

# DESIGN AND LAB EXPERIMENT OF A STRESS DETECTION SERVICE BASED ON MOUSE MOVEMENTS

*Research full-length paper*

*Track 12*

Kowatsch, Tobias, University of St.Gallen, St.Gallen, Switzerland, tobias.kowatsch@unisg.ch

Wahle, Fabian, ETH Zurich, Zurich, Switzerland, fwahle@ethz.ch

Filler, Andreas, University of St.Gallen, St.Gallen, Switzerland, andreas.filler@unisg.ch

## Abstract

*Workplace stress can negatively affect the health condition of employees and with it, the performance of organizations. Although there exist approaches to measure work-related stress, two major limitations are the low resolution of stress data and its obtrusive measurement. The current work applies design science research with the goal to design, implement and evaluate a Stress Detection Service (SDS) that senses the degree of work-related stress solely based on mouse movements of knowledge workers. Using van Gemmert and van Galen's stress theory and Bakker and Demerouti's Job Demands-Resource model as justificatory knowledge, we implemented a first SDS prototype that senses mouse movements and perceived stress levels. Experimental results indicate that two feature sets of mouse movements, i.e. average deviation from an optimal mouse trajectory and average mouse speed, can classify high versus low stress with an overall accuracy of 78%. Future work regarding a second build-and-evaluate loop of a SDS, then tailored to the field setting, is discussed.*

*Keywords: Health Information System, Human-Computer Interaction, Workplace Stress, Task Performance.*

## 1 Introduction

Work has changed over the past decades, resulting in steadily increasing workloads (Eurofound, 2012). Moreover, working environments are one of the dominant predictors of mental health disorders (WHO, 2013). By 2020, 50% of the top ten medical problems worldwide will be stress-related with work constituting a primary source of stress in modern society (Cartwright and Cooper, 2014). For example, a recent study by Health Promotion Switzerland indicates that 25.4% of employees perceive work-related stress and exhaustion, potentially leading to a loss of productivity to the amount of approximately 5.6 billion US Dollars (Igic et al., 2016).

Stress, defined as the “psychological and physical state that results when the resources of the individual are not sufficient to cope with the demands and pressures of the situation” (Michie, 2002, p. 67), negatively affects individual employees and with them, the performance of their organizations. In particular, it threatens the health condition, quality of life, work-related goal achievements, self-esteem, confidence and personal development (Michie, 2002; Nyberg et al., 2014). If job demands cannot be balanced by individual resources in the long run, diastolic blood pressure, heart rate, blood glucose, cholesterol concentration, escapist drinking, and smoking among other symptoms of stress show first evidence of serious health conditions such as cardiovascular diseases, Type 2 diabetes, or mental disorders (Boedeker and Klindworth, 2007; Geurts, 2014; Michie, 2002; Nyberg et al., 2014).

The detection of stress at the workplace represents therefore a prerequisite not only to anticipate any negative health effects in the long term but also to offer just-in-time (organizational) health promotion interventions (Mattke et al., 2013; Nahum-Shani et al., 2015; Nahum-Shani et al., 2016). Beyond organizational and psychological barriers to directly report stress to colleagues or supervisors (Demerouti et al., 2009), there exist several self-report instruments for measuring individual stress levels (Cohen et al., 1983; Demerouti et al., 2003; Kessler et al., 2003; Siegrist et al., 2009). However, applying these instruments has two major limitations: First, stress polls are usually conducted only two times per health intervention with several weeks or even months in between, if at all health promotion programs are implemented by corresponding organizations (Tims et al., 2013a; Tims et al., 2013b). Thus, the resolution of stress data is too low, i.e. short-term episodes of stress with serious negative health outcomes cannot be identified reliably. Second, data collection with self-reports, if collected in higher frequencies, is time-consuming, obtrusive and costly due to data collection, data analysis and data interpretation activities.

Recent Information Systems literature, however, indicates significant relationships between motor activity measured by mouse movements and emotional states (Grimes et al., 2013; Hibbeln et al., 2017). In Human-Computer Interaction literature, it has been even shown a relationship between mouse movements and perceived stress (Sun et al., 2014). However, with 71% the classification accuracy is limited, as is the underlying mass-spring-damper model regarding the set of potential mouse features from which perceived stress can be derived (ibid.). Our research question is therefore:

*Which features of mouse movements are significantly related to the degree of perceived stress such that a Stress Detection Service (SDS) is able to classify high versus low stress accurately?*

In order to address this question, we apply design science research (Astor et al., 2013; Gregor and Hevner, 2013; Hevner et al., 2004; Peffers et al., 2007; vom Brocke et al., 2013) with the goal to design, implement and evaluate a SDS that senses the degree of perceived stress in knowledge workers solely based on mouse movements. Combining the Job Demands-Resource (JD-R) model (Bakker and Demerouti, 2007) and van Gemmert and van Galen (1997)’s stress theory with a particular focus on its neuromotor noise concept as justificatory knowledge, we built the first two SDS modules that sense mouse movements and perceived stress. Mouse movement data was collected during a laboratory experiment and used to derive mouse movement features that can classify high versus low levels of perceived stress accurately. These features finally inform the design of a third SDS module, the stress classification module, which is part of our future work.

The remainder of this paper is structured as follows. Next, we present the conceptual foundations and hypotheses of the current work from which the design requirements for the SDS are derived. Then, the current implementation of the SDS is described. With a focus on the planned classification module and the identification of relevant mouse features, we describe a lab experiment that was conducted to test our hypotheses. We finally present the results of the lab experiment, discuss the limitations of this work and provide an outlook of a second build-and-evaluate loop of the envisioned SDS in the context of a longitudinal field study.

Finally, it must be noted that details about the motivation and related work of the envisioned SDS are presented elsewhere (Kowatsch et al., 2015). By contrast, the current paper is distinct from that protocol in so far as it adopts a design science approach and thus presents a build-and-evaluate loop including theoretically derived design requirements, a description of the design artefact, i.e. the SDS, and presents and discusses the actual empirical results from a first lab experiment.

## 2 Conceptual Foundations and Hypotheses

In this section, we present the conceptual foundations and hypotheses of the current work with the overall goal to outline the theoretical underpinnings of the SDS. The conceptual foundations are informed by the JD-R model (Bakker and Demerouti, 2007) and stress theory (van Gemmert and van Galen, 1997) with workplace stress being the focal theoretical construct. The JD-R model is adopted because it provides a holistic view of predictors and outcomes of work-related stress, i.e. it describes the broader context of this research. By contrast, van Gemmert and van Galen's stress theory is considered appropriate because it explains the relationship between workplace stress, variations in the motor system through the concept of neuromotor noise, and human performance. Thus, it represents the justificatory knowledge for the relationship between the motor system of an individual, i.e. the proxy to mouse movements, and the degree of workplace stress. An overview of the current work's research framework is depicted in Figure 1. The rationale of the theoretical constructs and their relationships are described in the following two subsections.

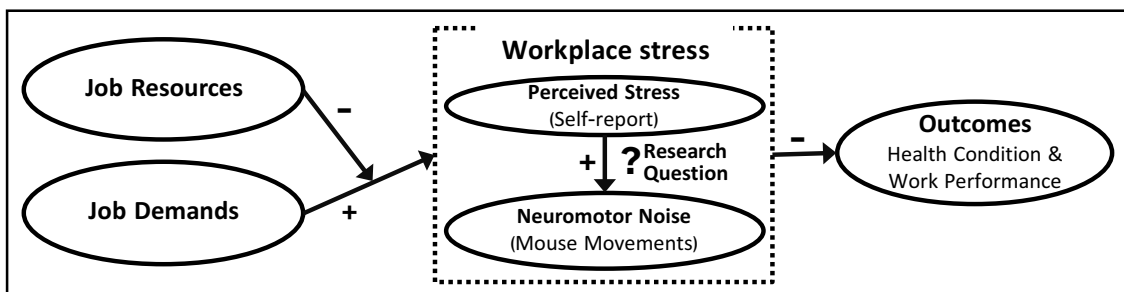


Figure 1. Research framework. Note: arrows indicate positive (+) and negative (-) relationships

### 2.1 Job Demands-Resource (JD-R) Model

The JD-R model proposes that the degree of work-related stress is positively influenced by the degree of mental, emotional and physical job demands and that job resources of employees, for example, their skills, negatively moderate this relationship (Bakker and Demerouti, 2007). Job resources can compensate the job demands in a way that workplace stress is reduced. Or, put in other words, an imbalance of job demands and job resources in which the demands are higher than the resources leads to increased levels of stress perceptions. Job resources do also positively impact the motivation of employees that, in turn, can increase the outcomes. It must be noted that this motivational aspect is not shown in Figure 1 as the focus of this research lies on the measurement of workplace stress only. By contrast, workplace stress is negatively associated with the outcome parameters on various levels such as the individual level, e.g. impacting health, or the organizational level, e.g. work performance (Bakker and Demerouti, 2007; Ganster and Rosen, 2013; Kuoppala et al., 2008; Michie, 2002).

## 2.2 Stress Theory and the Concept of Neuromotor Noise

In contrast to the JD-R literature, in which workplace stress is usually measured via self-report instruments, van Gemmert and van Galen's stress theory (van Gemmert and van Galen, 1997) puts a neurophysiological lens on workplace stress. The theory, which has been already adopted in human-computer interaction research (e.g. Lin and Wu, 2011), suggests that an imbalance of high job demands and low job resources is reflected by increased information processing demands. Moreover, these "increased processing demands (e.g. in dual-task situations) lead to increased levels of neuromotor noise and, therefore, to decreased signal-to-noise ratios in the [motor, the author(s)] system." (van Gemmert and van Galen 1997, p. 1300) Here, neuromotor noise is generated by cognitive activities in the brain. Particularly in high-demand work situations, neuromotor noise results from a competition of individuals' information processing resources. The resulting decrease of the signal-to-noise ratio has direct effects on the motor system, which can be measured by increased variations of human movements, i.e. micro-movements ("shivering"). For example, mouse movements have been shown to be valid proxies of cognitive and affective processing (Grimes et al., 2013; Maehr, 2008; Zimmermann et al., 2003) as they provide "continuous streams of output that can reveal ongoing dynamics of processing, potentially capturing the mind in motion with fine-grained temporal sensitivity." (Freeman et al., 2011, p. 1)

Against this background, if each mouse movement is guided by a target, for example, to hit the send button of an email application or to drop a file in a folder, then neuromotor noise may probably add so called micro-movements to the mouse trajectory that are not consciously processed by the individual. That is, we would not only expect mouse movement deviations from an *optimal* trajectory as shown Figure 2, but also an increase in mouse movement speed due to these micro-movements that increase the distance to the mouse movement target.



Figure 2. Hypothesized mouse trajectories in low (top) versus high (bottom) stress situations.

The overall time, however, to reach the target might be still comparable to a low stress situation without any micro-movements induced by neuromotor noise. We therefore formulate our hypotheses related to our research question as follows:

- H1:** Higher (lower) deviations from an optimal mouse trajectory are related to high (low) perceived stress.
- H2:** Higher (lower) speed of mouse movements are related to high (low) perceived stress.

## 3 Design Requirements for a Stress Detection Service (SDS)

The theoretical considerations and hypotheses allow us now to derive the requirements for the design of our envisioned SDS. First and foremost, the target audience and work context of the SDS must be defined to which the SDS should be tailored to. Consistent with related work (e.g. Koldijk et al., 2013; Sappelli et al., 2014), we consider knowledge workers in the current research as the target population, i.e. employees that are seen as "the most valuable asset of a 21st-century institution (whether business or non-business)" (Drucker, 1999, p. 79). Knowledge workers interact significantly with their mouse rendering these interactions a mirror of their work and relevant proxy of variations in the motor system (Sappelli et al., 2014). Since measuring stress itself could increase workplace stress and thus, negatively impact the work performance, the sensing must be designed in a way that is not obtrusive. The first requirement is therefore defined as follows:

**Requirement 1:** *The design artifact (SDS) has to provide an unobtrusive interface to sense mouse movements of knowledge workers that predominantly use their mouse to perform their work-related activities.*

Second, besides promising results with regard to the relationship between mouse movements and (a) perceived emotions (e.g. Grimes et al., 2013; Hibbeln et al., 2017; Koldijk et al., 2013) and (b) perceived stress (e.g. Sun et al., 2014), it is not yet conclusive whether neuromotor noise influences the motor system in a way such that the degree of workplace stress perceptions can be measured accurately, i.e. with accuracy levels close to 100%. Thus, the design artifact needs to measure the ground truth, i.e. perceived stress, too. This is not only a requirement to identify relevant features of mouse movements, as hypothesized above. It also makes the “static feature” SDS a rather self-learning SDS that improves its stress detection accuracy with each additional labeled mouse movement episode (Barata et al., 2016; Pejovic and Musolesi, 2015). However, bearing in mind the burden of intervention that increases with each labelling activity, i.e. indicating the perceived level of stress by the knowledge worker, it is proposed to conduct a lab experiment (see below) to initially derive relevant features of mouse movements that are correlated to perceived stress levels, before the SDS is used in the field to improve its feature set to increase classification accuracy in a training phase.

**Requirement 2:** *The design artifact (SDS) has to provide an interface to sense mouse movements of knowledge workers that predominantly use their mouse to perform their work-related activities.*

Third, the SDS must be able to classify the mouse movements into workplace stress perceptions (e.g. binary high vs. low stress levels). Two potential features of mouse movements have been hypothesized above and are evaluated in Section 5. In addition, to have an effect and to tackle any serious health condition in an early stage, the classification results must be communicated through an appropriate interface to various SDS endpoints. These endpoints could be quantified-self feedbacks tailored to the employee, relaxation tips to decrease short term episodes of stress or an (organizational) health intervention service, for example, offered by an organization’s physician or medical consultant (Cook et al., 2007; Mattke et al., 2013). The final requirement is therefore formulated as follows:

**Requirement 3:** *The design artifact (SDS) has to classify the mouse movements into classes of workplace stress and to communicate the result to health intervention services.*

## 4 Implementation of the SDS

The SDS consists of two sensing modules, one for mouse movements (Requirements 1) and one for perceived stress (Requirement 2). Moreover, it includes a stress classification module with an application programming interface that provides a trigger to (external) health intervention services (Requirement 3). An overview of the modules and their relationships is depicted in Figure 3.

In the current SDS implementation, the *sensing module for mouse movements* is written as a Java application that runs on several operating systems. It gathers a data stream of mouse coordinates and meta information (e.g. single-click, double click, etc.), attaches timestamps to it, and appends these data tuples to a text file. While previous work (Freeman and Ambady, 2010; Sun et al., 2014; Visser et al., 2004) employed artificially high sampling rates, for example, 500Hz and even higher, and special mouse hardware and software drivers to increase the resolution of mouse movements, we decided to utilize a non-artificial but standard sampling rate of today’s operating systems and of-the-shelf computer mouse hardware, i.e. about 125Hz on average. This does not only add external validity to the potential findings of the current work – e.g. artificially down-sampling high resolution mouse movement data as done by Sun et al. (2014) generates a potential systematic empirical bias – but may allow organizations to use standard mouse hardware they have already in place.

The *sensing module for perceived stress* relies on LimeSurvey, an open source platform for self-report data collection purposes (LimeSurvey.org). We have written a script for LimeSurvey that stores the timestamp of perceived stress self-report data with a pre-defined label. This allows us to synchronize

the self-report data on perceived stress at a particular point in time with the data stream of mouse movements for later analyses. This sensing module provides the ground truth to identify the relevant features of mouse movements. It is therefore in particular relevant for the design and training phase of the SDS. However, when the feature set results in accurate stress classifications it is no longer needed. The burden of measuring stress would therefore disappear completely over time.

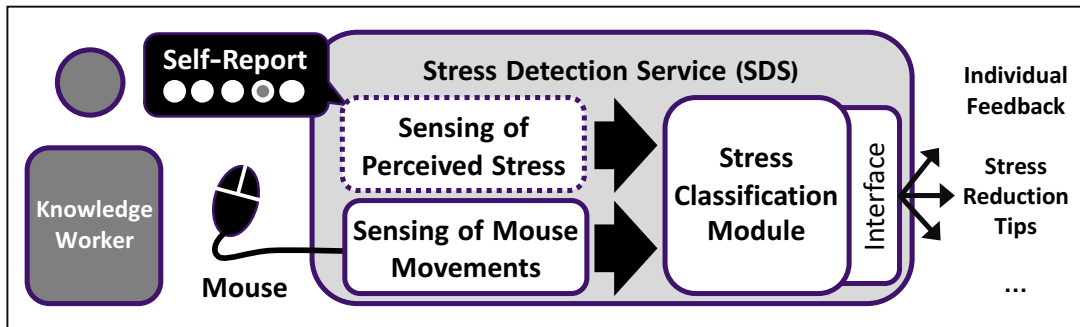


Figure 3. Schematic overview of the Stress Detection Service architecture. Note: The dotted sensing module is required only for the identification of features of mouse movements that are able to accurately classify work-related stress. Thereafter, this module can be used optionally to increase the classification accuracy with regard to individual employees in a training phase, too.

Finally, the *stress classification module* with an application programming interface is currently implemented in the form of a MatLab (R2016a) script.<sup>1</sup> With regard to the implementation described above it becomes obvious that the SDS implementation is currently tailored to controlled evaluation environments in the lab. However, if the evaluation indicates accurate classification results based on mouse movements, we plan to implement an SDS that better fits to a field and (organizational) setting, for example, an SDS that uses a database for storing the mouse coordinates and timestamps, and provides an appropriate user interface that does not distract from work activities.

## 5 Evaluation of Relevant Features of Mouse Movements

In this section, we describe a laboratory experiment that has the objective to identify features of mouse movements that are significantly related to perceived stress levels. For this purpose, we use the two hypothesized features from Section 2.2, i.e. the average deviation from an optimal mouse trajectory (H1) and the average speed of mouse movements (H2). Results will inform the design of SDS's stress classification module. We first outline the experiment and measures, before the results are reported.

### 5.1 Experimental Procedure and Measures

An overview of the lab experiment is shown in Figure 4 (left). The experiment was approved by the ethical committee of the authors' institution. After subjects gave their informed consent, they were randomly assigned to an experimental group (high stress condition) and a control group (low stress condition). They were seated in front of a computer screen and were given a mouse to work with until the end of the experimental session. In addition, physiological arousal was measured throughout the experimental session with the help of a skin conductance (SC) sensor on the non-dominant hand to control for any physiological differences in the subjects that might impact the neuromotor noise effect. For that purpose, a medical device (MindMedia's NeXus-10) was used and the total sum of standardized SC reactions (SCR) for each part of the experiment that involved mouse movements was calculated (see

<sup>1</sup> The authors are happy to provide full access to the source code and the compiled version of the sensing module of the current SDS implementation, the LimeSurvey and MatLab scripts.

Boucsein, 2012, p. 181). To derive the SCR values, we decomposed the electrodermal activity data by applying nonnegative deconvolution as described by Benedek and Kaernbach (2010).

Then, subjects read a cover story, which introduced a fictive objective of the study, i.e. to measure subjects' cognitive performance. All mouse movements were recorded from this point in time by the corresponding sensing module of the SDS.

After subjects have read the cover story, they were asked to follow the instructions related to a square task (Square Task 1). This square task was designed to test the hypotheses, i.e. to provide subjects with pre-defined mouse targets and trajectories. This first square task involved listening to calm music for one minute while tracing the 550px edges of a square on the screen using the mouse. Subjects were given the instruction to trace the edges as accurate as possible while maintaining a swift pace. They were also asked to perform a double click on each corner they pass, which induces a color change of the corner, a visual feedback for interaction success. Figure 4 (right) shows the visual instructions of this square task. Due to the characteristics of this square task, we could test the hypotheses regarding horizontal and vertical mouse movements separately in addition to their composite, i.e. the average of the hypothesized features in both dimensions. An overview of the derived features of mouse movements is provided in Table 1. Finally, the square task allowed us also to derive an objective performance measure from the mouse coordinates, i.e. the number of edges an individual could trace within the pre-defined time. This control variable is used to test the negative impact of stress on task performance as proposed by the JD-R model, and thus, to add external validity to the experimental procedure and findings.

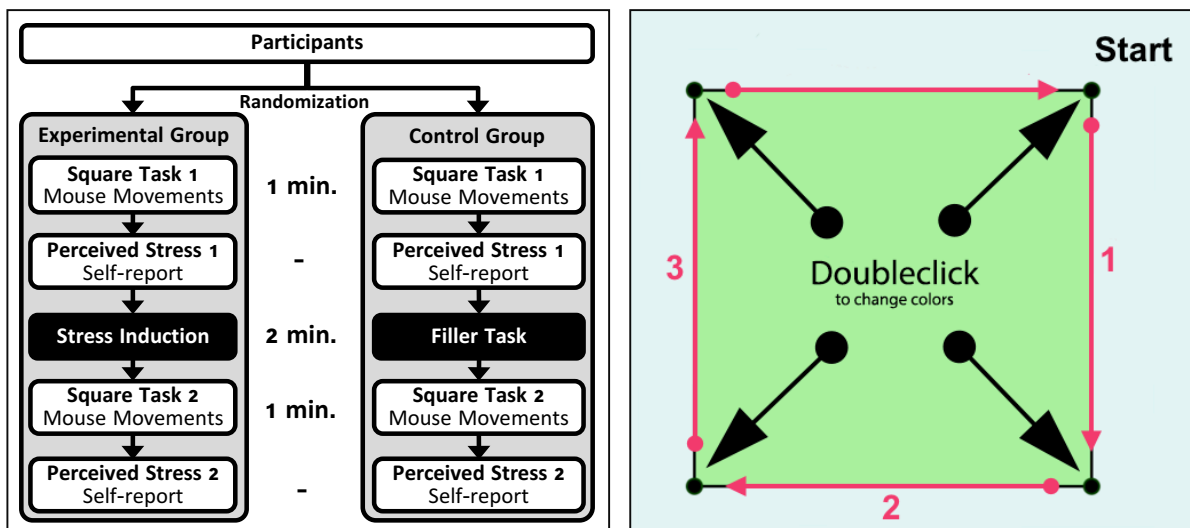


Figure 4. Study procedure (left) and visual instructions of the square task (right).

Consistent with prior work (Sun et al., 2014), participants were then asked to rate their perceived stress level on an 11-point Likert scale ranging from 0 (not stressed at all) to 10 (extremely stressed) with the corresponding sensing module of our SDS (Perceived Stress 1).

After this baseline measurement, stress was induced in the subjects of the experimental group (Stress Induction), whereas a filler task was given to the subjects of the control group (Filler Task). For the stress induction, we used techniques related to the Trier Social Stress Test (Kirschbaum et al., 1993). Accordingly, subjects of the experimental group were asked to prepare a five-minute presentation about a self-assessment of their cognitive capabilities, which, according to the instructions, would be critically examined by an academic expert on cognitive performance awaiting their presentation in the room next door. Furthermore, a video camera was placed beside the computer screen, the recording was started and subjects were told that the academic expert is observing them from now on.

After this instruction, subjects had two minutes to prepare the presentation with a pen and paper, thus no biasing mouse interactions were required. By contrast, subjects of the control group were given the filler task to write down with a pen calm moments of their last holidays. They were also given two minutes for this task.

Thereafter, subjects of both groups were asked again to conduct the one-minute square task (Square Task 2). In addition to the instructions of Square Task 1 both groups were now asked to memorize the different colors of the corners of the square after each double click together with the number of edges they successfully traced during the task to increase cognitive load and to counterbalance, to some degree, any learning effects from Square Task 1. Moreover, instead of relaxation music, both groups were exposed to acoustic office background noise (e.g. talking, phone ringing or and printing sounds from stylus printers) to create a more realistic acoustic workplace scenario. After Square Task 2 subjects were again asked to provide their perceived stress level (Perceived Stress 2).

Finally, all subjects were debriefed and the true objective of the study was revealed. Subjects of the experimental group were additionally asked to perform a relaxation exercise to calm down.

Feature	Description
Mouse $AD_{x/y}$	Averaged deviation from the horizontal ( $AD_x$ ) and vertical ( $AD_y$ ) edges of the square in pixels
Mouse $AD_{total}$	Average of $AD_x$ and $AD_y$
Mouse $AS_{x/y}$	Average speed on the horizontal ( $AS_x$ ) and vertical ( $AS_y$ ) mouse trajectory during the square task in pixels per millisecond
Mouse $AS_{total}$	Average of $AS_x$ and $AS_y$

Table 1. Derived feature set from mouse movements according to H1 and H2

## 5.2 Results

Overall, 19 students from a business university participated in the experiment. We had to drop one subject from further analyses due to incomplete physiological data. The descriptive statistics of the remaining 18 subjects are listed in Table 2. Boxplots of stress perceptions after the two square tasks for each of both groups are shown in Figure 5 (left). In order to test the success of the stress induction, we conducted a robust mixed ANOVA on the perceived stress data (see Field et al., 2012, p. 648; 10% trimming,  $t_{split}()$ ). As a prerequisite, Levene's test revealed that the variances of the perceived stress (PS) variables are similar for the control and the experimental group ( $F_{PS1}(1, 17) = 0.60, ns$ ;  $F_{PS2}(1, 17) = 0.01, ns$ ). The assumption of homogeneity of variance is thus not violated. ANOVA results show that the group allocation (main effect) is not significant ( $Q = 0.21, ns$ ), but both the time (main effect of PS1 and PS2) ( $Q = 6.69, p = .01$ ) and the interaction term of group allocation and time ( $Q = 6.71, p = .01$ ). Post-hoc tests revealed that there are no significant time differences in the control group ( $t(8) = -1.79, ns$ ) but only in the experimental group ( $t(8) = -3.49, p < .01$ ). Overall, this indicates that the induction of stress was successful.

Furthermore, boxplots of physiological arousal measured by the total sum of standardized skin conductance reactions during the two square tasks for each of both groups are shown in Figure 5 (right). Here, we found no significant differences in physiological arousal between both groups ( $Q=.04, p = .85$ ) and also no interaction effect between group and square task factors ( $Q=.04, p = .85$ ) by applying a robust mixed ANOVA with R's WRS2 package ( $bwtrim$  function). That is, physiological arousal did not bias any neuromotor noise effect between the groups. However, we found a significant difference in physiological arousal between the two square tasks ( $Q=17.9, p = .002$ ) which can be explained by the increase of cognitive load in both groups by memorizing the number of traced edges and colors of the square corners.



Finally, we evaluated the relative performance of both groups based on the number of edges that each subject traced during the square tasks. Results indicate, though statistically not significant ( $p > .05$ ), that the increase in performance from the first to the second square task can be attributed to learning effects. However, the increase in performance was lower for the experimental group (12.1%) compared to the control group (22.4%). This finding is consistent with the negative effects of workplace stress as outlined in the JD-R model and thus, adds external validity to the results of this lab experiment.

Variable	Experimental Group	Control Group
N	9	9
Gender	8 male, 1 female	7 male, 2 female
Age	23.78 (3.31)	26.22 (1.72)
<b>Square Task 1</b>		
Performance (Number of Edges)	19.06 (5.88)	22.52 (7.51)
Perceived Stress 1	2.67 (2.18)	2.78 (2.59)
Physiological Arousal (SCR)	994 (327)	977 (237)
Mouse AD <sub>x</sub>	6.24 (2.26)	6.89 (1.43)
Mouse AD <sub>y</sub>	5.01 (1.47)	5.60 (1.77)
Mouse AD <sub>total</sub>	5.62 (1.79)	6.24 (1.44)
Mouse AS <sub>x</sub>	0.17 (0.05)	0.22 (0.06)
Mouse AS <sub>y</sub>	0.19 (0.05)	0.23 (0.06)
Mouse AS <sub>total</sub>	0.18 (0.05)	0.23 (0.06)
<b>Square Task 2</b>		
Performance (Number of Edges)	23.33 (6.60)	27.56 (4.93)
Perceived Stress 2	6.11 (2.03)	3.67 (2.06)
Physiological Arousal	1514 (358)	1493 (294)
Mouse AD <sub>x</sub>	6.58 (3.00)	8.01 (2.90)
Mouse AD <sub>y</sub>	4.46 (1.33)	5.25 (1.81)
Mouse AD <sub>total</sub>	5.52 (2.13)	6.63 (2.18)
Mouse AS <sub>x</sub>	0.21 (0.07)	0.26 (0.06)
Mouse AS <sub>y</sub>	0.22 (0.06)	0.27 (0.04)
Mouse AS <sub>total</sub>	0.21 (0.06)	0.27 (0.05)

Table 2. Descriptive Statistics. Note: Means (Std. Dev.) are only provided for quantitative variables

Against this background, we now focus on the mouse movement features of the experimental group. Regarding the mouse trajectories and to provide an intuition about potential differences in low vs. highly stressed participants, Figure 6 visualizes mouse movement patterns of participants with a perceived stress from the low and the high end of the scale. Here, the differences in the deviation of the mouse trajectories to the optimal square edges is consistent with the hypothesized effects of van Gemmert and van Galen's stress theory about the concept of neuromotor noise and can be visually clearly observed.

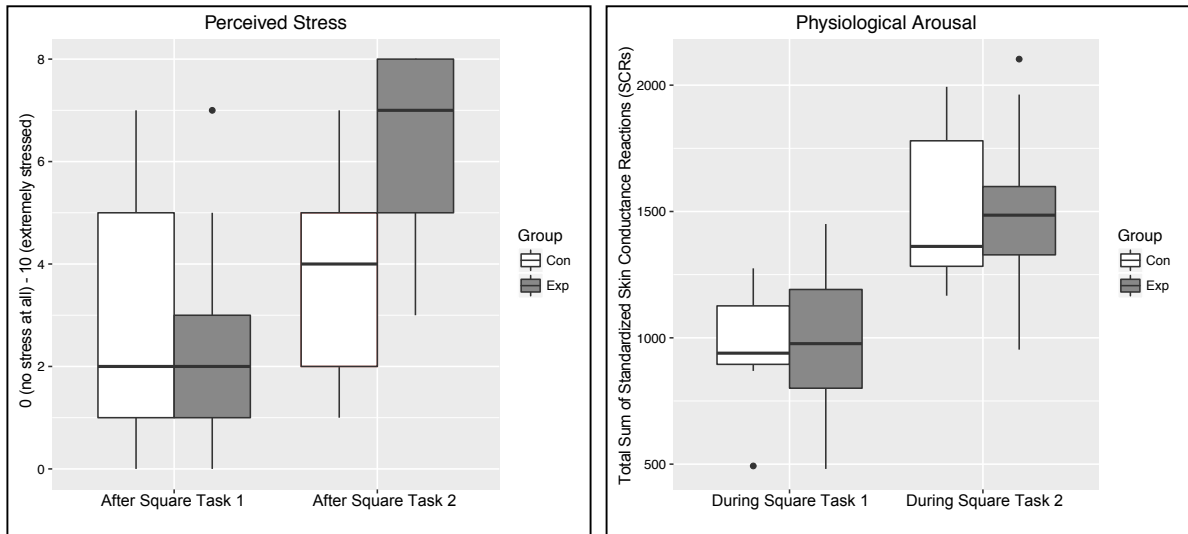


Figure 5. Boxplots for perceived stress (left) and physiological arousal (right). Note: “Con” stands for Control group and “Exp” stands for experimental group

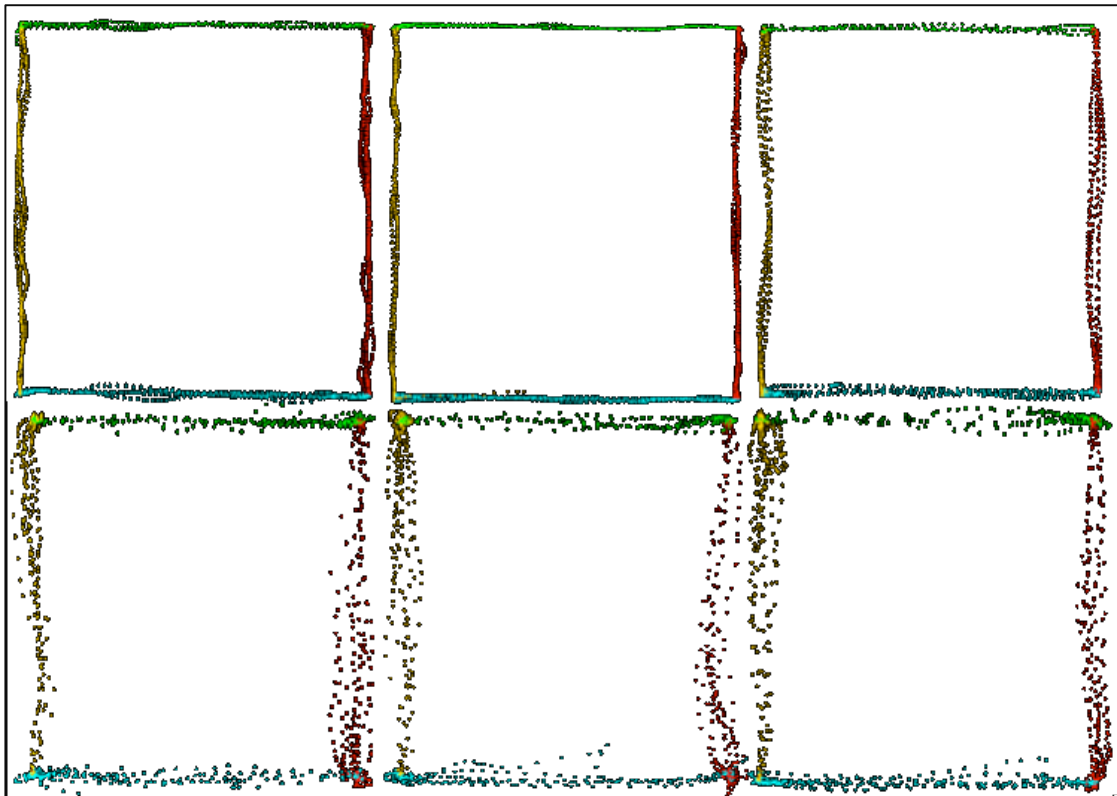


Figure 6. Six mouse movement patterns of subjects with low (top, baseline) and high (bottom, after stress induction) perceived stress levels.

To test our hypotheses, we first calculated Pearson’s product-moment correlations (Cohen et al., 2003) of perceived stress and the six mouse movement features for both, the experimental group and the control group. Table 3 summarizes the correlation coefficients for all six features. In the experimental group, we found five out of six (83%) positive and significant correlations with large and close to large effect sizes. By contrast, we found no significant correlations at the .05 level for the same relationships in the control group. We therefore conclude that these results support our hypotheses.

Pearson Correlation	Experimental Group	Control Group
Mouse AD <sub>x</sub> x perceived stress	.49* ( $p = .04$ )	-.05 ( $p = .85$ )
Mouse AD <sub>y</sub> x perceived stress	.35 ( $p = .15$ )	-.10 ( $p = .68$ )
Mouse AD <sub>total</sub> x perceived stress	.46* ( $p < .05$ )	-.08 ( $p = .75$ )
Mouse AS <sub>x</sub> x perceived stress	.54* ( $p = .02$ )	-.06 ( $p = .80$ )
Mouse AS <sub>y</sub> x perceived stress	.58* ( $p = .01$ )	-.07 ( $p = .77$ )
Mouse AS <sub>total</sub> x perceived stress	.57* ( $p = .01$ )	-.07 ( $p = .79$ )

Table 3. Pearson correlation coefficients of the six mouse movement features and perceived stress levels. Note: \* indicates  $p < .05$

To cross-validate these findings with respect to differences between both square tasks, we further conducted Wilcoxon signed-rank tests for the experimental and control group to compare the medians of the six mouse movement features. The results are shown in Table 4. They indicate that there are significant differences in all six features' medians between Square Task 1 vs. Square Task 2 for the subjects of the experimental group at the .95% confidence interval. By contrast, we could not identify any significant differences for subjects of the control group. Again, these results support H1 and H2.

Wilcoxon Signed-Rank Test	Experimental Group	Control Group
Difference of Mouse AD <sub>x</sub> between both square tasks	$p = .01^*$	$p = .96$
Difference of Mouse AD <sub>y</sub> between both square tasks	$p = .05^*$	$p = .89$
Difference of Mouse AD <sub>total</sub> between both square tasks	$p = .03^*$	$p = 1.0$
Difference of Mouse AS <sub>x</sub> between both square tasks	$p = .02^*$	$p = .68$
Difference of Mouse AS <sub>y</sub> between both square tasks	$p = .01^*$	$p = .55$
Difference of Mouse AS <sub>total</sub> between both square tasks	$p = .007^{**}$	$p = .68$

Table 4. Difference of the six mouse movement features between low vs. high perceived stress by applying Wilcoxon signed-rank tests. Note: \* / \*\* indicates  $p < .05 / .01$

To finally test our hypotheses with respect to SDS's classification module, i.e. to assess the applicability of the proposed mouse movement features to predict high versus low stress levels, we finally defined this task as a binary classification problem. Using receiver operating characteristic (ROC) analysis, we assessed the univariate classification performance of each feature for separating the two square tasks in both groups. Figure 7 shows the corresponding ROC curves for each of the six mouse movement features. Among the area under the curve (AUC), Table 5 lists accuracy, sensitivity and specificity scores of the optimal operating point on the ROC curve for the experimental group only, because the AUCs for the control group are close to .50 (see Table 5, column two in brackets) indicating a random and thus, inadequate classification.

While the AUCs for optimal line deviation features are consistently lower than features derived from mouse movement speed, performance at the optimal operating point is equal for all proposed features. The AD<sub>y</sub> feature and the AS<sub>total</sub> feature have the highest sensitivity score with 0.89, trading off specificity at 0.67. With AUC values starting at .75 and clearly above the random value of .50, and an average classification accuracy of .78 for both feature sets, i.e. average deviation and average speed, the empirical data supports the proposed hypotheses. It can be therefore concluded that both (a) average deviation from an optimal mouse trajectory and (b) average speed of mouse movements are potentially relevant mouse movement features of SDS's classification module.

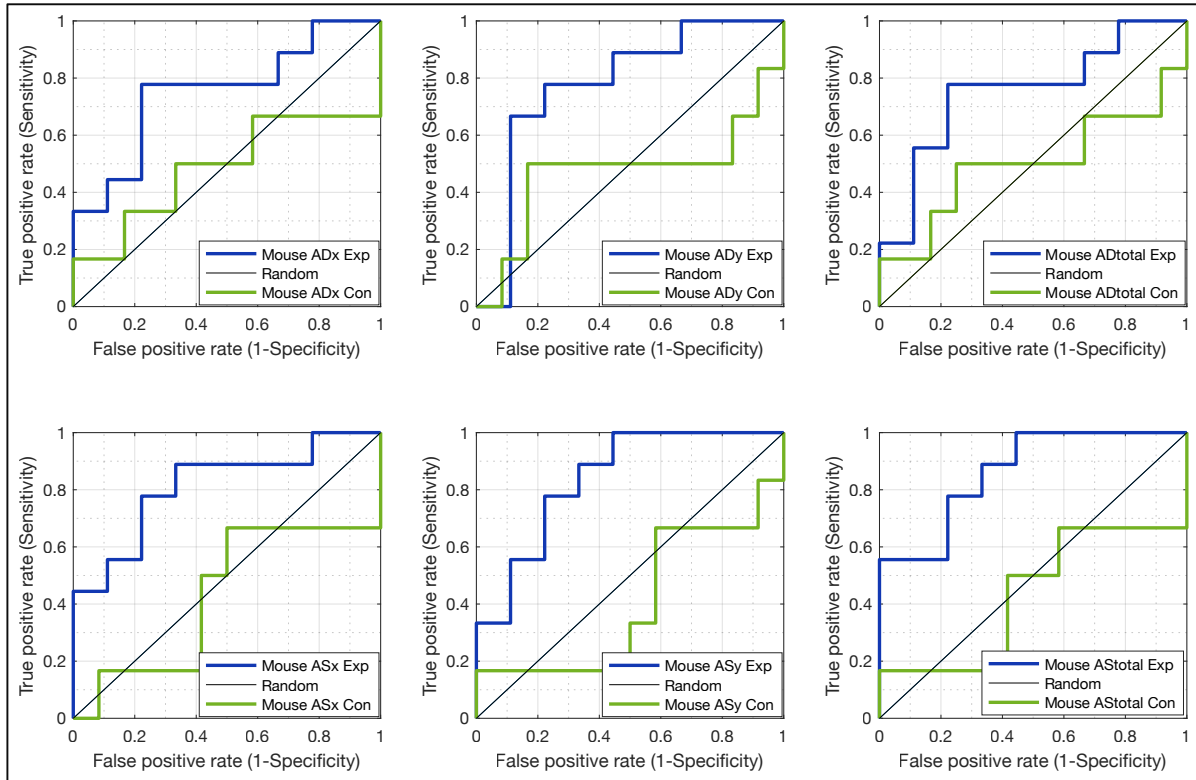


Figure 7. Receiver operator curves for the six mouse movement features for experimental group (Exp) and control group (Con).

Feature	Area Under the Curve	Accuracy	Sensitivity	Specificity
Mouse AD <sub>x</sub>	0.75 (Control Group 0.49)	0.78	0.78	0.78
Mouse AD <sub>y</sub>	0.78 (Control Group 0.47)	0.78	0.89	0.67
Mouse AD <sub>total</sub>	0.75 (Control Group 0.50)	0.78	0.78	0.78
Mouse AS <sub>x</sub>	0.81 (Control Group 0.43)	0.78	0.78	0.78
Mouse AS <sub>y</sub>	0.84 (Control Group 0.40)	0.78	0.78	0.78
Mouse AS <sub>total</sub>	0.86 (Control Group 0.43)	0.78	0.89	0.67

Table 5. Analysis of the receiver operating characteristic. Note: Accuracy, sensitivity and specificity values are provided for the optimal operating point; AUC: Area under the curve

## 6 Discussion

This research presents a first steps towards a novel Stress Detection Service (SDS), which has the objective to unobtrusively measure the degree of work-related stress in knowledge workers solely based on mouse movements. Relying on the conceptual foundations of van Gemmert and van Galen's stress theory and the JD-R model (Bakker and Demerouti, 2007), experimental results indicate that two feature sets of mouse movements, i.e. average deviation from an optimal mouse trajectory and average mouse movement speed, can classify high versus low stress with an overall accuracy of 78%. With 78%, the classification accuracy of our SDS lies also slightly above the accuracy of prior work (Sun et al., 2014, 71%) but requires only one feature to reach that accuracy score. That is, the feature set is more parsimonious compared to the work of Sun et al. as they have adopted the mass-spring-damper model with five features as justificatory knowledge for their classification algorithm development.

Moreover, the current work presents an experimental design with a control group compared to Sun et al. who chose a within-subjects design only. That is, we could not only test differences of repeated measures but also among the experimental and control groups by our within- and between-subjects design. The current work complements therefore the work of Sun et al. (2014) and shows for the very first time that mouse movements are indeed related to stress perceptions but are not necessarily related to another physiological indicator of arousal and stress, i.e. skin conductance responses, a common view of previous research (e.g. Chittaro and Sioni, 2014a; Liao et al., 2006; Moody and Galletta, 2015; Riedl, 2013; Riedl et al., 2013; Schnädelbach et al., 2012; Sun et al., 2014). That is, physiological arousal increased significantly in both groups from Square Task 1 to Square Task 2 but there was only a significant difference of the mouse movement features in the experimental group (see Table 4). Consistent with prior work (e.g. Grimes et al., 2013), this indicates that mouse movements might have the potential to differentiate between positive vs. negative arousal (and stress), a quality physiological arousal measured by skin conductance reactions alone is not capable of (e.g. Sheng and Joginapelly, 2012).

Another aspect must be pointed out that seems to be relevant for measures in general and thus, also for physiological mouse movements. There seems to be a discrepancy in the descriptive statistics of Table 2 (e.g. Mouse  $AD_{total} = 5.62$  during Task 1 vs.  $5.52$  during Task 2) and our hypotheses. One rationale of this observation, among the fact that stress detection via mouse movements does not work for everybody, may lie in intra-individual differences of subjects (e.g. Kehr and Kowatsch, 2015; Pitariu and Ployhart, 2010; Ployhart and Vandenberg, 2010; Tams et al., 2014; vom Brocke and Liang, 2014). That is, individuals' physiological reactions to externally induced stress may differ systematically, for example, with regard to their personality traits (Bibbeya et al., 2013) or the degree they are chronically stressed (Jones et al., 2016). It is therefore recommended to control for any additional physiological and psychological factors in future studies and to apply self-learning techniques (e.g. reinforcement learning) that adapt the feature set and their classification model to each individual knowledge worker over time (Kaelbling et al., 1996; Szita and Szepesvari, 2010).

Even though this work reports promising results regarding a SDS, two major limitations of the current work are the artificial square task and the limited study sample from which general findings cannot be drawn, particularly, with respect to field settings. Moreover, the binary classification approach adopted in this work yielded accurate results, but this demonstrates only a first step towards more complex models. That is, a combination of features can act as a base for decision models trained by machine learning algorithms to go beyond the binary classification case and to further increase performance levels.

## **7 Future Work**

In our future work, we use the findings of the current work to adapt the current SDS to a field setting and assess its technical, organizational and legal feasibility. In line with additional design requirements derived from knowledge workers of one of our industry partners, we will combine all current SDS modules into one integrated desktop application and compare detection accuracies with prior work.

Moreover, we will assess the accuracy of SDS's stress detection module with respect to multi-class classification problems (e.g. low, medium and high stress). We will also add a quantified-self module that allows self-reflection of perceived stress data in the form of a visual diary. This quantified-self module aims not only at improving self-determined stress reflection and regulation capabilities of employees but it has also the goal to motivate knowledge workers to continuously train their individual classification model such that the obtrusive and time-consuming self-report activities can be replaced by mouse movements in the long term.

We finally plan to integrate and evaluate existing interventions for the management of stress at the workplace (Ryan et al., 2017). Examples are relaxation exercises (Chittaro and Sioni, 2014a), breathing trainings (Chittaro and Sioni, 2014b), health literacy interventions (Jacobs et al., 2016) or job crafting interventions (Kooij et al., 2017) that are triggered just-in-time by our envisioned SDS with the overall goal to prevent chronic stress and serious mental health problems such as major depression.

## References

- Astor, P. J., M. T. P. Adam, P. Jercic, K. Schaaff and C. Weinhardt (2013) Integrating Biosignals into Information Systems: A NeuroIS Tool for Improving Emotion Regulation, *Journal of Management Information Systems* 30 (3), 247-277.
- Bakker, A. B. and E. Demerouti (2007) The Job Demands-Resources Model: State of the Art, *Journal of Managerial Psychology* 22 (3), 309-328.
- Barata, F., T. Kowatsch, P. Tinschert and A. Filler (2016) Personal MobileCoach: Tailoring Behavioral Interventions to the Needs of Individual Participants, in: *Proceedings of the UBICOMP 2016 Workshop Designing, Developing, and Evaluating The Internet of Personal Health (IoPH)*, Heidelberg, Germany.
- Benedek, M. and C. Kaernbach (2010) Decomposition of skin conductance data by means of nonnegative deconvolution, *Psychophysiology* 47, 647-658.
- Bibbeya, A., D. Carrola, T. J. Roseboom, A. C. Phillipsa and S. R. de Rooijb (2013) Personality and physiological reactions to acute psychological stress, *International Journal of Psychophysiology* 90 (1), 28-36.
- Boedeker, W. and H. Klindworth (2007) Hearts and Minds at Work in Europe: A European Work-related Public Health Report on Cardiovascular Diseases and Mental Ill Health, Essen: BKK Bundesverband.
- Boucsein, W. (2012) *Electrodermal Activity*, (2nd ed.) Heidelberg, Germany: Springer.
- Cartwright, S. and C. L. Cooper (2014), Towards Organizational Health: Stress, Positive Organizational Behavior, and Employee Well-being, in: *Bridging Occupational, Organizational and Public Health: A Transdisciplinary Approach*, G. F. Bauer and O. Hämmig (eds.), Springer, Dordrecht, 29-42.
- Chittaro, L. and R. Sioni (2014a) Affective computing vs. affective placebo: Study of a biofeedback-controlled game for relaxation training, *International Journal of Human-Computer Studies* 72 (8-9), 663-673.
- Chittaro, L. and R. Sioni (2014b) Evaluating mobile apps for breathing training: The effectiveness of visualization, *Computers in Human Behavior* 40, 56-63.
- Cohen, J., P. Cohen, S. G. West and L. S. Aiken (2003) *Applied multiple regression/correlation analysis for the behavioral sciences*, Mahwah, NJ, USA: Lawrence Erlbaum Associates Publishers.
- Cohen, S., T. Kamarck and R. Mermelstein (1983) A Global Measure of Perceived Stress, *Journal of Health and Social Behavior* 24 (4), 385-396.
- Cook, R. F., D. W. Billings, R. K. Hersch, A. S. Back and A. Hendrickson (2007) A Field Test of a Web-based Workplace Health Promotion Program to Improve Dietary Practices, Reduce Stress, and Increase Physical Activity: Randomized Controlled Trial, *Journal of Medical Internet Research* 9 (2:e17).
- Demerouti, E., A. B. Bakker, I. Vardakou and A. Kantas (2003) The convergent validity of two burnout instruments: A multitrait- multimethod analysis, *European Journal of Psychological Assessment* 19 (1), 12-23.
- Demerouti, E., P. M. Le Blanc, A. B. Bakker, W. B. Schaufeli and J. Hox (2009) Present but sick: a three-wave study on job demands, presenteeism and burnout, *Career Development International* 14 (1), 50-68.
- Drucker, P. F. (1999) Knowledge-Worker Productivity: The Biggest Challenge, *California Management Review* 41 (2), 79-94.
- Eurofound (2012) Fifth European Working Conditions Survey, Publications Office of the European Union, Luxembourg.
- Field, A., J. Miles and Z. Field (2012) *Discovering Statistics Using R*, (1st ed.) London: Sage.
- Freeman, J. B. and N. Ambady (2010) MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method, *Behavior Research Methods* 42 (1), 226-241.
- Freeman, J. B., R. Dale and T. A. Farmer (2011) Hand in motion reveals mind in motion, *Frontiers in Psychology* 2 (59), 1-6.

- Ganster, D. C. and C. C. Rosen (2013) Work Stress and Employee Health: A Multidisciplinary Review, *Journal of Management* 39 (5), 1085-1122.
- Geurts, S. (2014), Recovery from Work during Off-job Time, in: *Bridging Occupational, Organizational and Public Health: A Transdisciplinary Approach*, G. F. Bauer and O. Hämmig (eds.), Springer, Dordrecht, 193-208.
- Gregor, S. and A. Hevner (2013) Positioning and Presenting Design Science Research for Maximum Impact, *MIS Quarterly* 37 (2), 337-355.
- Grimes, G. M., J. L. Jenkins and J. S. Valacich (2013) Exploring the Effect of Arousal and Valence on Mouse Interaction, in: *Proceedings of the International Conference on Information Systems (ICIS)*, Milan, Italy.
- Hevner, A. R., S. T. March, J. Park and S. Ram (2004) Design Science in Information Systems Research, *MIS Quarterly* 28 (1), 75-105.
- Hibbeln, M., J. L. Jenkins, C. Schneider, J. S. Valacich and M. Weinmann (2017) How Is Your User Feeling? Inferring Emotion Through Human-Computer interaction Devices, *MIS Quarterly* 41 (1), 1-21.
- Igic, I., B. Brunner, A. Elfering and S. Wieser (2016) Job-Stress-Index 2016, Health Promotion Switzerland, Bern, Switzerland.
- Jacobs, R. J., J. Q. Lou, R. L. Ownby and J. Caballero (2016) A systematic review of eHealth interventions to improve health literacy, *Health Informatics Journal* 22 (2), 81-98.
- Jones, A., J. C. Pruessner, M. R. McMillan, R. W. Jones, G. T. Kowalik, J. A. Steeden, B. Williams, A. M. Taylor and V. Muthurangu (2016) Physiological adaptations to chronic stress in healthy humans – why might the sexes have evolved different energy utilisation strategies?, *The Journal of Physiology* 594 (15), 4297-4307.
- Kaelbling, L. P., M. L. Littman and A. W. Moore (1996) Reinforcement Learning: A Survey, *Journal of Artificial Intelligence Research* 4, 237-285.
- Kehr, F. and T. Kowatsch (2015) Quantitative Longitudinal Research: A Review of IS Literature, and a Set of Methodological Guidelines, in: *Proceedings of the 23rd European Conference on Information Systems (ECIS)*, Münster, Germany, 2015, Münster, Germany.
- Kessler, R. C., C. Barber, A. L. Beck, P. A. Berglund, P. D. Cleary, D. McKenas, N. P. Pronk, G. E. Simon, P. E. Stang, T. B. Üstün, et al. (2003) The World Health Organization Health and Work Performance Questionnaire (HPQ), *Journal of Occupational and Environmental Medicine* 45 (2), 156-174.
- Kirschbaum, C., K. M. Pirke and D. H. Hellhammer (1993) The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting, *Neuropsychobiology* 28 (1-2), 76-81.
- Koldijk, S., M. Sappelli, M. Neerincx and W. Kraaij (2013), Unobtrusive Monitoring of Knowledge Workers for Stress Self-regulation, in: *Proc. of the 21th International Conference on User Modeling, Adaptation, and Personalization (UMAP '13)*, S. Carberry, S. Weibelzahl, A. Micarelli and G. Semeraro (eds.), Springer, Berlin, Germany, 335-337.
- Kooij, D. T. A. M., M. van Woerkom, J. Wilkenloh, L. Dorenbosch and J. J. A. Denissen (2017) Job crafting towards strengths and interests: The effects of a job crafting intervention on person-job fit and the role of age, *The Journal of Applied Psychology* 102 (6), 971-981.
- Kowatsch, T., F. Wahle, A. Filler, F. Kehr, D. Volland, S. Haug, G. J. Jenny, G. Bauer and E. Fleisch (2015) Towards Short-Term Detection of Job Strain in Knowledge Workers with a Minimal-invasive Information System Service: Theoretical Foundation and Experimental Design, in: *Proceedings of the 23rd European Conference on Information Systems (ECIS)*, Münster, Germany.
- Kuoppala, J., A. Lamminpää and P. Husman (2008) Work Health Promotion, Job Well-being, and Sickness Absences - A Systematic Review and Meta-Analysis, *Journal of Occupational Environmental Medicine* 50 (11), 1216-1227.

- Liao, W., W. Zhang, Z. Zhu, Q. Ji and W. D. Gray (2006) Toward a decision-theoretic framework for affect recognition and user assistance, *International Journal of Human-Computer Studies* 64 (9), 847-873.
- Lin, C.-J. and C. Wu (2011) Factors affecting numerical typing performance of young adults in a hear-and-type task, *Ergonomics* 54 (12), 1159-1174.
- Maehr, W. (2008) eMotion: Estimation of User's Emotional State by Mouse Motions, Saarbrücken, Germany: VDM.
- Mattke, S., H. Liu, J. P. Caloyeras, C. Y. Huang, K. R. Van Busum, D. Khodyakov and V. Shier (2013) Workplace Wellness Programs Study, Santa Monica, CA, USA: RAND Corporation.
- Michie, S. (2002) Causes and Management of Stress at Work, *Occupational and Environmental Medicine* 59 (1), 67-72.
- Moody, G. D. and D. F. Galletta (2015) Lost in Cyberspace: The Impact of Information Scent and Time Constraints on Stress, Performance, and Attitudes Online, *Journal of Management Information Systems* 32 (1), 192-224.
- Nahum-Shani, I., E. B. Hekler and D. Spruijt-Metz (2015) Building Health Behavior Models to Guide the Development of Just-in-Time Adaptive Interventions: A Pragmatic Framework, *Health Psychology* 34 (Supplement), 1209-1219.
- Nahum-Shani, I., S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari and S. A. Murphy (2016) Just-in-Time Adaptive Interventions (JITAs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support, *Annals of Behavioral Medicine* ePrint ahead doi:10.1007/s12160-016-9830-8.
- Nyberg, S. T., E. I. Fransson, K. Heikkilä, K. Ahola, L. Alfredsson, J. B. Bjorner, M. Borritz, H. Burr, N. Dragano, M. Goldberg, et al. (2014) Job Strain as a Risk Factor for Type 2 Diabetes: A Pooled Analysis of 124,808 Men and Women, *Diabetes Care* 37 (8), 2268-2275.
- Peffers, K., T. Tuunanen, M. Rothenberger and S. Chatterjee (2007) A Design Science Research Methodology for Information Systems Research, *Journal of Management Information Systems* 24 (3), 45-77.
- Pejovic, V. and M. Musolesi (2015) Anticipatory Mobile Computing: A Survey of the State of the Art and Research Challenges, *ACM Computing Surveys* 47 (3), Article 47.
- Pitariu, A. H. and R. E. Ployhart (2010) Explaining Change: Theorizing and Testing Dynamic Mediated Longitudinal Relationships, *Journal of Management* 36 (2), 405-429.
- Ployhart, R. E. and R. J. Vandenberg (2010) Longitudinal Research: The Theory, Design, and Analysis of Change, *Journal of Management* 36 (1), 94-120.
- Riedl, R. (2013) On the biology of technostress: literature review and research agenda, *The DATA BASE for Advances in Information Systems* 44 (1), 18-55.
- Riedl, R., H. Kindermann, A. Auinger and A. Javor (2013) Computer breakdown as a stress factor during task completion under time pressure: Identifying gender differences based on skin conductance, *Advances in Human-Computer Interaction* Article ID 420169, 1-8.
- Ryan, C., M. Bergin, T. Chalder and J. S. Wells (2017) Web-based interventions for the management of stress in the workplace: Focus, form, and efficacy., *Journal of Occupational health* 59 (3), 215-236.
- Sappelli, M., S. Verberne, S. Koldijk and W. Kraaij (2014) Collecting a dataset of information behaviour in context, in: *Proceedings of the 4th Workshop on Context-Awareness in Retrieval and Recommendation*, Amsterdam, The Netherlands, 26-29.
- Schnädelbach, H., A. Irune, D. Kirk, K. Glover and P. Brundell (2012) ExoBuilding: Physiologically Driven Adaptive Architecture, *ACM Trans. Comput.-Hum. Interact.* 19 (4), 1-22.
- Sheng, H. and T. Joginapelly (2012) Effects of Web Atmospheric Cues on Users' Emotional Responses in E-Commerce, *AIS Transactions on Human-Computer Interaction* 4 (1), 1-24.
- Siegrist, J., N. Wege, F. Pühlhofer and M. Wahrendorf (2009) A short generic measure of work stress in the era of globalization: effort-reward imbalance, *International Archives of Occupational or Environmental Health* 82 (8), 1005-1013.



- Sun, D., P. Paredes and J. Canny (2014) MouStress: detecting stress from mouse motion, in: *Proceedings of the CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Canada, 61-70.
- Szita, I. and C. Szepesvari (2010) Model-based reinforcement learning with nearly tight exploration complexity bounds, in: *Proceedings of the Proceedings of the 27th International Conference on Machine Learning*, Haifa, Israel, 1-8.
- Tams, S., K. Hill, A. O. de Guinea, J. Thatcher and V. Grover (2014) NeuroIS - Alternative or Complement to Existing Methods? Illustrating the Holistic Effects of Neuroscience and Self-Reported Data in the Context of Technostress Research, *Journal of the Association for Information Systems* 15 (10), Article 1.
- Tims, M., A. B. Bakker and D. Derks (2013a) The Impact of Job Crafting on Job Demands, Job Resources, and Well-Being, *Journal of Occupational Health Psychology* 18 (2), 230-240.
- Tims, M., A. B. Bakker, D. Derks and W. van Rhenen (2013b) Job Crafting at the Team and Individual Level: Implications for Work Engagement and Performance, *Group & Organization Management* 38 (4), 427-454.
- van Gemmert, A. W. A. and G. P. van Galen (1997) Stress, neuromotor noise, and human performance: A theoretical perspective, *Journal of Experimental Psychology: Human Perception and Performance* 23 (5), 1299-1313.
- Visser, B., M. P. D. Looze, M. P. D. Graaff and J. H. V. Dieën (2004) Effects of precision demands and mental pressure on muscle activation and hand forces in computer mouse tasks, *Ergonomics* 47 (2), 202-217.
- vom Brocke, J. and T.-P. Liang (2014) Guidelines for Neuroscience Studies in Information Systems Research, *Journal of Management Information Systems* 30 (4), 211-234.
- vom Brocke, J., R. Riedl and P.-M. Léger (2013) Application Strategies for Neuroscience in Information Systems Design Science Research, *Journal of Computer Information Systems* 53 (3), 1-13.
- WHO (2013) Mental health action plan 2013-2020, World Health Organization, Geneva, Switzerland.
- Zimmermann, P., S. Guttormsen, B. Danuser and P. Gomez (2003) Affective computing - a rationale for measuring mood with mouse and keyboard, *International Journal of Occupational Safety and Ergonomics* 9 (4), 539-551.