), EuroHaptics 2004: Proceedings of the 4th International Conference EuroHaptics 2004. Technische Universität München, München Germany. Inst. of Automatic Control Engineering, Munchen, pp. 498-503.

Sebe, N., 2009. Multimodal interfaces: challenges and perspectives. Journal of Ambient Intelligence and Smart Environments 1, 23-30.

Sigurdsson, S.O., Austin, J., 2008. Using real-time visual feedback to improve posture at computer workstations. Journal of Applied Behavior Analysis 41, 365-375.

Slijper, H.P., Richter, J.M., Smeets, J.B., Frens, M.A., 2007. The effects of pause software on the temporal characteristics of computer use. Ergonomics 50, 178-191.

Van den Heuvel, S., De Looze, M.P., Hildebrandt, V.H., Thé, K.H., 2003. Effects of software programs stimulating regular breaks and exercises on work-related neck and upper limb disorder. Scandinavian Journal of Work Environment and Health 29, 106-116.

Van der Veer, G.C., Puerta Melguizo, M.C., 2002. Mental models. In: Jacko, J.A., Sears, A. (Eds.), The Human Computer Interaction Handbook. Fundamentals, Evolving Technologies and Emerging Applications. Lawrence Erlbaum Associates, Publishers, London, pp. 52-80.

Van Erp, B.F., Verschoor, M.H., 2004. Cross-modal visual and vibrotactile tracking. Applied Ergonomics 35, 105-112.

Vink, P., Bazley, C., Kamp, I., Blok, M., Possibilities to Improve the Aircraft Interior Comfort Experience, Applied Ergonomics, in press.

Vitense, H.S., Jacko, J.A., Emery, V.K., 2003. Multimodal feedback: an assessment of performance and mental workload. Ergonomics 46 (1-3), 68-87.

7

The digital stress coach. Total control over your mental health, or 'big brother is watching you'?[1]

[1] Published as: Van Lieshout, M., Wiezer, N., & De Korte, E. (2015). The digital stress coach. Total control over your mental Health, or 'Big Brother is watching you'?. In L. Kool, J. Timmer, & R. van Est (Eds.), *Sincere Support - The rise of the e-coach* (pp.161-176). Den Haag: Rathenau Instituut.

Marc van Lieshout
Noortje Wiezer
Elsbeth de Korte

7.1 Introduction

It's an appealing idea to have a digital coach that helps you reduce work-related stress and get more energy from your work. We all feel stress. And we all want to do work that energises us. Stress is also something we'd like to avoid. It's a societal problem and a digital stress coach might offer a solution that fits in with the 'future world of work'. The development of digital stress coaches is still in its infancy. That means that this is the right time to think about how we can guide their introduction in a responsible manner. This chapter describes current developments regarding stress coaches and discusses various issues related to their introduction.

7.1.1 Work-related stress

When we talk about stress, we often mean that we have a heavy workload or that we're feeling the strain of a poor relationship with one of our colleagues. Strictly speaking these are not examples of stress, but of *stress factors* that lead to stress. Stress is the body's natural response to an external stress factor, resulting in the 'fight or flight response'. Stress is not necessarily unhealthy, but it will become so if it goes on for too long, is too intense, or if we do not have enough time to recover from it. Unhealthy stress can lead to a variety of complaints. Minor complaints include forgetfulness, insomnia, worrying, irritation, or tension. Most people suffer from these every now and then. But the complaints can also become so serious that it is impossible to work or carry out normal everyday activities like grocery shopping. In that case, the person in question is suffering from burn-out. Fortunately, serious complaints such as these are uncommon, but the number of employees who indicate that they have symptoms of burn-out has risen sharply in recent years. More than thirteen percent of Dutch employees reported burn-out symptoms in 2012 (Koppes et al. 2013). What is most worrying is the growing number of young employees in this group.

Work-related stress is stress caused by someone's work situation. If an employee is unable to change a stressful situation at work, then that stress is likely to go on for a long time, making it unhealthy. Various factors can lead to work-related stress. Work pressure is one very important stress factor. This happens when an employee is unable to meet the demands made on him in the time available to him. Other stress factors include bullying in the workplace, conflicts with one's boss, aggressive customers or clients, job insecurity, boring or unchallenging work, and difficulty combining work and private life. A recent study (Koppes et al. 2013) shows that almost 30 percent of Dutch employees regularly work under enormous time pressure. Somewhat more than 15 percent of employees are occasionally harassed by colleagues, and almost 24 percent by external parties (customers, clients). Almost 50 percent of young employees (24 and younger) find their work monotonous, 8.5 percent of employees neglect family duties because of their work, and slightly more than 2 percent neglect their work occasionally owing to family obligations. In other words, the number of employees encountering stress factors at work is considerable.

Present-day social, economic and technological trends are only adding to work-related stress. Work is becoming more complex, flexible and individualised, and the boundary between our work and our private lives is blurring (Allvin et al. 2011; Houtman et al. 2012; Manyika et al. 2013). More complex work can be challenging, but it can also lead to stress. Flexible working practices can lead to job insecurity. They are also leading to a situation in which employees work for companies that increasingly operate in networks, in which they increasingly have to work in a variety of different business environments, with work becoming more individualised as a result (Wevers & Bongers 2013). Today, it is possible to work whenever and wherever we like (e.g. by means of telecommuting, 'The New Way of Working'; Pot et al. 2012). Because such flexibility makes it easier to strike a good work-life balance, it opens up the possibility of reducing work-related stress. It also blurs the boundaries between work and private life, however, with people running the risk of working all the time and everywhere - even on holiday. Some employees are unable to set and maintain such boundaries for themselves. Another risk is that flexible working practices eliminate one of the main buffers against work-related stress, namely the support of one's social network at work. All these changes can increase the risk of work-related stress (Houtman et al. 2012), and, what is more, they are making traditional interventions meant to improve working conditions increasingly irrelevant, since many of these are tailored to groups, departments, or entire organisations.

7.1.2 A role for a coach

Not everyone responds to stress factors in the same way. Some people cope with them better than others, and in fact the same person may have more trouble handling stress one day than the next. Whether stress leads to health complaints depends largely on whether the person concerned has the chance to recover from the stressful experience or situation. Does he have time to let his body return to its normal state, now that the 'fight or flight response' is no longer necessary? Employees who suffer stress symptoms often call in a coach for help. There are various forms of coaching for work-related stress. Mainstream coaching often concentrates on *coping with stress factors*. Examples include coaching that teaches people to deal with 'energy takers' or with disruptions and problems in the workplace, or that trains them in time management. Another form of coaching concentrates on *promoting recovery*, for example coaching combined with training in relaxation techniques or yoga. Less customary is coaching to learn to *remove* *or reduce stress factors*. That may not be possible in some extreme cases, but in many work situations the best way to prevent stress is to tackle it at the source. In this type of coaching, the idea is to find a solution to reduce or eliminate the stress factors rather than try to cope with them or their effects. At the moment, methods that address the sources of stress tend to focus on the circumstances affecting *groups* of employees (a team, department or organisation).

It is our observation that the risk of stress is increasing and as a result, there is a greater demand for stress-reduction support. Technology makes it possible to digitise such support, and digital support fits in well with the changing landscape of work, as it can be used whenever and wherever a person requires it. The development of the digital stress coach is still in its infancy, but that gives us time to think about how we can guide the introduction of such coaches in a responsible manner.

7.2 Digitising coaching

Technology is making major changes possible in traditional coaching practices. A decade ago we had never heard of a smartphone; today, we are surrounded by them. Technology is integrated, both visibly and invisibly, into the products that we use and the environments we inhabit. Technology recognises us as users and monitors and influences our behaviour, preferably in a personalised and intuitive manner, and sometimes without us even noticing it (Aarts & De Ruyter 2009).

The technological advances underpinning these changes are mobile Internet (the Internet made available on mobile devices like the smartphone), the 'Internet of Things' (data collected by sensors on devices or persons made accessible on the Internet) and cloud technology (applications and data that can be accessed and stored remotely). Mobile Internet makes it possible to track and influence all of a person's everyday routines. Mobile Internet technology is advancing rapidly, with intuitive interfaces and new formats emerging such as wearables, i.e. microchips that are worn on, underneath, or integrated into clothing. The Internet of Things makes it possible to collect data on a continuous basis, to track and link it to other data, and to produce increasingly reliable measurements. Those measurements might concern the user's bodily functions, feelings or behaviour, or his physical or social environment. Cloud technology, which makes it possible to request access to hardware, software and data online, allows us to store and process huge volumes of data and, for example, to use that data in mobile Internet applications (Manyika et al. 2013).

Coaching providers in the Netherlands are discovering the advantages - and the necessity - of a web presence and of the *digitisation of communication with the coach*. The advantages of digital interaction lie in its lower cost, its greater speed, and the possibility of leaving a message or filling in a form even when the other party is absent. Clients are growing more accustomed to digital forms of communication (e-mail, Skype and WhatsApp) to supplement their appointments or telephone consultations. Web-based counselling and support are also helping to digitise the coach. One important form of support is e-health software in which feedback is provided by a flesh-and-blood coach, but communication between the coach and the coachee is exclusively digital (for example the *Kleur je leven* [Colour your life] course by the Trimbos Institute[7]). That means that support is basically anonymous and is provided online when the client wants it and at his or her own pace. The digitisation of communication between coach and coachee is the first step along the road to digital coaching (Kool, Timmer and Van Est 2013).

There are numerous online stress tests or questionnaires that help visitors determine whether they are suffering from the symptoms of stress and whether they are expending more energy on their work than they get back from it. The results, which are often accompanied by suggestions or recommendations for improvement, are delivered immediately and are most likely pre-programmed and standardised.[8] Some of these tests are simply teasers for more elaborate programs (that users then have to pay for). Another version of such self-help tests are support apps that offer standard advice or coaching programmes. They supplement all the self-help tools. They can be used on a mobile platform and range from mindfulness exercises to breathing techniques and exercises to promote sleep.[9]

Another trend involves exploring the extent to which physiological signals can be used to indicate higher stress levels and to monitor stress. Sensors are constantly improving; they are getting smaller, have a much longer shelf life than before, can record much more data than in the past, and are easier to connect to storage devices. All these improvements make it easier to track physiological signals and link them to one another. In terms of stress, sensors can be used, for example, to measure heart rate variability (HRV) or galvanic skin response. Sensors such as eSense Skin Response make suggestions for improvement based on the temperature of the skin.[10] Tiny sensors coupled to a smartphone measure the skin's temperature. The data is fed into an app that makes standard recommendations based on the information it has received. We must be careful when interpreting the results of this kind of physiological measurements, however, since people often differ considerably in their physiological response to stress.

---

7 http://www.trimbos.nl/onderwerpen/preventie/depressie/kleur-je-leven.

8 https://www.sterkopjewerk.nl/Zelftest.

9 See, for example, www.digitalezorggids.nl, which offers and reviews the five best stress apps.

10 http://www.mindfield.de/en/products/eSense/eSense-Skin-Response.html.

Research on stress and the role that technology can play in alerting us to stress and tackling it early on investigates not only stress factors but also focuses on people's wellbeing. Work is one of the areas being explored. For example, Technology Foundation STW, the National Initiative Brain & Cognition (NIHC) and Philips Research will be investing three million euros over the next few years in the Healthy Lifestyle Solutions programme, with five research projects set up to digitise methods to help people get 'a good night's sleep, a balance between stress and relaxation, a healthy diet and sufficient exercise'.[11] The SWELL research programme (Smart Reasoning Systems for Wellbeing at Work and at Home, a COMMIT project[12]) focuses explicitly on the relationship between the work and home environments. SWELL aims to measure physical fitness, workload, and stress using advanced sensors that can provide individualised advice.[13] That makes it one of the few research programmes to specifically examine the work environment and to attempt to combine physiological features and perceived stress.

We can detect a number of trends from the forgoing. Coaching is becoming digitised. At first it was simply the style of communication that was digitised, but now new tools are being added, such as interactive questionnaires. The research instruments will grow more refined as time passes, and will combine physiological, situational, and personal features. Recommendations for improvement will be more personalised and responsive to the individual's changing circumstances. Finally, the instruments will collect aggregate data that is not necessarily traceable to any one person. This data can be used to generate information on groups of people, differentiated by a variety of different features.

The following sections discuss the possible effects that these trends may have on relevant individuals and, for example, on existing relationships between employers and employees. We explore how the advent of the stress coach fits into current regulatory frameworks and where problems may arise. In the concluding section, we recommend responsible ways of guiding the introduction of stress coaches in the workplace.

---

[11] http://www.newscenter.philips.com/main/research/news/press/2011/20111019-stw-nihc-research.wpd#.VGtu9zTF98E##

[12] COMMIT is a Dutch public private research programme focusing on ICT in the area of health and well-being, e-science, public safety and information processing. http://www.commit-nl.nl.

[13] http://www.commit-nl.nl/projects/swell-smart-reasoning-systems-for-well-being-at-work-and-at-home.

7.3 Exploring the societal impact of the digital stress coach

7.3.1 Immediate effects of digital stress coaches

Does the digital stress coach do what it claims to do? And does it do so effectively? Because there are as yet very few digital stress coaches, research into these questions is scarce. We simply do not know yet whether digital coaching produces better results than traditional coaching. To determine how effective and reliable digital coaches are, we need to know what they actually measure, what recommendations for improvement they give, and how good those recommendations are. These are methodological issues that apply equally in the case of non-digital coaching and, for example, with respect to the validity and reliability of questionnaires. A familiar problem is how to determine the boundary between two different recommendations: if the score is x, the user falls into one category; if the score is somewhat higher (or lower) than x, the user falls into the other category. Deciding on the dividing line is therefore crucial. The resulting categorisation is strict, but can be linked to minor variations in the measurement outcomes. Users will be inclined to look at the recommendations, and not at the underlying scores. Sloppy interpretation of the scores may worry users unnecessarily, or unadvisedly ease their minds.

There is another problem associated with using the physiological features of stress. Many studies have explored the relationship between *perceived* stress and *quantifiable* physiological stress, for example blood pressure and stress hormones. That relationship is exceptionally weak. If we were to divide people into one group that feels a lot of stress and another that feels little stress, we would discover that the two groups' stress hormone levels are by no means likely to differ as well (Van Doornen 2013). People vary considerably in their physiological stress response. One person may respond to stress with an altered heart rate, while the other produces more cortisol (the stress hormone). If an individual's data deviates from that of the overall mean in a population, that mainly says something about the extent to which the individual differs from the mean. It may not say anything about how he experiences stress. There are also too many steps between what is 'happening in the brain' and what can be observed with physiological measuring systems, and each and every step can influence observation. All of this makes it difficult to determine how much stress someone is feeling by taking physiological measurements. Caution is therefore advised when interpreting the results. Stress should preferably be measured in different ways (physiological and socio-psychological). Physiological measures reveal physiological responses that may constitute a health risk. They can function as a warning sign and encourage us to explore the causes of the physiological response, which may be work-related.

The problem of whether questionnaires and physiological data are reliable has consequences for the quality assurance of digital stress coaches. Medical devices must bear the CE marking. After all, a thermometer or blood pressure monitor has to work properly. But there is no such seal of approval for questionnaires and data analyses. It is difficult to standardise treatment protocols and empirical research methods in the field of coaching. The quality of coaching methods can be improved by means of internal quality assessment, supervision and peer review. Evidence-based research basically provides a 'gold standard', but that standard is more difficult to achieve in the case of coaching activities.

It is obvious that the provider of a digital stress coach is *liable* for its quality. In the case of digital coaches, however, quality assurance is still in its infancy. Usually, a good system of quality assurance and control only emerges after a method or service has been available commercially for a while, long enough to solve any teething troubles and gain experience in normal use. None of this is terribly important when digitisation is confined to communication, but as soon as measurement and feedback are entirely digitised, then quality assurance must be properly arranged in advance. This is something that the professionals (coaches and health and safety services) and technology developers should tackle. Those working in coaching would do well to validate the questionnaires that they use (scientifically) and to draft guidelines and rules concerning the validation and use of questionnaires (for example in the form of a seal of approval).

Research on the effectiveness of digital stress coaches could assist in drafting a set of proper quality standards for providers and clients.

7.3.2 The stress coach and the employer-employee relationship

Work-related stress is a troublesome topic in discussions between employers and employees about working conditions. That is largely because the two sides hardly ever agree about the point at which work-related stress becomes a health issue or on the causes of such stress (see also Wiezer et al. 2012). The Dutch Working Conditions Act [*Arbowet*] obliges employers to pursue a policy focusing on prevention; where this is not possible, the employer must limit work-related psychosocial strain. An employee must follow the employer's instructions concerning health and safety at work and inform the employer if he observes anything that could put health and safety in the workplace at risk. Employers could introduce a digital stress coach as part of their policy and instruct employees to use the stress coach for company health and safety purposes. It is even possible to imagine the digital coach's recommendations being regarded as instructions. The stress coach also creates the impression that it can determine objectively that there is a health risk. An employee who uses a digital stress coach and finds that he is running a health risk can hold his employer accountable. An employee who feels stressed but does not see any evidence in the physiological measurements will face more difficulties convincing his boss, unless the digital coach also measures 'perceived stress'. An employee who does not feel stressed but has sky-high stress values according to the digital coach will also have a difficult time. That in itself could be stressful!

A digital coach could be an interesting proposition for employers. The Dutch courts recently found in favour of an employer who wanted to fire an obese employee.[14] In the case concerned, there was in fact a specific reason that justified the dismissal (difficulty fitting the employee into work schedules). Employers might benefit from a digital coach that tracks whether employees stick to agreements that they have made (to lose weight, for example). At the moment, mandatory use of digital stress coaches in the workplace is still hypothetical. Serious objections can be raised (see below) and the commotion that might well ensue will prevent employers from introducing any such tool for the time being.

The foregoing raises a critical question: to what extent does the use of a digital stress coach constitute an unlawful violation of the coachee's privacy? Everyone - including employees - has a right to privacy. That right is laid out in the Universal Declaration of Human Rights and in the European Union's Charter of Fundamental Rights. The growing level of digitisation is putting pressure on this right, however, which manifests itself in two separate ways: the right to privacy on the one hand (no unlawful interference in one's communications, private life, family life or home) and the right to protection of personal data on the other. The Netherlands has exhaustive legislation concerning the protection of personal data that stipulates what is and is not permitted. The law takes sensitive personal data (such as information on race, religion, political and sexual preferences, medical data and so on) very seriously, with stricter guarantees applying.[15] In the case of the right to privacy, the law takes more of a case-by-case approach. There are no strict rules; whether or not someone's right to privacy has been violated depends on the situation.[16] Not every violation of privacy is unlawful. Work-related medical examinations are an example of an infringement of integrity that is permitted by law. The Working Conditions Act states that in highly specific cases, an employee can be

---

[14] See http://www.akd.nl/nl/kennis/publicaties/obesitas-een-reden-voor-ontslag.

[15] The Act indicates which data should be regarded as sensitive personal data. But personal data that is not defined as sensitive in the Act can still be perceived as such. For example, a woman who is in a safe house after fleeing domestic abuse will consider her address to be very sensitive data.

[16] The Personal Data Protection Act mainly governs procedural matters. This is known as the procedural approach to protection. The fundamental right to privacy set out by the United Nations and in the European Charter is always about substance. This is known as a 'substantive' approach (Gutwirth, Gellert & Bellanova et al. 2011).

ordered to undergo a medical examination in order to determine whether he is capable of doing his job or doing it safely (an example would be an eye test for airline pilots). Privacy guarantees fall into two categories covered by two different principles. The first is the principle of subsidiarity: are there other, less invasive ways that will achieve the same (or a satisfactory) result? The second is the principle of proportionality: is the intervention in proportion to the intended result? Each situation needs to be assessed in the light of these two principles.

Digital coaches may - unlawfully - violate both the right to privacy and the right to the protection of personal data. They enter our homes easily, while permanent monitoring by sensors placed on or in our bodies could constitute a violation of our physical integrity.

7.3.3 More data in employees' files

Another issue that will need to be considered carefully is that more data is available to add to employees' files. Data is collected on every employee that plays an important role in discussions concerning his or her career advancement and performance. With coaching being increasingly digitised, more data will also become available on employees' wellbeing, health, and physical and mental resilience. The growing number of sensors that attach to the body or to clothing will allow such data to be collected in real time (physiological signals via sensors, other information via apps), or much more often than is now the case. This information could also end up in the employee's file.

When a personnel file is set up, the employer must consider the employee's right to protection of his personal data as laid down in the Dutch Personal Data Protection Act [*Wet bescherming persoonsgegevens*]. According to this Act, the employer must meet a number of requirements when collecting personal data. We will not look at these requirements in detail here, but it is important to note that data collection must not be excessive and must be compatible with the designated aim of such collection (for example to assess the employee's performance and/or to satisfy certain statutory obligations). Imagine that an employer offers its employees a digital coach that allows the company to keep better track of their stress levels (in order to reduce work-related health risks). Certain data that is collected can be regarded as medical information (heart rate, blood pressure). Other data says something about the employee's state of mental health (perceived stress). This information may be relevant when assessing the employee's performance, but it is up to the employee to decide whether or not he wishes to share that information with his employer. An additional complication arises if the stress coach is used not only in the workplace (to monitor stress levels while someone is working) but also elsewhere, which is perfectly obvious. The employee could engage in activities intended to lower stress levels (e.g. digital therapy, sport, yoga, walking) in his private life as well, and those activities will also be monitored.

The pressure to deal responsibly with data provided by digital stress coaches is becoming more urgent for two reasons: first of all, it will be possible to collect much more data much more often than nowadays; second, a greater variety of data may become available. It is consequently important to ensure that this data is only added to the employee's file if doing so is compliant with the relevant rules and practices. Medical and other sensitive data should not be added to the employee's personnel file. Data related to his physical wellbeing is undoubtedly just as sensitive as medical data, and if possible even more closely associated with a specific context. An employee who is recently divorced will carry this personal, stress-inducing problem with him to work as well. Data on his stress levels will be retained, but the context in which that data should be considered may be forgotten. It is still very rare for people to talk about psychological problems at work. That is not entirely surprising. Recent research has shown that such problems play a much bigger role in the selection of new employees or in the retention of employees during a reorganisation than age or other health complaints, for example (Houtman, Koppes & Dekker 2013). Employers, employees, health and safety services, and occupational physicians must address the problem of how to deal with these trends on both fronts. There is still time and space for to address these issues. Employers, employees and occupational physicians will need to make firm agreements, preferably before digital stress coaches find their way into the workplace.

Efforts to improve privacy protection are being supported by impending changes to EU legislation. The Dutch Personal Data Protection Act will be replaced by the European Union's new Data Protection Regulation. When this will take place is unknown, and neither is the precise wording of the regulation clear. One instrument referred to in the regulation and relevant to us here is the Data Protection Impact Assessment. The current intention is to make that assessment mandatory under the regulation in certain cases. One such case could very well be the addition of data from a digital coach to an employee's file. In that event, the organisation would be obliged to conduct a risk assessment prior to introducing the digital coach and to take steps to tackle any risks found. A privacy officer, who would have an autonomous position within the company, would then have to ensure that the Data Protection Impact Assessment and any measures arising from it were implemented. And if significant changes were proposed at a later time, the assessment would have to be repeated.

7.3.4 Infringement of employee autonomy

One general aspect of every form of coaching is that the coachee must commit - at least morally - to performing the designated activities according to the designated schedules.

Those activities may consist of completing a questionnaire, allowing a certain bodily activity to be measured, or adhering to certain nutritional or exercise patterns (which are also monitored). Unlike a human coach, a digital coach may make a coachee feel subject to surveillance, even if he is using the stress coach voluntarily. After all, the digital stress coach is basically with him all the time and wherever he goes. It may cross the boundary between the person's working and private life because his private activities (for example exercise) become visible in the work context. As indicated by the example of the obese employee who was dismissed, employers have the tools to force an employee to adopt different behaviour. A digital stress coach is yet another such tool, and in that sense it affects employee autonomy. The more the digital stress coach can do, the more employee autonomy may be undermined. Increasingly, employees' positive freedom ('Being free to…') may be restricted and their negative freedom ('Being free from…') may decrease. In addition, many projects exploring the role that technology can play in stress reduction go way beyond work-related stress. A project such as SWELL focuses on people's 'wellbeing' and on physical fitness and vitality. The recommendations that a stress coach makes on the basis of measurements will target not only behaviour in the workplace but also healthy nutrition, rest periods, and physical activity. For the most part, these are matters that form part of the stress coach user's private life. That means that the digital stress coach (or the employer through the channel of the stress coach) is influencing behaviour in the employee's home, and that the employee's right to privacy is consequently being violated.

7.4 Policy implications

The assumption is that using a digital coach can help an employee recover from stress and learn to recognise stress early on. The opposite side of the coin is that this does not necessarily mean action being taken to ameliorate the causes of stress. In addition, the digital coach makes the employee more transparent, something that may weaken or strengthen his position. The volume of sensitive data that is being stored is increasing, but that data does not necessarily produce an accurate picture of the situation. Employees may encounter problems if the data stored on them paints an incorrect picture, or if it is used for the wrong purposes.

The digitisation of coaching is not limited to working hours, but extends into the private domain in various ways. While that may be necessary for the digital coach to function properly, it also undermines the employee's autonomy in the sense that he loses control over the images and data that may circulate about him.

In principle the Dutch Personal Data Protection Act combined with the Working Conditions Act address the foregoing privacy issues. These two Acts offer a good point of departure for assessing and tackling the privacy risks posed by stress coaches. It is important, however, to consider these risks carefully when setting up digital coaching practices. There are various tools available for such purposes, but they too require further development. One is the Data Protection Impact Assessment described earlier. Others involve systems in which privacy is already included as a factor in systems design (Privacy by Design), and methods for allowing data subjects (the people about whom data is collected) to have greater control of their own data (Van Lieshout et al. 2011). External factors will determine the pace at which these tools and methods are developed and whether they will be automatically included in systems design.[17] Such factors include the speed at which the new EU Data Protection Regulation is introduced, the obligations set out in that regulation, public pressure, the possibility of exploiting privacy commercially, and so on.

The idea behind stress coaches complies with the intentions of the Dutch Working Conditions Act. After all, stress coaches are meant to estimate risks as accurately as possible and to prevent harm (to an employee's health). Such coaches could be included in an organisation's health and safety policy, for example. Risk assessment is mandatory under the Act, but such assessments explicitly concern risks that arise in the workplace and during working hours. Recommendations to reduce the level of risk can only be work-related and implemented during working hours. Employees may ignore recommendations that extend into their private lives without suffering any negative consequences. Employees are thus also not obliged to use a stress coach if they do not wish to.

An employment court may deviate from the above in exceptional cases. An employer may make demands on an employee in connection with his performance on the job. If the employee's state of health negatively affects his performance, then the employer may require the employee to do something about his physical condition, even if that means taking action in the private domain (for example losing weight). If the employee fails and is unable to do his job as a result, and if it can be demonstrated that there is no other work for him in the organisation, then the employer may be justified in dissolving the employment contract.[18]

Based on the description of the issues addressed above, it appears that the present set of legal instruments (Working Conditions Act, Personal Data Protection Act) basically offer a satisfactory framework for the informed introduction of digital coaches in the workplace. If we work purely on the basis of the statutory framework, then there are

---

17 Various studies indicate that considerable efforts are being made to develop privacy-friendly solutions, but that the need for these solutions and their economic viability are a different matter (Cave et al. 2011).

18 See http://www.akd.nl/nl/kennis/publicaties/obesitas-een-reden-voor-ontslag.

certain guarantees against the misuse or improper use of data. Nevertheless, a recent study shows that three quarters of the employers surveyed did in fact have unlawful access to their employees' medical data.[19] New technologies undoubtedly give rise to new forms of improper use and misuse of data. Much will depend on the type of coach that is used, who uses it, and for what purpose. Digital coaches have options that will make existing forms of misuse easier. One example is a situation in which a supervisor uses data about an employee's health to build a case for the employee's dismissal. It is difficult to say how far such practices will go because they have not yet occurred. It is clear, however, that digital coaches offer opportunities for monitoring that go way beyond those presently available. Employee representatives and occupational physicians can play a role in this respect. The latter play a pivotal role in the relationship between employees and employers in work-related stress issues. Occupational physicians should be supported in this, for example by the Netherlands Society of Occupational Medicine (NVAB) taking a firm position and issuing guidelines for dealing with data produced by digital coaches.[20] Before employees are offered stress coaches, the employer, employee representatives, and occupational physicians should make firm agreements about how digital coaches and the data that they produce will be used, and about informing the employees properly.

Employees can themselves use digital coaches to demonstrate which situations they experience as stress-inducing or burdening. As far as we are aware, this has not yet happened, but there is nothing to exclude the possibility of specific digital coaches being customised for this purpose. Here too, it will be necessary for employers and employees to consult and reach agreement on this specific use of digital coaches.

A final point to consider in a broader context is that employees might in fact find it beneficial to have stress factors at work linked to those in their private lives. They can derive enormous benefits from activities outside the work environment, for example sport, walking, socialising, eating healthily, getting enough rest, and so on. There is a direct connection between feeling good in one's private life and feeling good at work. Someone who is happy at home will be better able to cope in the workplace. Flexible working practices and working conditions are causing the two domains to merge, but this is not necessarily a bad thing. It may also offer more leeway to manage matters, with 'de-stressing' as part of the package. Private coaching firms are already keying into this trend, as one of the above sections has shown, and that could have positive consequences for those involved. Employers need do nothing to capitalise on these

19 http://www.spitsnieuws.nl/archives/binnenland/2013/05/baas-schendt-privacy-werknemer.

20 In response to the news story on Spitsnieuws.nl (see previous footnote), Arbo Unie - a commercial firm that provides health & safety services to businesses - indicated that it had 'complied with the physician's duty to maintain confidentiality and with privacy legislation,' and that it had not been influenced by financial considerations.

advantages, and they can also encourage positive behaviour by offering appropriate options in their package of employment terms.

7.5 Conclusions

Who wouldn't want to have a digital coach to help them reduce work-related stress? Given the rapid pace of technology, it will not be long before we all have access to a digital stress coach. Fortunately, there is still enough time to prepare ourselves for the risks associated with using a tool of this kind. Time enough to think carefully about the purpose of collecting data with a digital stress coach and to consider the conclusions that we can derive from that data. Time enough to investigate all this and to determine whether we can achieve the goals we have set by using a digital stress coach. And there is still enough time for employers and employees to reach agreement on the use of such coaches and on data protection.

The risk of work-related stress remains high, and will only be exacerbated by changes in technology and society. Once we have studied the relationship between physiological measurements and perceived stress thoroughly, determined which conclusions we can and cannot draw from the data, and have firm agreements in place between employers, employees, and occupational physicians about the use of the stress coach and the data that it collects, then the digital stress coach can start making an important contribution to reducing that risk and improving employee wellbeing.

7.6 Recommendations

Research on the effectiveness of digital stress coaches and scientific validation of the methods used will help to draft a set of sound guidelines or quality standards, possibly in the form of a seal of approval, for providers and purchasers of stress coaches. The professionals (coaches and health and safety services) and technology developers ought to play a key role in this.

The Dutch Data Protection Act and the Working Conditions Act offer a satisfactory frame of reference for assessing the risks associated with the use of stress coaches. Before introducing stress coaches, a risk assessment should be performed and measures put into place to counteract the risks that are identified. Various tools offer guidelines in this respect, for example the Data Protection Impact Assessment, Privacy by Design, and giving data subjects more control over their data.

Digital stress coaches can violate a person's physical integrity. An assessment is needed as to whether a stress coach using sensors is necessary to collect data, and whether sensors are not collecting more data than strictly necessary. Employees should always be alerted to their right to privacy.

Employers, employees, health and safety services, and occupational physicians must address the problem of how to deal with the data provided by digital stress coaches. Before stress coaches are offered to employees, the employer, employee representatives, and occupational physicians must reach agreement on how the coaches are to be used, and the employees concerned must be properly informed.

7.7 Bibliography

Aarts, E. & B. de Ruyter (2009). 'New Research perspectives on Ambient Intelligence'. In: *Journal of Ambient Intelligence and Smart Environments* 1, pp. 5-14.

Allvin, M. et al. (2011). Work without Boundaries. *Psychological Perspectives of the New Working Life*. Wiley-Blackwell Sussex.

Cave, J., Robinson, N., Schindler, Bodea, G., Kool, L., Van Lieshout, M. (2011*) Does it help or hinder? Promotion of Innovation on the Internet and citizens' right to privacy*. Study for European Parliament - Industry, Research and Energy (ITRE).

Doornen, L. van (2013). 'De psychofysiologie van werkstress'. In: Schaufeli, W. & A. Bakker (red.). *De psychologie van arbeid en gezondheid*. 3e druk. Houten: Bohn, Stafleu, van Lochum, pp. 65-90.

Gutwirth, S., R. Gellert & R. Bellanova et al. (2011). 'Deliverable D1: Legal, social, economic and ethical conceptualisations of privacy and data protection'. Prescient project: http://www.prescient-project.eu/prescient/index.php.

Houtman, I. et al. (2012). 'Monitoring New and Emerging Risks'. (accepted in 2012).

Houtman, I., L. Koppes & F. Dekker (2013). *Verklaring van de stijging in WIA-instroom. Een werknemers- en werkgevers perspectief*. Hoofddorp: TNO.

Cave, J. et al. (2011). *Does it help or does it hinder? Promotion of innovation on Internet and citizen's right to privacy*. Brussels: European Parliament.

Kool, L., J. Timmer & R. van Est, R. (2013). *Keuzes voor de e-coach. Maatschappelijke vragen bij de automatisering van de coachingspraktijk*. Den Haag: Rathenau Instituut.

Koppes, L.L.J. et al. (2013). *Nationale Enquête Arbeidsomstandigheden 2012. Methodologie en globale resultaten*. Hoofddorp: TNO.

Lieshout, M. van et al. (2011). 'Privacy by Design. An alternative to existing practices in safeguarding privacy'. In: *INFO - The Journal of policy, regulation and strategy for telecommunications, information and media* 13, no. 6, pp. 55-68.

Manyika, J. et al. (2013). *Disruptive technologies. Advances that will transform life, business, and the global economy*. McKinsey Global Institute, May 2013. McKinsey & Company.

Pot, F. et al. (2012). 'Workplace innovation in the Netherlands. Chapter 7'. In: Houtman, I. (ed.). *Work Life in The Netherlands*. [S.l.]: TNO, 2012:173-190.

Wevers, C. & P. Bongers (2013). *Vitaal menselijk kapitaal. Vernieuwing in de relatie tussen arbeid en gezondheid*. Hoofddorp: TNO.

Wiezer, N. et al. (2011). *Exploring the link between restructuring and employee well-being*. Central Institute for Labour Protection-National Research Institute, Warsaw.

Wiezer, N. et al. (2012). *Werkdruk*. Hoofddorp: TNO.

8

General discussion and conclusion

8.1 Objective of this thesis

The objective of this thesis was to explore the potential of persuasive technology to improve health and wellbeing at work. To gain more insight in this potential, the following research questions were addressed: Do persuasive technologies for health and wellbeing at work incorporate theory or evidence-based principles and constructs (Chapter 2)? Which types of research methods are appropriate and useful to evaluate persuasive technologies for health and wellbeing at work (Chapter 3)? What are the effects of persuasive technologies on health and wellbeing at work (Chapter 4 to 6)? What is the societal impact of persuasive technology for health and wellbeing at work (Chapter 7)? This chapter presents an overview of the main findings. The studies in this thesis cover a long period of time and a wide range of subjects. Therefore, reflections on the main results and the implications of this thesis for research, design and application in practice are given in the context of recent related work.

8.2 Overview of the main findings

8.2.1 Theory and evidence base

The first step towards gaining insight into whether persuasive technologies are a powerful medium for delivering interventions in the workplace setting is to assess their consistency with evidence-based practices. Persuasive technologies collect data using sensors (sense), analyse and interpret these data accordingly (reason) to give adequate feedback (action). To do this, they draw on a variety of assumptions, theories and standards. However, it is often unclear whether these technologies are consistent with evidence-based practice, whether theories and models are used and if so, which theories or models are being used. Because persuasive technologies aim to change attitudes or behaviour through persuasion and social influence, it is important to integrate theories and constructs from the behavioural sciences. Research into non-digital interventions shows that interventions indeed are more likely to be effective if they are rooted in health behaviour theories (Abraham and Michie, 2008; West and Michie, 2016). Furthermore, persuasive technology interventions should be grounded in theoretical frameworks related to occupational health and wellbeing, such as stress models, physical workload models and physical activity models (Koldijk et al., 2016). In Chapter 2, a comparative assessment was conducted concerning the incorporation of behaviour change theory in 45 mHealth apps. Results show a limited presence of BCTs in general, limited use of potentially successful combinations of BCTs in apps and use of potentially unsuccessful combinations of BCTs. Because research shows that adding these theories enhances the effectiveness of non-digital interventions, this could also be a promising addition for mHealth interventions. The findings of our study indicate that there is indeed potential for improving apps related to health and wellbeing in the occupational setting by incorporating behaviour change theory in the design. The study in Chapter 2 thus adds to the existing knowledge on incorporating health behaviour change theory in persuasive technologies for the occupational context

8.2.2 Research methods

To gain insight into the potential of persuasive technologies in the occupational setting, it is important to select research methods appropriate to evaluate the effectiveness of these persuasive health technologies. To explore the potential for persuasive technology in the occupational setting, different methods have been applied in this thesis: a combination of laboratory studies (e.g. Chapters 4 and 6) and field studies (e.g. Chapters 3 and 5) have been used, as well as a combination of quantitative (e.g. Chapters 4–6) and qualitative research methods (e.g. Chapter 3). In addition, a review (Chapter 2) and a position paper (Chapter 7) also form part of this thesis. Applying a variety of methods is a strength of this thesis.

In Chapter 3 the added value of three different qualitative research methods was compared: interviews with users, focus groups with users and focus groups with experts. Qualitative methods are needed to gain understanding about how and why a system is or is not used (Klasnja et al., 2011). The results show that the type of evaluator (user or expert) and type of evaluation (interview or focus group) yield different results. Therefore, a qualitative evaluation might benefit from applying a combination of qualitative methods. In addition, it is concluded that factors to consider when selecting a qualitative research method are the design, the development stage of persuasive technology under research, and method of implementation of the mHealth app. The working context in which an mHealth application will be used, employees' mental models (someone's thought process about how something works), feasibility and the available resources also have to be considered. Finally, it is important to take into account which skills are required of experts and users to use a specific qualitative method.

8.2.3 Effectiveness

The effectiveness remains unclear of many persuasive technologies that have appeared on the market, which might hamper their adoption and application in the occupational setting. In Chapters 4, 5 and 6, the effectiveness of a specific persuasive technology (the Hoverstop computer mouse) is evaluated. These chapters showed that the mouse could be used without negatively affecting task performance, and also showed how feedback features influenced the effects. In addition, user satisfaction was evaluated.

Chapter 4 described a laboratory experiment with the Hoverstop computer mouse. The results showed that persuasive technology built into a computer mouse positively affected workers' behaviours. It also decreased static muscle loading of the fore-arm muscles, thereby reducing the risk of developing arm complaints. Neither positive nor negative effects on objectively measured productivity were found. However, subjective ratings did not support these findings: about half of the subjects judged the time necessary to complete a task to be extended compared to a traditional mouse. Another remarkable result was the large variety in user satisfaction using this mouse.

Because the conditions in the laboratory setting were only 15 minutes in duration, a short-term randomized controlled trial (RCT) was performed (Chapter 5) to assess whether the results of the laboratory study in Chapter 4 would hold true in the field. The field study in Chapter 5 also aimed to gain better insight into how users get accustomed to feedback, the effects on task performance and the user satisfaction during the initial phase of working with the mouse. In the field study, positive effects were again found on behaviour, and the technology appeared not to affect performance. The technology also showed a large variety in user satisfaction, in line with the laboratory experiment. The results also showed that participants who indicated that they already experienced physical complaints judged the mouse to be more useful. Apparently, for these users, the mouse seemed more relevant. This result confirms other work that found relevance to be an important factor for user satisfaction and adoption of technology (Kool et al., 2015; Wixom and Todd, 2005).

When applying persuasive technology to change attitudes or behaviour in the working context, it is important not to affect task performance negatively. For this purpose designing unobtrusive technology might offer possibilities, which is the topic of the study in Chapter 6. Four different types of unobtrusive feedback were evaluated – again using the Hoverstop computer mouse – in a laboratory experiment. The results from this study revealed that all types of feedback signal were equally effective to change behaviour, compared with absence of feedback, while task performance was not affected. However, user experience showed differences in user preferences for feedback signals. Tactile feedback was preferred over visual feedback, and a peripheral visual signal was preferred over a transparent visual signal. Although it appears to be difficult to optimize the balance between effectiveness (and thus noticeability) and unobtrusiveness, the results suggest that it is possible to provide effective and efficient feedback with a non-obtrusive signal.

The findings described in Chapters 4, 5 and 6 show that a computer mouse with feedback, as an example of persuasive technology, may have benefits to motivate computer workers to adopt healthy, comfortable and productive work behaviour, at least in the short term. These studies add to existing knowledge of the potential of persuasive technology for health and wellbeing at work. Persuasive technology might have potential as a preventive measure for musculoskeletal complaints, but long-term studies are still needed to link these findings to symptoms and establish clinical relevance.

The results in Chapters 4, 5 and 6 also show a large variety in terms of user satisfaction and preferences, thus revealing the importance of including user preferences in investigating and developing persuasive technology. The importance of studying user preferences was also shown in Chapter 3, which evaluated an mHealth app for health and wellbeing at work. The findings indicated that the extent to which the app addressed users characteristics, motivation and needs was an important factor for users. The relevancy and benefits should also be made clear to the user. The results of Chapter 3 indicate that developing tailored persuasive technology (i.e. technology personalized to the individual user) might increase acceptability and use, and thereby the impact on health and wellbeing.

8.2.4 Societal impact

Many stakeholders could benefit from persuasive technology for health and wellbeing at work. There might be shared benefits such as prevention of occupational risks (i.e. work-related stress) and enhanced health and wellbeing. This might lead to substantial gains in productivity, decreases in sick leave and decreases in healthcare costs (Kool et al., 2015). When applied in the working context, it is therefore not only important to explore the possible effects of persuasive technology on relevant individuals, but also how it may change the processes and relationships on the work floor. Chapter 7 discussed the societal impact of persuasive technology in the working context, how it fit into current regulatory frameworks and where possible ethical problems might arise.

First, Chapter 7 pointed out that data gathered with technology suggest a kind of certainty that employees and employers will act upon. That might be unjustified, because many persuasive technologies lack validity and reliability, which risks presenting an incorrect picture of the monitored data.

Second, using persuasive technologies such as stress coaches in the working context can influence the relationship between employer and employee. For instance, an employee can, using persuasive technology and finding that he is running a health risk, hold his employer accountable. On the other hand, employers can, using persuasive technology, track whether employees stick to agreements that they have made (e.g. to manage work-related stress). In addition, the emphasis placed on self-management by persuasive technology might cause a shift in the perceived responsibility towards the individual employee, although health and wellbeing is a joint responsibility of employer and employee. The Dutch Working Conditions Act obliges employers to pursue a policy focused on prevention. Where this is not possible, the employer must limit work-related risk factors. An employee must follow the employer's instructions concerning health and wellbeing at work and inform the employer if he or she observes possible risk factors for health and wellbeing. Persuasive technology could be introduced by employers as part of their policy. Although this might be beneficial for both employers and employees, serious objections can also be raised. Continuous monitoring is not always limited to working hours (i.e. sleep data or data on physical activity), but might easily interfere with the employees' private life. This might violate the employees' right to privacy or the right to protection of personal data.

Third, more and a greater variety of data become available to add to employees' personnel files. It is important to ensure that medical and other sensitive data is only added to the employee file if doing so is compliant with relevant rules and practices.

Finally, there is the risk for loss of employee autonomy. Persuasive technology might make an employee feel subject to surveillance, even if the use is voluntarily, because private activities might become visible in the working context (i.e. when exercising).

It is therefore important to develop and implement technologies responsibly. Research on the effectiveness and validation of methods will help to draft guidelines or quality standards to direct this. Dutch legislation on data protection and working conditions is currently used as a framework for responsible use of persuasive technology in the work context. Future research is necessary to determine whether this is sufficient.

8.3 Discussion

This thesis shows that persuasive technology may contribute to a reduction of health risks and may improve health and wellbeing. However, there are still challenges that need to be addressed. In this section, the findings of this thesis are put into perspective considering the methodological issues and findings of related research. Implications will be given for future research, design and application in practice.

8.3.1 Implications for research

The studies described in this thesis indicate that persuasive technology is promising, but there is still research needed to develop persuasive technology that is effective in changing behaviour and thereby increasing health and wellbeing at work.

Persuasive technology needs solid theoretical and evidence-based foundation

The first step to gain insight into whether persuasive technologies have potential as interventions in the workplace setting was to evaluate their consistency with evidence-based practices. It has been concluded above that there is as yet very little theoretical foundation in the apps – as an example of persuasive technology – that are now being used within the occupational context. Because persuasive technology aims to change attitudes or behaviour through persuasion and social influence, it is important to integrate theories and constructs from behavioural sciences. According to Orji and Moffatt (2018), the limited integration of behaviour theories and practice probably has to do with the lack of skills among persuasive system designers, who would need to translate the theoretical determinants of behaviour into technology design requirements. Research into non-digital interventions shows that interventions are more likely to be effective when they are rooted in health behavioural theories (Abraham and Michie, 2008; West and Michie, 2016). In the literature, it is cautiously concluded that this also applies to persuasive technology (digital interventions), although research is still limited and further research is needed (Orji and Moffat, 2018).

The findings of this thesis also indicate that the foundation in theoretical frameworks related to occupational health and wellbeing is limited. For instance, do stress management apps make use of evidence-based stress models? Some apps in the study discussed in Chapter 2 gave feedback which was not in line with current scientific insights, making the achievement of behavioural change objectives questionable. These findings suggest that just focusing on the presence of theories on health behaviour change cannot explain the quality or effectiveness of persuasive technologies for the occupational context. In the literature, next to a limited integration of behaviour theories, the lack of use of theoretical frameworks or models related to health and wellbeing in persuasive technology is also acknowledged (Orji and Moffatt, 2018; Van den Broek, 2017).

In summary, a stronger theoretical and evidence-based foundation will likely increase the potential and effectiveness of persuasive technologies. To move this field of research forward, developing a comprehensive framework for translating theoretical determinants into technology design components is suggested (Orji and Moffatt, 2018; Koldijk et al., 2016). Such a framework should address theories and practice related to occupational health and wellbeing, as well as behaviour change, and relevant behaviour change objectives can be defined based on theoretical frameworks related to health and wellbeing. To reach these change objectives, behavioural change theory has to be considered. The extent to which this might be addressed in the same way as non-digital interventions should also be studied. Such a framework might facilitate interdisciplinary collaboration between persuasive technology designers and researchers, which would strengthen the evidence base in this area (Pagliari, 2007).

Persuasive technology needs advanced research methods to assess impact

Better insights are needed in appropriate research methods when evaluating persuasive technologies. To convincingly demonstrate the effectiveness of such technologies for health and wellbeing at work, large-scale, longitudinal studies with control groups have long been considered the 'gold standard' capable of eliciting causal relationships between intervention and outcomes (Klasnja et al., 2011; Kumar et al., 2013; Schelvis et al., 2015; Stawarz and Cox, 2015; West and Michie, 2016). However, persuasive technology has a number of characteristics that pose a challenge when conducting research. First, it is important to ensure that an evaluation method matches the rapid pace of technology development. Second, some persuasive technologies are personalized and adapted to the users' needs. Third, some persuasive technologies continuously adapt to the user as he or she changes his or her behaviour ('just-in-time adaptive interventions'). As a result, persuasive technology is no longer 'one intervention', but 'a lot of different ones', which makes it difficult to draw conclusions on effective intervention ingredients. Therefore, typical research methods such as RCTs are not always a feasible option, because changes to an intervention during evaluation pose a threat to internal validity (Pham et al., 2016). In Chapter 5 of this thesis, a short-term (2 week) RCT was performed, showing positive intervention effects. This RCT was a feasible option, because the persuasive technology under research (a persuasive computer mouse) did not change during the 2-week trial. However, the short time duration of this trial makes it difficult to draw conclusions on sustained behaviour change and/or learning effects and, in line with that, long-term health and wellbeing effects.

It is important to recognize, though, that evaluation involves more than an estimation of effect size (West and Michie, 2016). It also involves understanding how and why a system is (or is not) used, which is critical for the adoption and use of persuasive technology (Klasnja et al., 2011). For this purpose, qualitative research methods are appropriate. The findings of the study in Chapter 3 show that a qualitative evaluation might benefit from a applying a combination of qualitative methods. These findings confirm results from other studies (e.g. Tan et al., 2009; Vermeeren et al., 2010; Zapata et al., 2015). Although this study adds to the existing knowledge on qualitative evaluations of mHealth applications, more scientific research is needed to determine which methods work best for which kind of persuasive technologies and in which context, as well as which methods work well together.

In the literature, factors considered to be important when selecting research methods to evaluate digital interventions include: practicability of the method; available resources to apply a method; and skills or experience required from experts and/or end users (e.g. will experts or end users need instruction or training before participation; Jeffries et al., 1991; Nayebi et al., 2012; Nielsen, 1993; Tan et al., 2009; Zapata et al., 2015). For evaluations in a working context, a particular point of attention is assessing the extent to which the working context poses constraints in research designs. Examples are employee recruitment, the setting in which testing takes place (professional, natural or simulated setting) or interference with ongoing work (Kjeldskov and Graham, 2003). Because evaluation and development of persuasive technology go hand in hand and are highly iterative, the development stage of technology also has to be considered (West and Michie, 2016).

Evaluation preferably starts at the earlier stages of development – as soon as the first concept has been established (concept phase). The basic idea has to be tested, preferably with stakeholders and end users. The development phase typically involves multiple cycles of revision and testing, until a point is reached where it is considered that the intervention can be implemented (West and Michie, 2016). At some point it is likely that evidence needs to be gathered as to whether the intervention has the planned effects and to assess engagement, usability and any side effects. It is best to do this initially in a pilot study or efficacy trial, under ideal conditions such as in laboratory (West and Michie, 2016; Tomlinson et al., 2013). Chapters 4 and 6 of this thesis described examples of such studies. Eventually assessment may take place in a full scale effectiveness trial under real-life conditions (West and Michie, 2016; Tomlinson et al., 2013), as done in Chapter 5 of this thesis. Following implementation, further testing might be done for optimization. The ideal evaluation strategy contains both steps in which user satisfaction is examined (and improved) and steps by which effectiveness is examined (and improved). In the literature, several suggestions for alternative methods for continuously evolving or adaptive technology can be found and will be discussed below.

Qualitative research methods might be applied even in the concept phase and can be repeated during the iterative development process (West and Michie, 2016). Examples of possible methods to apply are interviews or focus groups (e.g. guided, open, interactive discussion, facilitated by a researcher). These methods were used in the studies discussed in Chapter 3, which showed that using different types of evaluators and evaluation methods are complementary, which might be advantageous for the evaluation process. Next to focus groups or interviews, 'think aloud' (asking users to articulate their thoughts as they use [a part of] an intervention or mock-up) or 'observation of use' (observing users and recording their behaviour as they interact with the intervention) are examples of qualitative research methods that might be useful (Gulati, 2012; Jaspers, 2009; West and Michie, 2016; Zapata et al., 2015). 'Dog-fooding', which involves developers using their own products to identify bugs and gain a personal sense of whether it is achieving its goals, is a typical method applied in the optimization phase (West and Michie, 2016).

An example of a quantitative method is an N-of-1 study. The multiple cross-over single-subject design of an N-of-1 study is an experimental variant of a pre-post-test design: data are collected over a period of time at frequent intervals, and temporal trends are tested as a function of the introduction of changes to the intervention; they may be systematic or randomized (Kravitz et al., 2014; Kumar et al., 2013; West and Michie, 2016). N-of-1 trials are therefore an appropriate method that could even be used in a development phase. They can be used to test a proof of concept without the need of setting up a group comparison study or for studies in which randomization is not possible. N-of-1 trials are also suitable for development and evaluation of personalized interventions, because the effects of different features of an intervention might be tested (Kravitz et al., 2014; Kumar et al., 2013; West and Michie, 2016). For studies with multiple users, A-B testing might offer an option: sequential A-B testing (establishing a baseline for a key set of variables and then making a change to the intervention and determining the effects) or concurrent A-B testing (giving different versions of an intervention to different groups of users and establish what difference this makes; West and Michie, 2016).

Stepped-wedge design is an appropriate quantitative research design if the interventions are going to be implemented with all users and if it is not feasible to scale all at once. This method operates as a series of waiting lists and randomizes the order in which groups receive the intervention. The intervention group can be compared with both their pre-test measures and with measures from other subjects who have not yet received the intervention. The stepped-wedge design allows for improvement of the intervention based on lessons learned in previous steps (Kumar et al., 2013; Schelvis et al., 2015).

Another method for developing, improving and evaluating persuasive technology is the Multiphase Optimization Strategy (MOST), which is used to identify the most promising components of persuasive technology in a screening phase and to test these using ANOVA. Promising components can then be evaluated in a randomized trial (Collins et al., 2007; Kumar et al., 2013). MOST also incorporates the standard RCT, but before the RCT is undertaken, MOST uses a principled method for identifying which components of the intervention and which levels of the components lead to the best outcomes (Collins et al., 2007).

One of the promises of persuasive technology is the potential to personalize and adapt the intervention to the users' needs. To accomplish this, a better understanding of within-subject differences and the effects of mediating variables on outcomes is required. A research method that meets this need and has been developed for building time-varying adaptive interventions is sequential multiple assignment randomized controlled trials (SMART; Collins et al., 2007; Kumar et al., 2013; West and Michie, 2016). With SMART, individuals are randomized to receive different choices of features. Then, according to a decision rule, those who respond well are allocated to a new comparison, as are those who respond less well or not at all. This process may continue to identify efficient and effective features (tailoring variables) for different groups. The end goal of SMART is the development of evidence-based adaptive intervention strategies, which are then evaluated in a subsequent RCT (Kumar et al., 2013; West and Michie, 2016). However, there are limitations to this approach. Increasingly, persuasive technology will adapt in *real time* to users' needs (i.e. 'just-in-time adaptive intervention'), which means interventions are dynamically personalized during use and each individual receives essentially a different intervention. To evaluate this kind of intervention, other research methods are needed.

In summary, the methodology for evaluating the impact of persuasive technology is still at an early stage of development, and many studies do not provide a clear indication as to whether the intervention being evaluated was effective or which components determined effectiveness (Michie et al., 2018; West and Michie, 2016). These considerations underline the need for application of relevant and timely evaluation methods to advance the current state of the evidence (Kumar et al., 2013; Michie, 2016; Schelvis et al., 2015; West and Pham et al., 2016). Future research should aim to determine which methods work best for which kinds of persuasive technology and for which context, as well as which methods work well together. A few research designs that might be useful to evaluate persuasive technology have been discussed. Application of these methods to assess the impact of persuasive technology for work health and wellbeing is recommended.

Persuasive technology has potential but lacks evidence

Insight is needed as to whether persuasive technology is indeed a powerful medium for delivering interventions at work. Chapters 4 to 6 of this thesis add to the existing body of evidence showing that persuasive technology positively influences workers to adopt healthy work behaviour, at least in the short term, which might provide an effective contribution to prevention. These results are in line with research that can be found in the literature. For instance, based on their empirical review, Orji and Moffatt (2018) concluded that persuasive technologies are effective at promoting various health and wellness–related behaviour, with 92% of all the reviewed studies reporting some positive outcome (fully and partially positive). Their study also underlines the need for advanced research methods. A closer look regarding the data analysis methods used in these reviewed studies shows that 46% employed a mixed method evaluation combining both quantitative and qualitative approaches. This was followed by the quantitative approach, accounting for 39%. The most commonly used approach for collecting quantitative data was setting out a questionnaire/survey. A fully qualitative approach was the least applied with only 15% of all the studies using this approach. The most frequently used qualitative methods were interviewing, focus-group discussion and observation of participants' behaviours and persuasive technology use. Furthermore, only a few of the reviewed studies conducted longitudinal evaluations of the effectiveness of their persuasive technology, and the majority of the studies did not conduct a follow-up study beyond the initial (feasibility) study. Therefore, Orji and Moffatt (2018) concluded that it would be difficult to establish the long-term effects of persuasive technology for health and wellness from existing studies, and persuasive technology would benefit from research into objective evaluation approaches. This once more underlines the need for advanced research methods, as discussed above.

The review of Orji and Moffatt (2018) included persuasive technology for health and wellness and not specifically for the occupational context, however, making it more difficult to draw conclusions on the effectiveness of persuasive technology in the work setting. A few examples of research on persuasive technology for the occupational context can be found in the literature. For example, studies on rest break software showed mixed effects on health (Slijper et al., 2007; Van den Heuvel et al., 2003). Van Drongelen et al. (2014) showed positive effects of a tailored mHealth intervention on physical activity, snacking behaviour and sleep among airline pilots. In their review on persuasive technologies to reduce prolonged sedentary behaviour at work, Wang et al. (2018) included eight studies. In four studies, positive effects on reducing sedentary behaviour at work where found in at least one outcome compared to control groups. Three studies showed no differences or very small differences with control groups, and one study found a small increase of sedentary behaviour. Based on their reviewed studies, Wang et al. (2018) found that reminders were the most frequently used persuasive designs; their findings showed that prompts or reminders alone had no significant effect on reducing sedentary behaviour at work, while the combination of such reminders with education or other informative session seemed to be more promising.

The findings on user experiences in the experimental studies with the computer mouse in Chapter 4 to 6 showed that, although positive effects on behaviour change can be seen, poor usability can be problematic for adoption and use. When persuasive technology is not adopted or used, it will be impossible to reach positive outcomes for health and wellbeing.

To summarize, currently available research into persuasive technology specifically applied within the occupational setting shows potential but is still limited. In this thesis only a few examples of persuasive technology applications were subject of research, such as apps and a computer mouse. We should be careful about generalizing results to other persuasive technology applications and to general conclusions on persuasive technology. In addition, only short-term studies were performed in this thesis, which is therefore just a starting point that leaves many questions unanswered. This thesis does, however, show that it is important to continue research in this area. There is a lack of long-term scientific evaluations of the effectiveness of persuasive technology, which is acknowledged in literature (Albrecht, 2016; Orji and Moffatt, 2018; Van den Broek, 2017). As discussed above, an important reason is that the characteristics of persuasive technology challenge the way research is conducted; applying advanced research methods might contribute to building an evidence base. Further research should aim at long-term studies to assess behaviour change and learning effects using such methods over time. In addition, future evaluations should look into possible unintended side effects of persuasive technology for the occupational context (Pagoto and Bennett, 2013; Stacey et al., 2018; West and Michie, 2016).

8.3.2 Implications for design and development

The ultimate goal of the persuasive technology studied in this thesis is to provide solutions that stimulate health and wellbeing for workers. Persuasive technologies need to be not only effective (does it actually work?), but should be meaningful as well: is it usable? Does it meet users' needs? Do users use it (Stawarz and Cox, 2015)?

Persuasive technology needs improvements to enhance satisfaction and technology acceptance

It is impossible to reach the desired positive outcomes for health and wellbeing when persuasive technology is not adopted or used. Research shows that uptake and adherence are low, even where there is evidence for effectiveness (West and Michie, 2016). Therefore, engaging users remains a major challenge.

Although Chapter 2 was not a usability evaluation of mHealth apps, apps with reasonable or good usability characteristics were reviewed, but apps with poor usability characteristics were found as well. Using the apps ourselves, we came across apps with system failures, apps in which it was difficult to find certain content or apps were the accuracy of measurements was open to serious doubt. The measures on user experiences in the experimental studies with the computer mouse in Chapter 4 to 6 also showed that poor usability could be a real showstopper. It has long been known in the literature that usability is important for the adoption and acceptance of technology (Wixom and Todd, 2005). The studies with the computer mouse (Chapters 4–6) showed that large variations exist between users on the preferences for the properties of and feedback provided by the persuasive technology tested. Also, from the mixed-method qualitative study in Chapter 3, in which an mHealth app was evaluated, several lessons can be learned. The following factors that influence user satisfaction and technology acceptance should be taken into account for the design and development of persuasive technology:

› **Technology**
  System failures or poor system performance might negatively influence adoption and adherence. Accuracy (actual as well as perceived – i.e. users' perception that the information is correct) largely influences the confidence of users in the

application and thereby influences its use. The quality of feedback influences the possibility to reach behaviour change and, in line with that, it influences health and wellbeing effects. This is in line with the work of Wixom and Todd (2005).

› **User characteristics**
The relevance and benefits of persuasive technology should be made clear to the (potential) user within the application itself, as well as in communication guiding the implementation of the application. This was substantiated in Chapter 5, which revealed differences in satisfaction with technology between users with physical complaints, for whom the technology appeared to be more relevant than for users without physical complaints. Often persuasive technology fails to apply personalisation techniques (Manyika et al., 2015; Schwab, 2016). Persuasive technology should address users' characteristics (age, condition, health, function, [work] activities), motivation and needs (e.g. occupational health [risks] and wellbeing) to increase acceptability and use. This is in line with studies by, for example, Koldijk et al. (2016), West and Michie (2016) and Stacey et al. (2018). Furthermore, the opportunity in developing persuasive technology is to design systems that continuously adapt to individual worker preferences to gain a maximum effect, and the challenge is to determine how to personalize these systems accordingly. Therefore, a next step in developing persuasive technology should be using machine learning and learning algorithms to tailor the technology to user characteristics on the basis of information gathered as the person uses it (West and Michie, 2016). Another issue for just-in-time adaptive interventions is determining when to intervene and how to be proactive. This is a major area of study that persuasive technology designers should consider (West and Michie, 2016). Finally, once implemented, it is advisable to study which characteristics of persuasive technology are important to ensure users adhere to a persuasive intervention.

› **Context**
It is advised to take the work-context in which the technology is being used into account, not only because it influences the possibilities of using persuasive technology in general (e.g. working in clean rooms, operation rooms, oil platforms) or the accuracy of the measurements (e.g. the working environments of firefighters can affect sensors), but it is also important to study user preferences for use in his/her specific working context. Coursaris and Kim (2011, page 130) also suggest designing interfaces and applications that fit to particular contextual settings, while being flexible to accommodate others: 'focus beyond the interface when developing applications'. In addition, implementation plays a large role in the adoption and use of an application: this should be planned carefully, considering communication about the intervention, how and to what extent the

management should be involved or whether an application has to be embedded within a larger intervention programme. These are important factors influencing adoption and use (see e.g. Michie et al., 2016). Research on the implications of persuasive technology in the working context is very limited and should be further developed.

› **Privacy/autonomy in data management**
It is important to take users' autonomy into account for application use and data management and privacy, such as in settings, defining goals and feedback methods. Employees should always be alerted to their right to privacy and, before applications are offered, employees must be properly informed. In addition, within an organization, it should be made clear what happens with the data. Users should be given ownership and autonomy in data management. Various approaches might be used already in early stages of development of technology, such as privacy by design (Schaar, 2010) or privacy impact assessment (Koldijk et al., 2014).

In summary, the following recommendations are important for the development and design of persuasive technology: first, system performance influences adoption and adherence and should be tested thoroughly before implementation. Second, the relevance and benefits of persuasive technology should be clear to the user and should address users' characteristics within the application itself, as well as in the communication guiding the implementation of the application. Designers should engage with the application of tailoring techniques, such as machine learning and learning algorithms, and evaluate when to intervene and how proactive to be. Future research should focus on which characteristics of persuasive technology are important to ensure users adhere to a persuasive intervention after implementation and optimize these characteristics accordingly. Third, persuasive technology should be designed to fit to particular occupational settings. To ensure a good fit, using participative design approaches with end users and to test persuasive technology in real-life settings is advised. Finally, employees should be alerted to their right to privacy and use of personal data, preferably in early stages of development. Approaches that might be used are privacy by design or privacy impact assessment.

Persuasive technology design and development requires multidisciplinary collaboration

Chapter 2 showed that mHealth apps make limited use of theories on health behaviour change. A general lack of theoretical constructs included in persuasive technology (such as apps) might not be entirely unexpected given that app developers' expertise relates to software development and might not include health behaviour theory or theories on mental and physical health (Pagliari, 2007; Orji and Moffatt, 2018). To increase potential effectiveness, it is necessary to incorporate behaviour change theory and provide content with a solid evidence base, which in turn would increase the behaviour change potential of applications. The findings in Chapter 2 suggest a need for multidisciplinary collaboration between app developers, health behaviour change professionals and experts on physical and mental health: combinations of expertise could provide higher quality persuasive technology (Koldijk et al., 2016).

Others subscribe the necessity to collaborate in multidisciplinary teams to reflect on the use of theoretical models in the development of persuasive technology, which could foster persuasive technology superior in theory, which would result in better health and wellbeing outcomes (e.g. Cowan et al., 2013; Direito et al., 2014; Pagoto and Bennett, 2013). In addition, as shown in Chapter 3, it is also important to involve the targeted workers in the development and implementation of persuasive technology. This is the only way to develop persuasive technology that meets the needs and demands of those who will use it later (Albrecht, 2016). Many possible methods are available to cooperate in multidisciplinary design/development teams that include the end users. It is recommended that, for the design and development of persuasive technology, participatory design approaches be adopted to enable the involvement of the target population. This is also stressed by Stacey et al. (2018) in their European 'Foresight study on new and emerging occupational health risks associated with digitalisation by 2025'. Examples of participatory design approaches are context mapping or user observations (Orji and Moffatt, 2018; Van Boeijen et al., 2013).

To conclude, multidisciplinary collaboration in the design and development process is crucial to increase the potential of persuasive technology for health and wellbeing in the occupational setting. To support these processes, making use of participatory design approaches involving all stakeholders, including end users, is recommended.

8.3.3 Implications for the occupational practice

Several practical enablers and barriers exist that stimulate or hamper adoption of persuasive technology in the context of work. Certain conditions must be met and obstacles need to be overcome to exploit the potential of this technology.

Persuasive technology needs standards and guidelines

Developments in technology move fast, which means that now is the time to think about how we can guide the introduction and implementation of technologies for health and wellbeing at work in a responsible matter. Questions that have to be answered are: how far can we permit technology to go in influencing our behaviour? Can employers and employees trust the persuasive technology? Who is actually profiting from the collected data? Is the user informed about all this (Kool et al., 2015)? For this, standards and regulations might offer a solution.

First, as discussed in Chapter 7 of this thesis, persuasive technology interventions are being developed without a clear way for employers or employees to establish which ones are reliable and provide evidence to support the claimed benefits. The limited amount of evidence-based results concerning efficacy, effectiveness and implementation strategies of persuasive technologies raises concerns about the ability of the involved stakeholders (e.g. employees, employers, insurers and occupational health and safety services) to assess its usefulness. This could limit the effective uptake of persuasive technology to benefit the physical and mental health of workers (Albrecht, 2016; Kool et al., 2015; Tomlinson et al., 2013). In addition, employers and employees might have limited ICT skills and may be barely familiar with the criteria to ensure the quality of persuasive technology (Das et al., 2020).

Second, it is important to have insight into whether persuasive technologies can be trusted, and there is a lack of evidence on their quality and reliability. For instance, persuasive technology might create a semblance of precision or false security (e.g. reporting the exact number of calories burnt) which could influence the recommendations to the user. How are these data being judged by the user? Is this a question of creating problems that do not exist? Are our ideas about our performances on which we base the system measurements harmful for our mental and physical wellbeing? Careless interpretation might worry users unnecessarily (false positives) or ease their minds while they should worry (false negatives). Moreover, it is questionable whether there is sufficient room for personal differences and whether algorithms take this into account, as discussed in Section 8.3.2. Finally, humans can only be captured in data to a limited extent, and one of the reasons for this is the difficulty in making complex matters such as health or stress measurable (Das et al., 2020). To what extent can human behaviour or health be predicted? There is a risk that the focus will only be on measurable health elements and other aspects might be lost sight of. Therefore, realism about predictive value or found connections is important (Das et al., 2020).

Third, monitoring technology changes work processes and work relations, as discussed in Chapter 7 of this thesis. In some cases, the organization or even the technology provider receives reports with aggregated, anonymized data. Technology has increased the possibilities for workers to take responsibility concerning their health and wellbeing at work (self-management/do-it-yourself mentality) and gives employers a more motivational and facilitating role. The employer might use the aggregated data of individuals to better understand the emergence of symptoms and effects of interventions. This might lead to worker empowerment. However, does the worker have a real choice to use or not use applications that are offered by the employer? This might negatively affect cooperation, sharing of responsibilities and the professional autonomy of the employee. While it has to be noted that autonomy is not a characteristic of the system, but rather of how it is deployed (Das et al., 2020; Kool et al., 2015).

Fourth, as Chapter 7 discussed, the continuous gathering of information raises questions about safeguarding privacy and the responsible use of personal data. This is especially the case because emerging technologies often cross the boundaries between work and private life, such as measurements of workers' sleep quality (Das et al., 2020; Kool et al., 2015). Also, employers – as well as employees – are not always in the position to take relevant security measures by themselves to protect their own or others' data (Das et al., 2020; Kool et al., 2015). Responsible use of persuasive technology therefore requires critical reflection, not only for reasons of privacy, but also because it can have consequences for tasks, processes and relationships in the occupational context.

Currently, Dutch legislation on data protection and working conditions is now used as a framework for responsible application of persuasive technology in the occupational context, as mentioned in Chapter 7 of this thesis. This concerns the right of privacy (laid out in the Universal Declaration of Human Rights and in the European Union's Charter of Fundamental Rights), the right to data protection (General Data Protection Regulation [GDPR]) and the Working Conditions Act (rules for healthy and safe work and responsibilities for employer and employee). Future research is necessary on whether this legislation is adequate to cover the risks of persuasive technology. The Dutch Personal Data Authority and the Inspectorate of the Ministry of Social Affairs and Employment might jointly provide clarification (Das et al., 2020).

For employers and employees it still remains unclear whether their rights to privacy and autonomy are respected by persuasive technologies and how to assess that situation. This might be remedied through appropriate guidance such as setting standards and providing assessment guidelines. In addition, recommendations might help developers to mitigate risks on quality, user autonomy, privacy and data security (Albrecht, 2016; European Commission, 2014; Kool et al., 2015; Schaar, 2010). Developers of persuasive technology should use privacy-enhancing strategies in the early stages of development (e.g. privacy by design). Developers should also be transparent about the methods they use to persuade users to change behaviour (Kool et al., 2015). Employers and employees might benefit from a regulatory framework to clarify liabilities and responsibilities related to such new systems (Stacey et al., 2018). The government, funders, private organizations and academia must cooperate to set these standards (Manyika, 2015; Stacey et al., 2018; Tomlinson et al., 2013).

It is therefore recommended that assessment criteria or a seal of approval that provides end users with information about the purpose and the target groups of persuasive technology – what it is (not) capable of, whom it is (and is not) intended for and what purpose it serves – be developed. In addition, this seal of approval should provide end users with information on the effectiveness of the technology, what evidence base is available, how it arrives at its recommendations and what methods are being used. The latter explicitly relates to unobtrusive and persuasive technologies that influence users subconsciously.

Examples of such initiatives already do exist. An example is the EU Expert group on mHealth assessment guidelines. In February 2016, the European Commission appointed a working group to draft mHealth assessment guidelines. The group included representatives of patients, citizens, health professionals and providers, payers, industry, academia and public authorities. The group sought to provide common quality criteria and assessment methodologies that could help different stakeholders, in particular end users, in assessing the validity and reliability of mobile health applications.

Currently, ISO (the International Organization for Standardization) is finalizing a new standard on the quality and reliability of health and wellness apps, prepared by the Technical Committee CEN/TC 251, Health informatics, in collaboration with experts from the standardization organizations. It is due to be completed in 2021 (prCEN ISO/TS 82304-2). This standard applies to health apps, which are a special form of health software and do not fall under medical devices regulations. It provides quality criteria for health apps and defines a quality label to visualize the quality and reliability of such apps. It will include a set of quality criteria and cover the app project life cycle through development, testing, releasing and updating of an app, including native, hybrid and web-based apps, apps associated with wearables, ambient and other health equipment and apps that are linked to other apps. It will also address fitness for purpose and the monitoring of usage. It is intended for use by app developers. Often, developers might not have prior experience in a health informatics context, so the standard might compensate for this by providing a set of requirements that have been through a rigorous standards development process. End users, professionals and the wider public will be able to use the quality label when recommending or selecting a health app for use (see http://www.ehealth-standards.eu/quality-reliability-for-health-and-wellness-apps/).

To conclude, practical and ethical barriers exist that might hamper the adoption of persuasive technology in the context of work. Employers and employees need to be aware of the effects persuasive technology might have on the employer-employee relationship. An open discussion between these groups on this subject would be advisable. In addition, there is a lack of clarity on the validity of the persuasive technology, efficacy, effectiveness and implementation strategies, whether they can be trusted and whether privacy and use of personal data are guaranteed. Employers and employees might benefit from a regulatory framework to clarify the liabilities and responsibilities in relation to new systems. It is advised to evaluate whether current legislation on working conditions, privacy and data protection provide a sufficient framework to cover the risks of persuasive technology. The Dutch Personal Data Authority and the Inspectorate of the Ministry of Social Affairs and Employment might jointly provide clarification. It still remains unclear how employers and employees can assess whether the rights to privacy and autonomy are respected by persuasive technology, which means that appropriate guidance is necessary. The development assessment criteria or a seal of approval is therefore recommended, particularly criteria that provide end users with information about the purpose and the target groups of persuasive technology, what it is (not) capable of, whom it is (and is not) intended for and what purpose it serves. In addition, this seal of approval should provide end users with information on the effectiveness of the technology, what evidence base is available, how it arrives at its recommendations and what methods are being used. Currently, initiatives by organizations such as the ISO are being taken to meet this need.

8.4 Concluding statements

The studies in this thesis explored the potential of persuasive technology for self-management in the context of health and wellbeing at work. The thesis added to existing knowledge by reviewing the theory and evidence base in existing applications. In addition, insight has been provided in using qualitative methods to assess persuasive technology. Furthermore, this thesis provided insights into the short-term effectiveness of persuasive technology on behaviour, health-related outcomes, performance and usability. Finally, this thesis explored the societal impact of persuasive technology in the work context. The results provide implications for further research, design, development and practice. Further research is necessary to develop effective persuasive technology–based interventions. Given the current status of persuasive technology and its supporting theories and evidence base, it can be concluded that the application of persuasive technology within the occupational context is in its early stages and huge challenges remain. However, we need to address these challenges to keep up with the fast developments of technology and the changes in work and workers to preserve – or preferably to enhance – the health and wellbeing of workers. However, such technological change is already affecting the occupational context and will become even more relevant in the future.

De technologie staat voor niets en de mens voor de keuze'

[Technology stands for nothing and man stands for choice]

-Loesje-

8.5 Literature

Abraham, C. & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology*; 27(3), 379-87.

Albrecht, U.-V. (Ed.) (2016). *Chances and risks of mobile health apps (CHARISMHA)*. Hannover: Hannover Medical School.

Direito, A., Pfaeffli Dale, L., Shields, E., Dobson, R., Whittaker, R., & Maddison, R. (2014). Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? *BMC Public Health, 14,* 646.

Collins, L.M., Murphy, S.A., Stretcher, V. (2007). The Multiphase Optimization Strategy (MOST) and the Sequential Multiple Assignment Randomized Trial (SMART). *Am J Prev Med;* 32(5S): S112-S118

Coursaris, C.K., & Kim, D.J. (2011). A meta-analytical review of empirical mobile usability studies. *Journal of Usability Studies, 6*(3), 117-171.

Cowan, L.T., Van Wagenen, S.A., Brown, B.A., Hedin, R.J., Seino-Stephan, Y., Hall, P.C., & West, J.H. (2013). Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. *Health Educ Behav, 40,* 133-139.

Das, D., R. de Jong en L. Kool, m.m.v. J. Gerritsen (2020). *Werken op waarde geschat - Grenzen aan digitale monitoring op de werkvloer door middel van data, algoritmen en AI.* Den Haag: Rathenau Instituut (Dutch publication).

Direito, A., Pfaeffli Dale, L., Shields, E., Dobson, R., Whittaker, R., & Maddison, R. (2014). Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? *BMC Public Health, 14,* 646.

Gulati, A., Kumar Dubey, S. (2012). Critical analysis on usability evaluation techniques. *International Journal of Engineering Science and Technology.* 4(3): 990-997.

Jaspers, M.W.M. (2009). A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *International Journal of Medical Informatics, 78,* 340-353.

Jeffries, R.J., Miller, J.R., Wharton, C., & Uyeda, K.M. (1991). User interface evaluation in the real world: a comparison of four techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New Orleans, LA, April 29-May 2* (pp. 119-124). New York: ACM.

Kjeldskov, J., & Graham, C. (2003). A review of mobile HCI research methods. In L. Chittaro (Ed.), *Proceedings of the 5th International Mobile HCI 2003 Conference* (pp. 317-335). Germany: Springer-Verlag.

Klasnja, P., Consolvo, S., & Pratt, W. (2011). How to evaluate technologies for health behavior change in HCI research. *CHI 2011*, May 7-12, Vancouver, Canada.

Koldijk, S., Koot, G., Neerincx, M., & Kraaij, W. (2014). Privacy and trust in context-aware systems. In V. Dimitrova, T. Kuflik, D. Chin, F. Ricci, P. Dolog, & G.J. Houben (Eds.), *22nd International Conference, UMAP 2014, Aalborg, Denmark, July 7-11, 2014. Proceedings* (pp. 134-145). Cham (etc.): Springer. LNCS 8538.

Koldijk, S., Kraaij, W., & Neerincx, M.A. (2016). Deriving Requirements for Pervasive Well-Being Technology From Work Stress and Intervention Theory: Framework and Case Study. *JMIR Mhealth Uhealth, 4*(3), e79.

Kool, L., Timmer, J., & Van Est, R. (Ed.) (2015). *Sincere support, the rise of the e-coach*. The Hague: Rathenau Instituut.

Kravitz, R.L., Duan, N., Vohra, S., Li, J., the DEcIDE Methods Center N-of-1 Guidance Panel (2014). Introduction to N-of-1 Trials: Indications and Barriers. In: Kravitz RL, Duan N, eds, and the DEcIDE Methods Center N-of-1 Guidance Panel (Duan, N., Eslick, I., Gabler, N.B., Kaplan, H.C., Kravitz, R.L., Larson, E.B., Pace, W.D., Schmid, C.H., Sim, I., Vohra, S.). *Design and Implementation of N-of-1 Trials: A User's Guide.* AHRQ Publication No. 13(14)-EHC122-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2014: Chapter 1, pp. 1-11.

Kumar, S., Nilsen, W.J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., Riley, W.T., Shar, A., Spring, B., Spruijt-Metz, D., Hedeker, D., Honavar, V., Kravitz, R., Lefebre, R.C., Mohr, D.C., Murphy, S.A., Quinn, C., Shusterman, V., & Swendeman, D. (2013). Mobile health technology evaluation: the mhealth evidence workshop. *American Journal of Preventive Medicine, 45*(2), 228-236.

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., & Marrs, A. (2013). *Disruptive technologie: Advances that will transform life, business, and the global economy*. [S.l.]: McKinsey Global Institute.

Michie, S., West, R., Sheals, K., Godinho, C.A. (2018). Evaluating the effectiveness of behavior change techniques in health behavior: a scoping review of methods used. *Transl Behav Med*, Mar 1;8(2):212-224. doi: 10.1093/tbm/ibx019.

Nayebi, F., Desharnais, J.-M., & Abran, A. (2012). The state of the art of mobile application usability evaluation. *IEEE 25th Canadian Conference on Electrical and Computer Engineering (CCECE), April 29-May 02 2012, Montreal, QC, Canada*. DOI: 10.1109/CCECE.2012.6334930.

Nielsen, J. (1993). *Usability engineering*. London: Academic Press.

Orji, R., Moffatt, K. (2018). Persuasive technology for health and wellness: state-of-the-art and emerging trends. *Health Informatics Journal*, 24(1), 66-91. Doi: 10.1177/1460458216650979.

Pagliari, C. (2007). Design and Evaluation in eHealth: Challenges and Implications for an Interdisciplinary Field. *J Med Internet Res*; 9(2):e15. DOI: 10.2196/jmir.9.2.e15

Pagoto, S., & Bennett, G.G. (2013). How behavioral science can advance digital health. *Transl Behav Med, 3,* 271-276. DOI: 10.1007/s13142-013-0234-z.

Pham, Q., Wiljer, D., Cafazzo, J.A. (2016). Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. *JMIR Mhealth Uhealth* Sep 09;4(3):e107 doi: 10.2196/mhealth.5720.

Schaar, P. (2010) Privacy by design. *Identity in the information society, 3*(2), 267-274.

Schelvis, R.M.C., Oude Hengel, K.M., Burdorf, A., Blatter, B.M., Strijk, J.E., & Van der Beek, A.J. (2015) Evaluation of occupational health interventions using a randomized controlled trial: challenges and alternative research designs. *Scan J Work Environ Health, 41*(5), 491-503.

Slijper, H.P., Richter, J.M., Smeets, J.B.J., & Frens, M.A. (2007). The effects of pause software on the temporal characteristics of computer use. *Ergonomics, 50*(2), 178-191.

Stacey, N., Ellwood, P., Bradbrook, S.l, Reynolds, J., Williams, H., Lye, D.(2018). *Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025.* European Risk Observatory Report, Luxembourg: publications office of the European Union. Doi:10.2802/515834.

Stawarz, K., & Cox, A.L. (2015). Designing for health behavior change: HCI research alone is not enough. *CHI 2015*, April 18-23, Seoul, Republic of Korea. Workshop on HCI and Health: Advancing health and wellness technology research in home and community settings.

Schwab, K. (2016). *The fourth industrial revolution*. Geneva: World Economic Forum.

Tan, W., Liu, D., & Bishu, R. (2009). Web evaluation: Heuristic evaluation vs user testing. *International Journal of Industrial Ergonomics, 39*(4), 621-627.

Tomlinson, M., Rotheram-Borus, M.J., Swartz, L., & Tsai, A.C. (2013). Scaling up mHealth: where is the evidence? PLoS Med, 10(2), e1001382. DOI: 10.1371/journal.pmed.1001382.

Wixom, B.H., & Todd, P.A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information systems research, 16*(1), 85-102.

Van Boeijen, A., Daalhuizen, J., Zijlstra, J., & Van der Schoor, R. (Eds.) (2013). *Delft Design Guide.* Amsterdam: BIS Publishers.

Van den Broek, E.L. (2017) *Monitoring technology: The 21st century's pursuit of well-being?* Report number: EU-OSHA Discussion paper, affiliation 21860. DOI: 10.13140/RG.2.2.26713.01123.

Van den Heuvel, S.G., De Looze, M.P., Hildebrandt, V.H., & Thé, K.H. (2003). Effects of soft¬ware programs stimulating regular breaks and exercises on work-related neck and upper-limb disorders*. Scan J Work Environ Health*, 29(2), 106-116.

Van Drongelen, A., Boot, C.R.L., Hynek, H., Twisk, J.W.R., Smid, T., & Van der Beek, A.J. (2014). Evaluation of an mHealth intervention aiming to improve health-related behavior and sleep and reduce fatigue among airline pilots. *Scand J Work Environ Health, 40*(6), 557-568.

Vermeeren, A.P.O.S., Lai-Chong Law, E., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, October 16-20, Reykjavik, Iceland* (pp. 521-530). New York: ACM.

Wang, Y., Wu, L., Lange, J-P., Fadhil, A., Reiterer, H. (2018). Persuasive technology in reducing prolonged sedentary behavior at work: a systematic review. *Smart Health* 7-8: 19-30. Doi: 10.1016/j.smhl.2018.05.002.

West, R., & Michie, S. (2016). *A guide to development and evaluation of digital behavior change interventions in healthcare*. London: Silverback Publishing.

Wixom, B.H., & Todd, P.A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information systems research, 16*(1), 85-102.

Zapata, B.C., Fernandez-Alemán, J.L., Idri, A., & Toval, A. (2015). Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst, 39,* 1. DOI 10.1007/s10916-014-0182-2.

Summary

The objective of this thesis was to explore the potential of persuasive technology in the context of health and wellbeing at work. Developments in technology have brought about many changes in work, and these changes will continue as technologies evolve at an exponential pace. Up until now, technology has made our work easier, faster and more efficient. At the same time, technological developments have also posed new risks; despite developments in technology and work, many workers suffer from occupational health problems. With the ageing population it is necessary to rethink current approaches to enhance health and wellbeing at work.

Persuasive technology refers to interactive systems developed to change the attitudes or behaviours – or both – of users through persuasion and social influence. It shows real potential to drive improvements in working life, to reduce or better manage risk factors. However, particularly for the working context, there is still a lack of insight about when, where and for whom persuasive technology is effective and how this might be evaluated.

In this thesis, the potential of persuasive technology for health and wellbeing at work was addressed alongside four scientific challenges: (1) to what extent do persuasive technology applications incorporate theory or evidence-based principles and constructs; (2) which types of research methods are appropriate for evaluating persuasive technologies for health and wellbeing at work; (3) what are the effects of persuasive technologies on behaviour, health and wellbeing at work; and (4) what is the societal impact of persuasive technology for health and wellbeing at work.

Theory and evidence base

The first step to gain insight into whether persuasive technologies are a powerful medium for delivering interventions in the occupational setting is to assess their consistency with evidence-based practices. Persuasive technologies collect data using sensors (sense), analyse and interpret these data accordingly (reason) to give adequate feedback (action). To do this, such technologies draw on a variety of assumptions, theories and standards. However, it is often unclear which theories or models are being used. **Chapter 2** provided an overview of behaviour change theory incorporated in 45 mHealth applications (apps) for mental and physical health of employees. In particular, this study evaluated which behaviour change techniques (BCTs) could be identified and which combinations of BCTs were present. The results showed a limited presence of BCTs in general, limited use of potentially successful combinations of BCTs in apps and use of potentially unsuccessful combinations of BCTs. These findings indicate that the

potential of apps for health and wellbeing in the occupational setting might be substantially improved by incorporating behaviour change theory.

Research methods

The next step to gain insight into the potential of persuasive technologies in the occupational setting is to select research methods appropriate to evaluate the effectiveness of persuasive health technologies. This also involves understanding how and why a system is (or is not) used, which is critical for the adoption and use of persuasive technology. In **Chapter 3**, three different qualitative research methods were compared: interviews with employees, focus groups with employees and a focus group with experts. The objective of this study was: (1) to gain insight into the opinions and experiences of employees and experts on drivers and barriers using an mHealth app in the working context and (2) to assess the added value of three different qualitative methods to evaluate mHealth apps in a working context related to user satisfaction and technology acceptance. The results in Chapter 3 show that the type of evaluator (user or expert) and type of evaluation (interview or focus group) yielded different results. These findings indicate that qualitative evaluation of mHealth applications might benefit from combining more than one method. In addition, factors to consider when selecting a qualitative research method are the design, the development stage of technology, the working context in which it is being used, employees' mental models, practicability, resources and the skills required of experts and users to use a specific qualitative method.

Effectiveness

Insight in the effectiveness of persuasive technology on behaviour, health, wellbeing and performance was explored. For many of the persuasive technologies which appear on the market, the effectiveness remains unclear, which might hamper the adoption for application in the occupational setting. The purpose of persuasive technologies is to help users achieve their goals. In **Chapter 4**, whether a persuasive computer mouse was in fact able to do that was evaluated. The Hoverstop Mouse is a computer mouse that aims to change unnecessary, unfavourable postures of the lower arm and wrist which cause sustained muscle tension by giving tactile feedback signals to the user as signal to change those unfavourable postures. This mouse was used to study the effects of persuasive technology on behaviour, short-term health effects, performance and user friendliness. Fifteen subjects participated in a comparative, experimental study with repeated measures. Evidence was found that a tactile feedback signal could positively impact workers' behaviours. It decreased static muscle loading of the fore-arm muscles, thereby decreasing the risk of developing arm complaints. Neither positive nor negative effects on objectively measured productivity were found. However, subjective ratings did not support these findings: about half of the subjects judged the time necessary to complete a task to be greater compared to when using a traditional mouse. A large variety in user satisfaction was found, which indicated differences in personal preferences and needs.

In **Chapter 5**, a short-term randomized controlled trial (RCT) was performed to assess whether the results of the laboratory study in Chapter 4 would hold true in the field. This field study aimed to gain better insight into how users get accustomed to feedback, the effects on task performance and user satisfaction during the initial phase of working with the Hoverstop computer mouse. A total of 91 employees participated in the study. As in the laboratory study, positive effects were shown on behaviour in the field study as well. In addition, participants seemed to change behaviour over time: in two weeks, movement patterns changed from static to more dynamic. Again, the technology appeared not to affect performance. Usability ratings were mixed. The use of the mouse seems promising for preventing neck, shoulder and arm complaints.

In **Chapter 6**, four different types of feedback were evaluated with the aim of making feedback as unobtrusive as possible for application in a working environment and thereby exploring the possibilities to design systems that do change behaviour, but do not negatively impact task performance. In a laboratory experiment with 24 participants, the effects of two visual and two tactile feedback signals were compared to a no-feedback condition for a computer task. Results from the objective measures showed that all types of feedback were equally effective in reducing lifted hand/finger behaviour compared to absence of feedback, while task performance was not affected. In contrast to the objective measures, subjective user experience was significantly different for the four types of feedback signals. Continuous tactile feedback appeared to be the preferred signal; not only the effectiveness and efficiency were rated reasonable, it also scored best on perceived match between signal and required action. This study showed the importance of including user experiences when investigating persuasive technology. Although it seems difficult to optimize the balance between effectiveness (and thus noticeability) and unobtrusiveness, the results indicate that it might be possible to provide effective and efficient feedback with a non-obtrusive signal.

The findings in Chapter 4 to 6 show that a computer mouse with feedback, as an example of persuasive technology, may have benefits to motivate computer workers to adopt healthy, comfortable and productive work behaviours, at least in the short term. These studies add to the existing knowledge to determine the potential of persuasive technology as an effective contribution to the prevention of occupational health problems.

Societal Impact

Many stakeholders could benefit from persuasive technology for health and wellbeing at work. There might be shared benefits such as prevention of occupational risks (i.e. work-related stress) and enhancing health and wellbeing. In turn, this might lead to substantial gains in productivity, decreases in sick leave and decreases in healthcare costs. However, when applied in the working context, it is also important to explore the possible effects of persuasive technology on processes and relationships on the work floor. **Chapter 7** investigated the impact of persuasive technology in the working context, how it fit into current regulatory frameworks and where possible ethical problems might arise using the example of a digital stress coach. First, because many persuasive technologies lack validity and reliability, a risk of unjustified actions taken based upon the data may arise. Second, a digital stress coach might influence the relationship between employer and employee. An employee who discovered through persuasive technology that he or she is running a health risk, can hold his or her employer accountable. Employers can, using persuasive technology, track whether employees stick to agreements that they have made (e.g. to manage work-related stress). In addition, the emphasis placed by persuasive technology on self-management might cause a shift in perceived responsibility towards the individual employee, although health and wellbeing is a joint responsibility of the employer and employee according to the Dutch Working Conditions Act. Persuasive technology could be introduced by employers as part of their workplace policies. Although this might be beneficial for both employers and employees, it also risks interference with the employees' private life (e.g. monitoring sleep data). This might violate the employees' right to privacy or the right to protection of personal data. Third, with more available data, it is important to ensure that medical and other sensitive data are only added to the employee's file if doing so is compliant with relevant rules and practices. Finally, there is the risk for loss of autonomy of employees. Persuasive technology might make an employee feel subject to surveillance, even if the use is voluntarily, because private activities might become visible in the working context. It is therefore important to develop and implement technologies responsibly. Research on the effectiveness and validation of methods will help to draft guidelines or quality standards to direct this. At the moment, Dutch legislation on data protection and working conditions is now used as a framework for responsible incorporation of persuasive technology in the work context. Future research is necessary to look into the sufficiency of this framework.

Implications for research, design and occupational practice

Finally, in **Chapter 8,** a reflection of the main findings of all chapters in this thesis was provided. Alongside with methodological considerations, implications for future research, design and application in the occupational practice are given in the context of recent related research. The studies in this thesis explored the potential of persuasive technology for self-management in the context of health and wellbeing at work:

- Persuasive technology needs a solid theoretical and evidence-based foundation. Developing a comprehensive framework for translating theoretical determinants into persuasive technology design components is recommended. This framework should address theories related to occupational health and wellbeing as well as behaviour change. Such a framework might also facilitate interdisciplinary collaboration between persuasive technology designers and researchers, which would strengthen the evidence base in this area.
- Persuasive technology needs advanced research methods to assess impact. Persuasive technology challenges the way to conduct research, because the development of technology is characterized by a highly iterative development process and persuasive technologies are personalized and adapted to users' needs. Future research should aim to determine which quantitative and qualitative methods work best for which kind of persuasive technology, for which context and which methods work well together. Examples of research designs that show potential are N-of-1 studies, stepped wedge design or SMART. Applying these methods to assess the impact of persuasive technology for work health and wellbeing is recommended.
- More evidence on the effectiveness of persuasive technology is needed. Although this thesis provided insights into the short-term effectiveness of persuasive technology on behaviour, health-related outcomes, performance and usability and the results are promising, we have to be careful generalizing results to other persuasive technology applications. Research in this area is still in its early stages. Future research should focus on long-term studies to assess behaviour change and learning effects. Furthermore, it is important to look into possible unintended side effects of persuasive technology.
- To reach positive outcomes for health and wellbeing, it is important to address user satisfaction and acceptance of persuasive technology. First, testing system performance thoroughly before implementation is recommended, because it influences adoption and adherence. Second, the relevance and benefits of persuasive technology should be clear to the user and should address users' characteristics within the application itself, as well as in communications guiding implementation. Future research should focus on which characteristics of persuasive technology are important for users and how to personalize it accordingly. Third, persuasive technology should be designed to fit particular occupational contexts. To ensure a good fit, using participative design approaches

with end users and testing persuasive technology in real-life settings are recommended. Finally, employees should be alerted to their right to privacy and use of personal data, preferably in the early stages of development. Approaches that might be used are privacy by design or privacy impact assessment.

- Persuasive technology design and development needs multidisciplinary collaboration. To support these processes, making use of participatory design approaches involving all stakeholders, including end-users, is recommended.
- Persuasive technology requires standards and guidelines. Practical and ethical barriers exist that might hamper adoption of persuasive technology in the work context. Employers and employees might benefit from a regulatory framework to clarify liabilities and responsibilities in relation to new systems. Evaluating whether current legislation on working conditions, privacy and data protection provide sufficient framework to cover the risks of persuasive technology is advisable. The Dutch Personal Data Authority and the Inspectorate of the Ministry of Social Affairs and Employment might jointly provide clarification. In addition, developing assessment criteria or a seal of approval that provides end users with information about the purpose and the target groups of persuasive technology, what it is (not) capable of, whom it is (and is not) intended for and which purpose it serves is recommended. This seal of approval should also provide end users with information on the effectiveness of the technology, what evidence base is available, how it arrives at its recommendations and what methods are used. Currently, initiatives by organizations such as the ISO are being pursued to meet this need.

Conclusion

Persuasive technology is already affecting the occupational context and will become increasingly relevant in the future. However, given the current status of persuasive technology, the state of theory and evidence base, it may be concluded that application of persuasive technology within the occupational context is in its early stages and huge challenges remain. We need to address these challenges to keep up with the rapid developments of technology and the changes in work and workers to preserve, or preferably to enhance, the health and wellbeing of workers.

Samenvatting

Het doel van dit proefschrift is te verkennen of en hoe persuasive technology zinvol kan worden ingezet voor gezondheid en welzijn op het werk. Technologische ontwikkelingen hebben veel veranderingen in het werk teweeggebracht en deze ontwikkelingen zetten zich in een exponentieel tempo voort. Tot nu toe hebben technologische ontwikkelingen ons werk gemakkelijker, sneller en efficiënter gemaakt. Tegelijkertijd brengen technologische ontwikkelingen ook nieuwe risico's met zich mee; ondanks ontwikkelingen in de technologie en veranderingen in het werk hebben veel werknemers te kampen met werk gerelateerde gezondheidsproblemen. Mede door de vergrijzende beroepsbevolking is het belangrijk om eens opnieuw te kijken naar de huidige benaderingen voor het verbeteren van gezondheid en welzijn op het werk. Persuasive technology verwijst naar interactieve systemen die zijn ontwikkeld om de attitude of het gedrag, of beide, van gebruikers te veranderen door middel van overtuiging en sociale beïnvloeding. Het heeft reële potentie om verbeteringen in de werkomgeving te stimuleren, risicofactoren te verminderen of risico's beter te managen. Met name voor de werkcontext is er echter een gebrek aan inzicht in wanneer, waar en voor wie persuasive technology effectief kan zijn en hoe dit geëvalueerd zou kunnen worden.

In dit proefschrift wordt het potentieel van persuasive technology voor gezondheid en het welzijn op het werk verkend aan de hand van vier wetenschappelijke vraagstukken: (1) in hoeverre bevatten persuasive technologieën theoretische kennis, of op bewijs gebaseerde principes en constructen; (2) welke soorten onderzoeksmethoden zijn geschikt om persuasive technology voor gezondheid en welzijn op het werk te evalueren; (3) wat zijn de effecten van persuasive technology op gedrag, gezondheid en welzijn op het werk en (4) wat zijn de maatschappelijke gevolgen van het inzetten van persuasive technology voor de gezondheid en het welzijn op het werk?

S

Theorie en wetenschappelijke onderbouwing

De eerste stap om inzicht te krijgen of persuasive technologies een krachtig middel zijn als interventies in de werksetting, is het beoordelen van hun consistentie met wetenschappelijke onderbouwing. Persuasive technologies verzamelen gegevens met behulp van sensoren ('sense'), analyseren en interpreteren deze gegevens dienovereenkomstig ('reason') om adequate feedback te geven ('act'). Idealiter wordt hiervoor gebruik gemaakt van verschillende veronderstellingen, theorieën en normen. Het is echter vaak onduidelijk welke theorieën of modellen worden gebruikt. **Hoofdstuk 2** geeft een overzicht van de gedragsveranderingstheorie die is verwerkt in 45 mHealth-applicaties (apps) voor de mentale en fysieke gezondheid van medewerkers. In het bijzonder wordt in deze studie geëvalueerd welke gedragsveranderingstechnieken (Behavioral Change Techniques, oftewel BCT's) kunnen worden geïdentificeerd en welke combinaties van BCT's aanwezig zijn. De resultaten laten een beperkte aanwezigheid zien van BCT's in het algemeen, een beperkt gebruik van potentieel succesvolle combinaties van BCT's in apps en het gebruik van potentieel niet-succesvolle combinaties van BCT's. Deze bevindingen geven aan dat het potentieel van apps voor gezondheid en welzijn op het werk aanzienlijk kan worden verbeterd door gedragsveranderingstheorieën te integreren.

Onderzoeksmethoden

De volgende stap om inzicht te krijgen in het potentieel van persuasive technologies in de werksetting, is het selecteren van onderzoeksmethoden die geschikt zijn om de effectiviteit van persuasive technologies te evalueren. Dit omvat ook het begrijpen van de wijze waarop en waarom een systeem wordt gebruikt (of niet), wat van cruciaal belang is voor de adoptie en het gebruik van persuasive technology. In **hoofdstuk 3** worden drie verschillende kwalitatieve onderzoeksmethoden vergeleken: interviews met werknemers, focusgroepen met werknemers en een focusgroep met experts. Het doel van dit onderzoek is: (1) inzicht te krijgen in de meningen en ervaringen van werknemers en experts over drijfveren en barrières in het gebruik van een mHealth-app voor de werkcontext en (2) de toegevoegde waarde te beoordelen van drie verschillende kwalitatieve methoden om mHealth-apps voor een werkcontext te evalueren op gebruikerstevredenheid en acceptatie van technologie. De resultaten in hoofdstuk 3 laten zien dat het type beoordelaar (werknemer of expert) en het type evaluatie (interview of focusgroep) verschillende resultaten hebben opgeleverd. Deze bevindingen geven aan dat een kwalitatieve evaluatie van mHealth-applicaties baat kan hebben bij het combineren van meer dan één methode. Factoren waarmee rekening moet worden gehouden bij het selecteren van een kwalitatieve onderzoeksmethode zijn het ontwerp van de technologie, de ontwikkelingsfase van de technologie, de werkcontext waarin deze wordt toegepast, de mentale modellen van de werknemers, de praktische uitvoerbaarheid, de beschikbare middelen en de vaardigheden van experts en werknemers die nodig zijn om een specifieke kwalitatieve methode te gebruiken.

Effectiviteit

Verder is inzicht in de effectiviteit van persuasive technology in relatie tot gedrag, gezondheid, welzijn en taakprestatie onderzocht. Van de vele persuasive technologies die op de markt verschijnen is de effectiviteit onduidelijk, wat toepassing in de werksetting zou kunnen belemmeren. Het doel van persuasive technologies is de gebruikers te helpen hun doelen te bereiken. In **hoofdstuk 4** is geëvalueerd of een persuasive computermuis daartoe in staat was. De Hoverstop muis is een computermuis die tot doel heeft onnodige, ongunstige houdingen van onderarm en pols die aanhoudende spierspanning veroorzaken te veranderen, door tactiele feedbacksignalen aan de gebruiker te geven als teken om ongunstige houdingen te veranderen. Deze muis is gebruikt als een middel om de effecten van persuasive technology op gedrag, korte termijn gezondheidseffecten, taakprestatie en gebruiksvriendelijkheid te evalueren. Vijftien proefpersonen hebben deelgenomen aan een vergelijkend, experimenteel onderzoek met herhaalde metingen. De resultaten laten zien dat het mogelijk is met een tactiel feedbacksignaal het gedrag van werknemers positief te beïnvloeden. Het vermindert de statische spierbelasting van de onderarmspieren, waardoor het risico op het ontwikkelen van armklachten afneemt. Er zijn noch positieve, noch negatieve effecten gevonden op de objectief gemeten productiviteit. De subjectieve beoordelingen laten echter een ander beeld zien: ongeveer de helft van de proefpersonen schat in dat de tijd die nodig is om een taak uit te voeren langer is dan bij gebruik van een traditionele muis. Verder is een grote variatie in gebruikerstevredenheid gevonden, wat laat zien dat er verschillen bestaan in persoonlijke voorkeuren en behoeften.

In **hoofdstuk 5** is een korte termijn Randomized Controlled Trial (RCT) uitgevoerd om te evalueren of in het veld dezelfde resultaten worden gevonden als in het laboratoriumonderzoek van hoofdstuk 4. Dit veldonderzoek heeft tot doel beter inzicht te krijgen in hoe gebruikers wennen aan feedback, de effecten op de taakuitvoering en de gebruikerstevredenheid in de beginfase van het werken met de Hoverstop computermuis. In totaal namen 91 medewerkers deel aan het onderzoek. Ook in het veldonderzoek zijn positieve effecten op het gedrag aangetoond. Daarnaast lijken de deelnemers hun gedrag in de loop van de tijd te veranderen: in twee weken tijd veranderde het bewegingspatroon van statisch naar meer dynamisch. Ook hier lijkt de technologie geen invloed te hebben op de taakprestatie. De gebruikerstevredenheid liet gemengde resultaten zien. Het gebruik van de muis lijkt veelbelovend voor het voorkomen van nek-, schouder- en armklachten.

Op zoek naar het zo onopvallend mogelijk maken van feedback voor toepassing in een werkomgeving, en daarbij de mogelijkheden onderzoekend om systemen te ontwerpen die wel gedrag veranderen, maar geen negatieve invloed hebben op de taakuitvoering, worden in **hoofdstuk 6** vier verschillende soorten feedback geëvalueerd. In een laboratoriumexperiment met 24 deelnemers zijn de effecten van twee visuele en twee tactiele feedbacksignalen vergeleken met een conditie zonder feedback bij het uitvoeren van een computertaak. De resultaten van de objectieve metingen laten zien dat alle soorten feedback even effectief zijn om het onnodig heffen van handen en vingers te verminderen in vergelijking met de afwezigheid van feedback, terwijl de taakuitvoering niet wordt beïnvloed. In tegenstelling tot de objectieve metingen was de subjectieve gebruikerservaring voor de vier soorten feedbacksignalen significant anders. Een continu tactiel feedback signaal bleek het voorkeurssignaal te zijn; niet alleen werden de effectiviteit en efficiëntie als redelijk beoordeeld, het werd ook het best beoordeeld op match tussen feedbacksignaal en de vereiste actie. Dit onderzoek liet het belang zien van het evalueren van gebruikerservaringen bij het onderzoeken van persuasive technology. Hoewel het moeilijk lijkt om de balans tussen effectiviteit (en dus waarneembaarheid) en onopvallendheid te vinden, geven de resultaten aan dat het mogelijk is om effectieve en efficiënte feedback te geven met een niet-opvallend signaal.

De bevindingen in hoofdstuk 4 tot en met 6 laten zien dat een computermuis met feedback, als voorbeeld van persuasive technology, kan bijdragen aan het motiveren van computerwerkers tot gezond, comfortabel en productief werkgedrag, in ieder geval op de korte termijn. Deze studies dragen daarmee bij aan bestaande kennis over de rol van persuasive technology bij preventie van arbeidsrisico's.

Maatschappelijke impact

Verschillende belanghebbenden zouden kunnen profiteren van persuasive technology voor gezondheid en welzijn op het werk. Er kunnen gemeenschappelijke voordelen zijn, zoals preventie van arbeidsrisico's (bijvoorbeeld werk gerelateerde stress) en verbetering van gezondheid en welzijn. Dit zou op vervolgens kunnen leiden tot productiviteitswinst, een daling van het ziekteverzuim en een daling van kosten van gezondheidszorg. Bij toepassing in de werkcontext is het echter ook belangrijk om de mogelijke effecten te onderzoeken van persuasive technology op processen en relaties op de werkvloer. In **hoofdstuk 7** is aan de hand van een voorbeeld van een digitale stresscoach onderzocht wat de impact kan zijn van persuasive technology in de werkcontext, hoe deze past binnen de huidige regelgeving en waar mogelijk ethische problemen kunnen ontstaan. Ten eerste kan er een risico ontstaan doordat onterechte acties worden ondernomen op basis van de data, omdat het veel persuasive technologies aan validiteit en betrouwbaarheid ontbreekt. Ten tweede kan een digitale stresscoach de relatie tussen werkgever en werknemer beïnvloeden. Een werknemer die door middel van persuasive technology ontdekt dat hij een gezondheidsrisico loopt, kan zijn werkgever hiervoor aansprakelijk stellen. Werkgevers kunnen met behulp van persuasive technology bijhouden of werknemers zich houden aan gemaakte afspraken (om bijvoorbeeld werk gerelateerde stress te managen). Daarnaast kan de nadruk die persuasive technology legt op zelfmanagement leiden tot een verschuiving in de ervaren verantwoordelijkheid in de richting van de individuele werknemer, hoewel gezondheid en welzijn volgens de Arbowet een gezamenlijke verantwoordelijkheid is van werkgever en werknemer. Persuasive technology zou door werkgevers kunnen worden ingevoerd als onderdeel van hun beleid. Hoewel dit zowel voor werkgevers als voor werknemers gunstig kan zijn, kan dit ook het privéleven van de werknemers binnendringen (bijvoorbeeld bij het monitoren van de slaapgegevens). Dit kan een inbreuk zijn op de privacy van werknemers of de bescherming van persoonsgegevens. Ten derde is het bij de grotere beschikbaarheid aan data belangrijk om ervoor te zorgen dat medische en andere gevoelige data alleen aan een personeelsdossier worden toegevoegd als dit in overeenstemming is met de relevante wet- en regelgeving. Ten slotte bestaat het risico dat werknemers hun autonomie verliezen. Door persuasive technology kan een werknemer zich gecontroleerd voelen door monitoring, zelfs als het gebruik vrijwillig is, omdat privéactiviteiten zichtbaar kunnen worden in de werkcontext. Het is daarom belangrijk om technologieën op verantwoorde wijze te ontwikkelen en te implementeren. Validering van de gebruikte methoden en onderzoek naar de effectiviteit kan helpen bij het opstellen van richtlijnen of kwaliteitsnormen om dit te sturen. Op dit moment wordt de Nederlandse wetgeving voor privacy, voor bescherming persoonsgegevens en voor arbeidsomstandigheden gebruikt als kader voor een verantwoord gebruik van persuasive technology in de werkcontext. Verder onderzoek is nodig om na te gaan of dit kader toereikend is.

Implicaties voor onderzoek, design en praktijk

Tot slot wordt in **hoofdstuk 8** een reflectie gegeven op de belangrijkste bevindingen van alle hoofdstukken in dit proefschrift, gebruik makend van recent, aanverwant onderzoek. Naast methodologische overwegingen worden implicaties voor toekomstig onderzoek, design en toepassing in de praktijk gegeven. De resultaten uit dit proefschrift leiden tot de volgende aanbevelingen:

› Persuasive technology heeft een solide theoretische en wetenschappelijke basis nodig. Het wordt aanbevolen om een raamwerk te ontwikkelen waarmee theoretische determinanten kunnen worden vertaald naar ontwerprichtlijnen voor persuasive technology. Dit raamwerk zou zich moeten richten op theorieën over bevordering van gezondheid en welzijn op het werk en theorieën over gedragsverandering. Een dergelijk raamwerk zou ook de interdisciplinaire samenwerking tussen ontwerpers en onderzoekers van persuasive technology kunnen vergemakkelijken, wat de wetenschappelijke basis op dit gebied zal versterken.

- Voor het evalueren van de impact van persuasive technology zijn geavanceerde onderzoeksmethoden nodig. De manier van onderzoek doen met persuasive technology is een uitdaging, omdat de ontwikkeling van technologie wordt gekenmerkt door een iteratief proces en omdat persuasive technologies worden gepersonaliseerd en aangepast aan de behoeften van de gebruikers. Toekomstig onderzoek zou zich moeten richten op het bepalen van welke kwantitatieve en kwalitatieve onderzoeksmethoden het meest geschikt zijn voor welke persuasive technologies, voor welke context en welke methoden goed samengaan. Voorbeelden van interessante onderzoeks-designs zijn N-of-1 studies, stepped wedge design of SMART. Het is aan te bevelen om deze methoden toe te passen om de impact van persuasive technology op de gezondheid en het welzijn op het werk te beoordelen.

- Er is meer bewijs nodig voor de effectiviteit van persuasive technology. Hoewel dit proefschrift inzicht geeft in de effectiviteit van persuasive technology op gedrag, gezondheids-gerelateerde uitkomsten, prestaties en bruikbaarheid op korte termijn en de resultaten veelbelovend zijn, moeten we voorzichtig zijn met het generaliseren van de resultaten naar andere persuasive technology toepassingen. Het onderzoek op dit gebied bevindt zich nog in het beginstadium. Toekomstig onderzoek moet gericht zijn op lange termijn studies om gedragsverandering en leereffecten te evalueren. Verder is het belangrijk om te kijken naar mogelijke onbedoelde neveneffecten van persuasive technology.

- Om positieve resultaten voor gezondheid en welzijn te kunnen bereiken, is het belangrijk om aandacht te besteden aan de gebruikerstevredenheid en de gebruikersacceptatie van persuasive technology. Ten eerste wordt aanbevolen om de prestaties van het systeem grondig te testen voordat het wordt geïmplementeerd, aangezien dit van invloed is op de adoptie en het gebruik. Ten tweede moeten de relevantie en de voordelen van persuasive technology voor de gebruiker duidelijk zijn. Ook dient rekening worden gehouden met de kenmerken van de gebruikers, zowel in de applicatie zelf als in de communicatie bij implementatie. Toekomstig onderzoek dient zich te richten op de vraag welke kenmerken van persuasive technology belangrijk zijn voor de gebruikers en hoe deze kunnen worden gepersonaliseerd. Ten derde moet de persuasive technology zo worden ontworpen dat zij op bepaalde arbeidsomstandigheden is afgestemd. Om een goede match te garanderen tussen persuasive technology, gebruiker en gebruiksomgeving, wordt aanbevolen om participatieve ontwerpbenaderingen met eindgebruikers toe te passen en om persuasive technology in een reële omgeving te testen. Ten slotte moeten werknemers worden gewezen op hun recht op privacy en rechten in het gebruik van persoonsgegevens, bij voorkeur in een vroeg stadium van de ontwikkeling. Aanpakken die hierbij kunnen worden ingezet zijn Privacy by Design of Privacy Impact Assessment.

- Het ontwerp en de ontwikkeling van persuasive technology vereist multidisciplinaire samenwerking. Ter ondersteuning van deze samenwerking wordt aanbevolen om gebruik te maken van participatieve ontwerpbenaderingen waarbij alle belanghebbenden, inclusief de eindgebruikers, worden betrokken.
- Persuasive technology heeft behoefte aan normen en richtlijnen. Er bestaan praktische en ethische drempels die de toepassing van persuasive technology in een werkcontext kunnen belemmeren. Werkgevers en werknemers kunnen baat hebben bij een regelgevend kader om de verplichtingen en verantwoordelijkheden met betrekking tot dit soort nieuwe systemen te verduidelijken. Er wordt geadviseerd om te evalueren of de huidige wetgeving inzake arbeidsomstandigheden, privacy en gegevensbescherming voldoende kader biedt om de risico's van persuasive technology te dekken. Het College Bescherming Persoonsgegevens en de Inspectie van het Ministerie van Sociale Zaken en Werkgelegenheid zouden gezamenlijk duidelijkheid kunnen verschaffen. Daarnaast wordt geadviseerd om beoordelingscriteria of een keurmerk te ontwikkelen die de eindgebruikers informatie geven over het doel en de doelgroepen van persuasive technology, wat het (niet) kan, voor wie het (wel en niet) bedoeld is en welk doel het dient. Daarnaast moet dit keurmerk de eindgebruikers informatie verschaffen over de effectiviteit van de technologie, welke wetenschappelijke basis er beschikbaar is, hoe het tot de aanbevelingen komt en welke methoden er worden gebruikt. Momenteel worden initiatieven door onder andere ISO genomen om in deze behoefte te voorzien.

S

Conclusie

Persuasive technology heeft nu al invloed op de werkcontext en dat zal in de toekomst nog relevanter worden. Gezien de huidige status van persuasive technology, de stand van de wetenschap en de wetenschappelijke onderbouwing van de technologie, kan echter worden geconcludeerd dat de toepassing van persuasive technology binnen de werkcontext nog in de kinderschoenen staat en dat er nog enorme uitdagingen zijn. We zullen deze uitdagingen aan moeten gaan om gelijke tred te kunnen houden met de snelle ontwikkelingen van technologie en veranderingen in het werk en de werknemers om de gezondheid en het welzijn van de werknemers te behouden, of bij voorkeur te verbeteren.

About the author

Elsbeth de Korte was born on June 1st 1973 in Wageningen, the Netherlands. After completing secondary school at the Ignatius College in Purmerend, she started to study Occupational Therapy at Amsterdam University of Applied Sciences in 1993. She graduated in 1997 with a research project on "The effect of arm- and wrist supports on the load of the upper extremity during VDU work" and was awarded the 'Jan van Eyck Award' for excellent quality. In 1997, Elsbeth started to study at the Faculty of Human Movement Sciences at the VU University Amsterdam and graduated in 2000 with a specialisation in Ergonomics. Between July 1998 and September 1999, Elsbeth worked part-time as an Occupational Therapist at Waterlandziekenhuis (hospital) in Purmerend. Since September 2000 she has been employed at TNO. Elsbeth started her PhD at Delft University of Technology Faculty Industrial Design Engineering in 2007. During her PhD, she remained working at TNO; various research projects that she conducted at TNO form the basis for this thesis. In her current position as senior research scientist in the Team Work, Health & Technology, Elsbeth continued to do research in the area of health and wellbeing at work. In this context, she published articles and book chapters on workplace innovation and evaluation, physical and psychosocial workload and human-technology interaction. Elsbeth participated as an expert in international working groups for the European Commission, PEROSH and EU-OSHA. Besides her expertise, she uses her planning and organisational skills to set up and manage research programs with national and international public and private partners. She organized sessions and symposia for international conferences such as HCI International Congress, IEA Congress, Wellbeing at Work Conference and AHFE Conference. She also has been editor for the Dutch Journal of Human Factors. Currently she is editorial board member for ICT & Health.

Elsbeth is living together with Floris and their daughters Wende and Marin in Wijdewormer.

Contact:

elsbeth.dekorte@tno.nl

List of publications

International scientific articles

Van de Ven, H., Hulsegge, G., Zoomer, T., De Korte, E.M., Burdorf, A., Oude Hengel, K.M. (Submitted) The acute effects of shift schedule characteristics on fatigue and sleep quality using daily measurements of 6.195 observations among 223 workers. *Scandinavian Journal of Work, Environment & Health.*

Van de Ven, H., Van Drongelen, A., De Korte, E., Van der Klink, J., Bültmann, U, Brouwer, S. (Submitted). Shiftworkers who nap in between shifts are relatively unhealthy. *Chronobiology International.*

Niks, I.M.W., Van Drongelen, A., De Korte, E.M. (2020). Promoting employees' recovery during shift work: protocol for a workplace intervention study. *JMIR Research Protocols*, 9(7): e17368. Doi: 10.2196/17368

Kraaij, W., Verberne, S., Koldijk, S., De Korte, E., Van Dantzig, S., Sappelli, M., Shoaib, M., Bosems, S., Achterkamp, R., Bonomi, A., Schavemaker, J., Hulsebosch, B., Wabeke, T., Vollenbroek-Hutten, M., Neerincx, M., Van Sinderen, M. (2019). Personalized support for well-being at work: an overview of the SWELL project. *User Model User-Adap Inter*. Doi: 10.1007/s11257-019-09238-3

De Korte, E., Wiezer, N., Bakhuys Roozeboom, M., Vink, P., Kraaij, W. (2018). Behavior change techniques in mHealth apps for the mental and physical health of employees: systematic assessment. *JMIR Mhealth Uhealth*, 6(10): e167. Doi: 10.2196/mhealth.6363.

De Korte, E., Wiezer, N., Janssen, J., Vink, P., Kraaij, W. (2018). Evaluating an mHealth app for health and wellbeing at work: mixed-method qualitative study. *JMIR Mhealth Uhealth,* 6(3): e72. Doi: 10.2196/mhealth.6335.

De Korte, E.M., Spiekman, M., Hoes-Van Oeffelen, L., Van der Zande, B., Vissenberg, G., Huiskes, G., Kuijt-Evers, L.F.M. (2015). Personal environmental control: effects of pre-et conditions for heating and lighting on personal settings, task performance and comfort experience. Building and Environment 86: 166-176. Doi: https://doi.org/10.1016/j.buildenv.2015.01.002

De Korte, E.M., Huysmans, M.A., De Jong, A.M. Van de Ven, J.G.M., Ruijsendaal, M. (2012). Effects of four types of non-obtrusive feedback on computer behavior, task performance and comfort. *Applied Ergonomics*, 43(2): 344-353. https://doi.org/10.1016/j.apergo.2011.06.010.

Vink, P., Blok, M., Formanoy, M., De Korte, E., Groenesteijn, L. (2012). The effects of new ways of work in The Netherlands: national data and a case study. Work 41 (Supplement 1): 5081-5058.

Blok, M., Groenesteijn, L., Formanoy, M., De Korte, E., Vink, P. (2010). Innovative ways of working: A comparison of a traditional and a flexible office environment on office space use, performance, privacy and satisfaction. Advances in Occupational, Social and Organizational Ergonomics: 2018-2227. Doi: 10.1201/EBK1439835074-25

De Korte, E.M., De Kraker, H., Bongers, P.M., Van Lingen, P. (2008). Effects of a feedback signal in a computer mouse on movement behavior, muscle load, productivity, comfort and user friendliness. *Ergonomics,* 51 (11): 1757-1775.

De Kraker, H., De Korte, E.M., Van Mil, F.L.K., Rijs, B.P., Bongers, P.M. (2008). The effect of a feedback signal in a computer mouse on hovering behavior, productivity, comfort and usability in a field study. *Ergonomics*, 51(2): 140-155.

Van den Heuvel, S.G., IJmker, S., Blatter, B.M., De Korte, E.M. (2007). Loss of Productivity Due to Neck/Shoulder Symptoms and Hand/Arm Symptoms: Results from the PROMO-Study. *J Occup Rehabil* 17, 370–382. https://doi.org/10.1007/s10926-007-9095-y.

De Korte, E.M., Van Lingen, P. (2006). The effect of speech recognition on working postures, productivity and the perception of user friendliness. *Applied Ergonomics,* 37(3): 341-347. https://doi.org/10.1016/j.apergo.2005.06.011.

Hoozemans, M.J.M., De Looze, M.P., Kingma, I., Reijneveld, K.C.N., De Korte, E.M., Van der Grinten, M.P., Van Dieën, J.H. (2005). Workload of window cleaners using ladders differing in rung separation. *Applied Ergonomics,* 36(3): 275-282. https://doi.org/10.1016/j.apergo.2005.01.013

Visser, B., De Looze, M.P., Veeger, D.H.E.J., Douwes, M., Groenesteijn, L., De Korte, E., Van Dieën, J.H. (2003). The effects of precision demands during a low intensity pinching task on muscle activation and load sharing of the fingers. *Journal of Electromyography and Kinesiology*, 13(2): 149-157.

Visser, B., De Korte, E., Van der Kraan, I., Kuijer, P. (2000). The effect of arm and wrist supports on the load of the upper extremity during VDU work. *Clinical Biomechanics,* 15 (Supplement 1): S34-S38. https://doi.org/10.1016/S0268-0033(00)00058-9.

Dutch scientific articles

Van de Ven, H., Hulsegge, G., De Korte, E., Van Drongelen, A. (Submitted). De Persoonlijke Klok. *Tijdschrift voor Human Resource Management.*

Koopmans, L., De Korte, E. (2015). Beter presteren op je werk door Het Nieuwe Werken? *Tijdschrift voor Human Factors*. 4(40): 15-18.

De Korte, E. (2013). Geavanceerde technologie en gezond werken. Tijdschrift voor Ergonomie 3(38): 23.

De Korte, E.M., Kuijt-Evers, L.F.M., Vink, P. (2006). Afleiding beïnvloedt productiviteit bij een eenvoudige taak niet: het effect van visuele en auditieve afleiding op de prestaties bij een eenvoudige taak. Tijdschrift voor Ergonomie,, 5(31): 4-10.

Jans, M.P., De Korte, E.M., Heinrich, J., Hildebrandt, V.H. (2006). Intermitterende vervolgbehandeling Oefentherapie Cesar bij subacute of chronische aspecifieke rugklachten: een RCT. *Nederlands Tijdschrift voor Fysiotherapie*, 116(5): 111.

Kraker, H., Hendriksen, I.J.M., De Korte, E.M., Van der Maas, E.M.M. (2005). The effect of a campaign to stimulate walking during lunch break on the physical activity behavior of employees. In Dutch: Het effect van een lunchwandelcampagne op het beweeggedrag van werknemers. *Geneeskunde en Sport,* 6(38): 172-178.

Korte, E.M., Van Lingen, P., Thé, K.H. (2002). Schrijven met je stem: een pilotstudie naar de rol van spraakherkenning bij het verminderen van RSI-risico's. *Tijdschrift voor Ergonomie,* 6(27): 12-19.

Reijneveld, C.N., De Looze, M.P., Hoozemans, M.J.M., Van der Grinten, M.P., De Korte, E.M., Kingma, I. (2002). De optimale sportafstand voor de ladder van de glazenwasser. *Tijdschrift voor Ergonomie,* 2(27): 10-16.

De Korte, E., Van der Kraan, I., Kuijer, P.P.F.M., Visser, B. (1997). Ondersteuningen steunen? Een onderzoek naar het effect van ondersteuningen op de belasting van de bovenste extremiteit bij beeldschermwerk. *Tijdschrift voor Ergonomie,* 6: 173-179.

Book chapters

Van Lieshout, M., Wiezer, N., De Korte, E. (2015). The Digital Stress coach. Total control over your mental health, or 'Big Brother is watching you'? Chapter 6. In: Kool, L., Timmer, J., Van Est, R. (Eds.), *Sincere Support - The rise of the e-coach* (pp. 155-178) The Hague: Rathenau Instituut. ISBN/EAN: 978-90-77364-64-2

Van Lieshout, M., Wiezer, N., De Korte, E. (2014). De digitale stress coach: Totale controle over je mentale gezondheid of 'Big Brother is watching you'? Hoofdstuk 6. In: Kool, L., Timmer, J., Van Est, R. (Red.), *Eerlijk advies – De opkomst van de e-coach* (pp. 167-192) Den Haag: Rathenau Instituut. ISBN/EAN: 978-90-77364-59-8.

Pot, F., Dhondt, S., De Korte, E.M., Oeij, P., Vaas, F. (2012). *Workplace innovation in the Netherlands*. Chapter 8. In: I. Houtman (Ed.), Work Life in The Netherlands (pp 173-190). TNO, Hoofddorp, The Netherlands.

Huysmans, M.A., De Korte, E.M., Van Lingen, P. (2013). Invoermiddelen naast het toetsenbord. In: Peereboom, K.J., Van Scheijndel, P.A.M., (Red.) Handboek Ergonomie/Human Factors (pp 243-268). Alphen aan den Rijn: Vakmedianet.

Vink, P., De Korte, E.M. (2009). *Comfort*. In: Vink, P. (Red.), Aangetoond effecten van het kantoorinterieur: naar comfortabele, innovatieve, productieve en duurzame kantoren (pp 1-11). Alphen aan den Rijn: Kluwer.

De Korte, E.M., Vink, P., Zenk, R. (2009). *Intelligente producten en omgevingen*. In: Vink, P. (Red.), Aangetoond effecten van het kantoorinterieur: naar comfortabele, innovatieve, productieve en duurzame kantoren (pp 1-11). Alphen aan den Rijn: Kluwer.

Huysmans, M., De Korte, E.M., Van Lingen, P. (2005). Invoermiddelen naast het toetsenbord voor computerwerk. In: Voskamp, P., Van Scheijndel, P.A.M., Peereboom, K.J. (Red.), Handboek Ergonomie 2005 (pp 165-178). Alphen aan den Rijn: Kluwer.

Popular articles

Bakhuys Roozeboom, M., Huijs, J., Van der Beek, D., Goede, H., Bouwens, L., Van der Meer, L., De Korte, E., Van Empelen, P. (2020). Gedrag is de sleutel bij het verlagen van gezondheids- en veiligheidsrisico's op het werk. Whitepaper. *TNO.* (http://publications.tno.nl/publication/34637499/swOrkb/TNO-2020-gedrag.pdf)

De Korte, E., Blok, M., Kuijt-Evers, L., Iske, P., Vink, P. (2012). Innovatiemanagement en FM. Creativiteit bevorderen met future centers*. Facility Management Magazine* 198(25): 30-33.

Commissaris, D., Hitters, K., De Korte, E., Mol, P.J. (2012). Beweegmomentjes maken zittend werk dynamischer! Fit op kantoor deel 2. *OfficeMagazine*, 5(8): 16-18.

Commissaris, D., Hitters, K., De Korte, E., Mol, P.J. (2012). Beweegmomentjes maken zittend werk dynamischer. *Arbo Magazine*, 6: 18-19.

De Korte, E., Kuijt, L., Van der Kleij, R. (2011). Creatief gedrag gekleurd. Hoe ontmoetingsruimten creativiteit kunnen beïnvloeden. *Kleur en Visie*, December.

Groenesteijn, L., Formanoy, M., Blok, M., De Korte, E. (2010). Inrichten van een creatieve werkomgeving. *Facility Management Informatie*, juni: 28-33.

De Kraker, H., Formanoy, M., De Korte, E.M. (2009). Gezond computeren bij elfjarige kinderen. *Beweegreden*: 12-17

Van den Heuvel, S., IJmker, S., Blatter, B., De Korte, E. (2008). Op halve kracht: RSI veroorzaakt productiviteitsverlies. *Personeelsbeleid*, 5(44): 16-17.

Conference proceedings

De Korte, E.M. (2019). *Comfort. What's behaviour got to do with it?* (Keynote). International Comfort Congress 2019. Delft University of Technology, Delft, 29 August 2019.

Niks, I., Van Drongelen, A., Kranenborg, K., De Korte, E., Wiezer, N. (2019). *Design of a workplace intervention for promoting employees' recovery during shiftwork.* Fifth International Conference on Wellbeing at Work 2019, Paris, France, 22-23 May 2019.

De Korte, E.M. (2015). *Evaluating an mHealth application for workers' health and wellbeing: differences between three qualitative methods*. Seminar 1st PEROSH Research Exchange Warsaw, 21-22 Oct 2015. CIOP.

De Korte, E.M., Kuijt-Evers, L., Spiekman, M., Hoes-Van Oeffelen, L., Van der Zande, B., Vissenberg, G., Huiskes, G. (2013). *Evaluating comfort levels of a workstation with an individually controlled heating and lighting system*. In: Duffy V.G. (Ed.) Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management. Human Body Modeling and Ergonomics. DHM 2013. Lecture Notes in Computer Science, vol 8026, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39182-8_25.

De Korte, E., Kuijt, L., Van der Kleij, R. (2011). *Effects of meeting interior design on team performance in a creativity task*. In: Robertson, M.M. (ed.). International Conference on Ergonomics and Health Aspects of Work with Computers. EHAWC 2011. Lecture Notes in Computer Science, vol 6779, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-21716-6_7.

Groenesteijn L., Blok M., Formanoy M., de Korte E., Vink P. (2009) *Usage of Office Chair Adjustments and Controls by Workers Having Shared and Owned Work Spaces*. In: Karsh BT. (eds) Ergonomics and Health Aspects of Work with Computers. EHAWC 2009. Lecture Notes in Computer Science, vol 5624. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02731-4_3

Blok, M.M., De Korte, E.M., Groenesteijn, L., Formanoy, M., Vink, P. (2009). *The effects of a task facilitating working environment on office space use, communication, concentration, collaboration, privacy and distraction*. Proceedngs of the 17th World Congress on Ergonomics (IEA 2009), 9-14 August 2009, Beijing, China.

De Kraker, H., Formanoy, M., De Korte, E. (2008). *The effect of a healthy computing educational program on knowledge, attitude and computer behavior in middle school children*. In: Halgadottir, B. (Ed.), NES 2008 Ergonomics is a lifestyle, Reykjavik, Iceland 11-13 August, 225.

De Korte E., Kuijt-Evers L., Vink P. (2007) *Effects of the Office Environment on Health and Productivity: Auditory and Visual Distraction.* In: Dainoff M.J. (eds) Ergonomics and Health Aspects of Work with Computers. EHAWC 2007, held as part of HCI International 2007, Beijing, China. Lecture Notes in Computer Science, vol 4566. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73333-1_4.

Vink, P., De Korte, E., Blok, M., Groenesteijn, L. (2007). Effects of the Office Environment on Health and Productivity: Effects of the Coffee Corner Position. In: Dainoff M.J. (eds) Ergonomics and Health Aspects of Work with Computers. EHAWC 2007, held as part of HCI International 2007, Beijing, China. Lecture Notes in Computer Science, vol 4566. Springer, Berlin, Heidelberg. Doi: 10.1007/978-3-540-73333-1_20

De Korte, E.M., De Kraker, H., Bongers, P.M., Van Lingen, P. (2006). *Changing workstyle: healthy behaviour through feedback signals during computer work.* In: Pikaar, R.N., Koningsveld, E.A.P., Settels, P.J.M. (Eds.). IEA 2006, 16th World Congress on Ergonomics 'Meeting Diversity in Ergonomics', Maastricht, The Netherlands. Oxford, Elsevier.

Commissaris, D.A.C.M., Douwes, M., Schoenmaker, N., De Korte, E.M. (2006). *Recommendations for sufficient physical activity at work: promoting physical activity in low intensity static jobs*. In: Pikaar, R.N., Koningsveld, E.A.P., Settels, P.J.M. (Eds.). IEA 2006, 16th World Congress on Ergonomics 'Meeting Diversity in Ergonomics', Maastricht, The Netherlands. Oxford, Elsevier.

P

Acknowledgements