

Fred D. Davis

René Riedl

Jan vom Brocke

Pierre-Majorique Léger

Adriane B. Randolph *Editors*

# Information Systems and Neuroscience

Gmunden Retreat on NeuroIS 2017



Springer

# **Lecture Notes in Information Systems and Organisation**

Volume 25

## **Series editors**

Richard Baskerville, Atlanta, USA

Marco De Marco, Milano, Italy

Nancy Pouloudi, Athens, Greece

Paolo Spagnoletti, Roma, Italy

Dov Te'eni, Tel Aviv, Israel

Jan vom Brocke, Vaduz, Liechtenstein

Robert Winter, St. Gallen, Switzerland

More information about this series at <http://www.springer.com/series/11237>

Fred D. Davis · René Riedl  
Jan vom Brocke · Pierre-Majorique Léger  
Adriane B. Randolph  
Editors

# Information Systems and Neuroscience

Gmunden Retreat on NeuroIS 2017

 Springer

*Editors*

Fred D. Davis  
Information Systems and Quantitative  
Sciences (ISQS)  
Texas Tech University  
Lubbock, TX  
USA

Pierre-Majorique Léger  
Department of Information Technology  
HEC Montréal  
Montreal, QC  
Canada

René Riedl  
University of Applied Sciences Upper  
Austria  
Steyr, Oberösterreich  
Austria

Adriane B. Randolph  
Department of Information Systems  
Kennesaw State University  
Kennesaw, GA  
USA

and

University of Linz  
Linz  
Austria

Jan vom Brocke  
Department of Information Systems  
University of Liechtenstein  
Vaduz  
Liechtenstein

ISSN 2195-4968                      ISSN 2195-4976 (electronic)  
Lecture Notes in Information Systems and Organisation  
ISBN 978-3-319-67430-8              ISBN 978-3-319-67431-5 (eBook)  
<https://doi.org/10.1007/978-3-319-67431-5>

Library of Congress Control Number: 2017952029

© Springer International Publishing AG 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature  
The registered company is Springer International Publishing AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

NeuroIS is a field in Information Systems (IS) that makes use of neuroscience and neurophysiological tools and knowledge to better understand the development, adoption, and impact of information and communication technologies. The *Gmunden Retreat on NeuroIS* is a leading academic conference for presenting research and development projects at the nexus of IS and neurobiology (see <http://www.neurois.org/>). This annual conference has the objective to promote the successful development of the NeuroIS field. The conference activities are primarily delivered by and for academics, though works often have a professional orientation.

The conference is taking place in Gmunden, Austria, a much frequented health and summer resort providing an inspiring environment for the retreat. In 2009, the inaugural conference was organized. Established on an annual basis, further conferences took place from 2010–2016. The genesis of NeuroIS took place in 2007. Since then, the NeuroIS community has grown steadily. Scholars are looking for academic platforms to exchange their ideas and discuss their studies. The *Gmunden Retreat on NeuroIS* seeks to stimulate these discussions. The conference is best characterized by its “workshop atmosphere.” Specifically, the organizing committee welcomes not only completed research, but also work in progress. A major goal is to provide feedback for scholars to advance research papers, which then, ultimately, have the potential to result in high-quality journal publications.

This year is the third time that we publish the proceedings in the form of an edited volume. A total of 24 research papers are published in this volume, and we observe diversity in topics, theories, methods, and tools of the contributions in this book. The 2017 keynote presentation entitled “Why do we need animals to understand the neurobiology of economic decision-making?” was given by Tobias Kalenscher, professor of comparative psychology at the University of Düsseldorf, Germany. Moreover, we invited the EEG and brain-computer interfacing expert Gernot Müller-Putz, Graz University of Technology, Austria, to give a “hot topic talk” entitled “The Power of EEG: From Single Channel to High Resolution Derivations”. Moreover, a panel entitled “NeuroIS 2007–2017: Hot Topics and the Future of NeuroIS” was held. Altogether, we are happy to see the ongoing progress

in the NeuroIS field. More and more IS researchers and practitioners have been recognizing the enormous potential of neuroscience tools and knowledge.

Lubbock, USA  
Steyr, Linz, Austria  
Vaduz, Liechtenstein  
Montreal, Canada  
Kennesaw, USA  
June 2017

Fred D. Davis  
René Riedl  
Jan vom Brocke  
Pierre-Majorique Léger  
Adriane B. Randolph

# Contents

<b>The Psychophysiological Effect of a Vibro-Kinetic Movie Experience: The Case of the D-BOX Movie Seat</b> . . . . .	1
Horea Pauna, Pierre-Majorique Léger, Sylvain Sénécal, Marc Fredette, François Courtemanche, Shang-Lin Chen, Élise Labonté-Lemoyne and Jean-François Ménard	
<b>Reinforcement Sensitivity and Engagement in Proactive Recommendations: Experimental Evidence</b> . . . . .	9
Laurens Rook, Adem Sabic and Markus Zanker	
<b>The Choice Is Yours: The Role of Cognitive Processes for IT-Supported Idea Selection</b> . . . . .	17
Isabella Seeber, Barbara Weber, Ronald Maier and Gert-Jan de Vreede	
<b>Blood Pressure Measurement: A Classic of Stress Measurement and Its Role in Technostress Research</b> . . . . .	25
Thomas Fischer, Gerhard Halmerbauer, Eva Meyr and René Riedl	
<b>On the Role of Users' Cognitive-Affective States for User Assistance Invocation</b> . . . . .	37
Celina Friemel, Stefan Morana, Jella Pfeiffer and Alexander Maedche	
<b>Measuring and Explaining Cognitive Load During Design Activities: A Fine-Grained Approach</b> . . . . .	47
Barbara Weber, Manuel Neurauter, Andrea Burattin, Jakob Pinggera and Christopher Davis	
<b>How Product Decision Characteristics Interact to Influence Cognitive Load: An Exploratory Study</b> . . . . .	55
Sylvain Sénécal, Pierre-Majorique Léger, René Riedl and Fred D. Davis	



<b>Why and How to Design Complementary NeuroIS and Behavioral Experiments</b> . . . . .	65
Anthony Vance, Jeffrey L. Jenkins, Bonnie Brinton Anderson, C. Brock Kirwan and Daniel Bjornn	
<b>The Impact of Age and Cognitive Style on E-Commerce Decisions: The Role of Cognitive Bias Susceptibility</b> . . . . .	73
Nour El Shamy and Khaled Hassanein	
<b>Expertise as a Mediating Factor in Conceptual Modeling</b> . . . . .	85
Christopher J. Davis, Alan R. Hevner, Élise Labonte-LeMoyné and Pierre-Majorique Léger	
<b>A Neuro-Cognitive Explanation for the Prevalence of Folder Navigation and Web Browsing</b> . . . . .	93
Ofar Bergman and Yael Benn	
<b>Physiological, Psychological, and Behavioral Measures in the Study of IS Phenomena: A Theoretical Analysis of Triangulation Strategies</b> . . . .	101
Kevin Hill and Stefan Tams	
<b>The Psychophysiology of Flow: A Systematic Review of Peripheral Nervous System Features</b> . . . . .	109
Michael T. Knierim, Raphael Rissler, Verena Dorner, Alexander Maedche and Christof Weinhardt	
<b>Predicting Properties of Cognitive Pupillometry in Human–Computer Interaction: A Preliminary Investigation</b> . . . . .	121
Pierre-Majorique Léger, Patrick Charland, Sylvain Sénécal and Stéphane Cyr	
<b>Human Versus Machine: Contingency Factors of Anthropomorphism as a Trust-Inducing Design Strategy for Conversational Agents</b> . . . . .	129
Anna-Maria Seeger and Armin Heinzl	
<b>Affective Processing Guides Behavior and Emotions Communicate Feelings: Towards a Guideline for the NeuroIS Community</b> . . . . .	141
Peter Walla	
<b>Beyond Traditional Neuroimaging: Can Mobile fNIRS Add to NeuroIS?</b> . . . . .	151
Caspar Krampe, Nadine Gier and Peter Kenning	
<b>Decision Inertia and Arousal: Using NeuroIS to Analyze Bio-Physiological Correlates of Decision Inertia in a Dual-Choice Paradigm</b> . . . . .	159
Dominik Jung and Verena Dorner	

**IAT Measurement Method to Evaluate Emotional Aspects of Brand Perception—A Pilot Study** . . . . . 167  
Harald Kindermann and Melanie Schreiner

**Inferring Web Page Relevance Using Pupillometry and Single Channel EEG** . . . . . 175  
Jacek Gwizdka

**Measuring Biosignals of Overweight and Obese Children for Real-Time Feedback and Predicting Performance** . . . . . 185  
Nurten Öksüz, Russa Biswas, Iaroslav Shcherbatyi and Wolfgang Maass

**Stationarity of a User’s Pupil Size Signal as a Precondition of Pupillary-Based Mental Workload Evaluation** . . . . . 195  
Ricardo Buettner, Ingo F. Scheuermann, Christian Koot, Manfred Rössle and Ingo J. Timm

**Towards Reconceptualizing the Core of the IS Field from a Neurobiological Perspective** . . . . . 201  
Lars Taxén

**Using EEG Signal to Analyze IS Decision Making Cognitive Processes** . . . . . 211  
Nabila Salma, Bin Mai, Kamesh Namuduri, Rasel Mamun, Yassir Hashem, Hassan Takabi, Natalie Parde and Rodney Nielsen

# Tobias Kalenscher—Keynote

## **Why do we need animals to understand the neurobiology of economic decision-making?**

Despite the still frequently made assumption that humans are rational, consistent, sophisticated and selfish decision-makers, decades of research in the behavioral sciences suggests that individuals are often much less rational and egoistic as originally assumed. Yet, it is still elusive what causes these systematic deviations from the rational choice ideal. Interestingly, not only human decision-makers, but also nonhuman animals often act seemingly inconsistent with their revealed preferences, for example, when foraging for food. Humans and animals often make similar, maybe even identical decision “errors”. These intriguing parallels in human and animal choice patterns support the premise that they may share evolutionary roots. In my talk, I will argue in favor of the idea that the reality of decision-making with all its facets, including action against one’s own preferences, has to be understood in light of the nature, constraint and evolution of the neural apparatus supporting its function. I propose that the neural architecture of choice has evolved to its current state because it provided decision-makers with an adaptive advantage. This means that, even though there might exist a many-to-one mapping of neural implementations to choice processes, careful comparisons across species can complement human microeconomics research by supplying possible answers to the question why we make decisions as we do.

## **Gernot Müller-Putz—Hot Topic Talk**

### **The power of EEG: From single channel to high resolution derivations**

The talk briefly describes the neurophysiological foundation of EEG, recording methods, artifacts, type of electrodes and amplifiers. The main part will contain the discussion of using EEG depending on the number of derivations used and type of application as there are (i) single channel EEG and neurofeedback, (ii) medium number of channels to differentiate between brain states and (iii) high resolution EEG for functional brain imaging. A brief outlook to future applications of EEG will conclude this talk.

# The Psychophysiological Effect of a Vibro-Kinetic Movie Experience: The Case of the D-BOX Movie Seat

Horea Pauna, Pierre-Majorique Léger, Sylvain Sénécal,  
Marc Fredette, François Courtemanche, Shang-Lin Chen,  
Élise Labonté-Lemoyne and Jean-François Ménard

**Abstract** Watching a film in a movie theater can be an immersive experience, but to what extent does the experience differ when the moviegoer is using a vibro-kinetic seat, i.e., a seat providing motion and vibration feedback synchronized with the movie scenes? This paper seeks to measure the effect of a multi-sensory cinema experience from a psychophysiological standpoint. Using electroencephalography, galvanic skin response, heart rate, and facial micro-expression measures, this study compares the difference between two movie viewing experiences, i.e. one without movement and one with artistically enhanced vibro-kinetic feedback. Results of a within-subject experiment suggest that there are significant differences in psychophysiological states of users. Users exhibit more positive emotions, greater arousal, and more cognitive immersion in the vibro-kinetic condition. Therefore, multi-sensory stimulation, in the context of cinema, appears to produce an enhanced experience for spectators.

---

H. Pauna (✉) · P.-M. Léger · S. Sénécal · M. Fredette · F. Courtemanche · S.-L. Chen ·  
É. Labonté-Lemoyne  
HEC Montréal, Montréal, Canada  
e-mail: horea.pauna@hec.ca

P.-M. Léger  
e-mail: pml@hec.ca

S. Sénécal  
e-mail: sylvain.senecal@hec.ca

M. Fredette  
e-mail: marc.fredette@hec.ca

F. Courtemanche  
e-mail: francois.courtemanche@hec.ca

S.-L. Chen  
e-mail: shang-lin.chen@hec.ca

É. Labonté-Lemoyne  
e-mail: elise.labonte-lemoyne@hec.ca

J.-F. Ménard  
D-BOX Technologies Inc, Longueuil, Canada  
e-mail: jfmenard@d-box.com

**Keywords** Multi-sensory cinema · Vibro-kinetic · Multi-sensory Moviegoer experience · Immersion · Psychophysiological · Movie

## 1 Introduction

In order to maintain their market share, movie theaters have to invest in their customers' moviegoing experience. Technologies such as IMAX, 3D and Dolby Atmos are key ingredients in enhancing this experience by making it more immersive. More recently, vibro-kinetic movie seats have been used to provide moviegoers with an even more immersive experience. These seats are equipped with specialized hardware and software that store, manage, and transmit motion codes to the movie seat; these motion codes are synchronized with the movie scenes. For instance, the seat could generate a trembling motion during an earthquake movie scene or the feeling of weightlessness during a zero gravity movie scene. There have been several studies on the effect of many technologies on immersion and the general cinematic experience [1–5]. However, to our knowledge, no research has yet investigated the effect of whole-body vibro-kinetic or motion feedback technologies from the standpoint of the moviegoer's psychophysiological reactions. Many of these immersive technologies aim to provide an enhanced multi-sensory experience to the moviegoer [1]. However, the lack of data-driven research fails to utilize implicit measures of the user experience [6] to assess multi-sensorial reactions. We suggest that this is a major gap since movie theaters invest large sums in these technologies without having a clear understanding of their effect on an ecologically valid experience. Do these technologies really have an effect on the moviegoer's movie experience?

In order to answer this question, this study investigates the extent to which the experience differs when a moviegoer uses a vibro-kinetic movie seat rather than a traditional non-moving seat. Results of our within-subject experiment suggest significant differences between a traditional and a vibro-kinetic movie experience in terms of the viewer's emotional and cognitive reactions.

## 2 Prior Research

Since the 1950s, filmmakers have been experimenting with different techniques to bridge the gap between the spectator's reality and the reality of the movie to thereby increase immersion [4, 7]. Immersion has been described by Carù and Cova [8] as the capacity of an individual to eliminate the distance between the self and the experience. This concept is specified by Recuber [4] as being dependent on spatial features that enhance the absorbing efforts of a camera's perspective.

Multi-sensorial experiences are considered a driver in delivering a new form of cinematic immersion and have the potential to enhance the spectator's overall

experience [1]. Thus, the movie industry has been trying to enhance immersion on the basis of multi-sensorial experiences [1]. The addition of so called “dimensions” [2] aims to trigger more than just the auditory and visual senses to produce an enhanced experience. The cinema industry has experienced with aromatic output as an additional “dimension”, meant to trigger the olfactory senses and to increase viewer immersion [1, 7, 9]. One of the most popular examples of added dimensions is stereoscopic imagery (3D), increasing the realism of a two-dimensional screen by adding depth and spatial immersion [2].

We also find the addition of cinema seats mimicking the movement of the projected movie [2]. However, the addition of realistic tactile sensations to enhance the movie watching experience has been sparsely researched. For instance, [9] tested the effects of wind, vibrations emitted by a “wrist rumbler”, and light effects on self-reported sequence quality rating. They found that multi-sensory experiences were rated higher by participants. Thus, there is a research interest and development potential as to the effect of multi-sensory technology on the movie viewer’s actual emotional and cognitive reactions.

### 3 Experimental Design and Sample

An experiment was conducted with 43 participants (22 males, 21 females). Participants were screened for neuropsychological diagnostics and other physical conditions such as the need to wear glasses to watch a movie. The study was approved by the University’s ethics board.

We performed a  $3 \times 2$  within-subject experimental design. The first manipulated factor was the movie. The first 8 min and 26 s of *The Martian* (2015), the first 11 min and 43 s of *Skyfall* (2012) and 5 min and 41 s of a Formula 1 racing scene from the movie *Rush* (2013) were used. These sequences were chosen for the following reasons: (1) they are action-oriented, (2) they are self-contained stories (with a beginning and an end), and (3) they are short enough to accommodate the research design (less than 15 min).

The second manipulated factor was the movement (experiencing no movement during the sequence or having a vibro-kinetic experience). To manipulate this factor, a D-BOX (Longueuil, Canada) motion-enabled recliner chair was used. This particular seat is artistically enhanced as its movements are manually designed by specialized movement artists. The D-BOX seat has a vibro-kinetic spectrum ranging from 0 to 100 Hz and was calibrated to synchronize the motion with the audio signal within 10 ms at the fixed viewing distance. Given the available time, participants were randomly assigned to view only two of the three movies. One sequence had the D-BOX seat activated and the other had it disabled (no movement condition). All participants viewed both sequences seated on the same D-BOX seat and the movie/movement pairs were randomized.

## 4 Measures

Three types of psychophysiological variables were used to assess the participants' emotional and cognitive reactions: emotional valence, arousal, and cognitive states. Methodologies and guidelines have been presented by Riedl and Léger [6] to measure activity in the central nervous system and the peripheral nervous system in the context of Information Systems. We have used their recommendations concerning tools used to measure both nervous systems.

The participants' arousal level during movie sequences was measured using electrodermal activity (EDA). Though EDA has been widely used as an indicator of arousal [10–14], it cannot, on its own, determine if the activation is positive or negative when presented with audio-visual stimulus intended to trigger both spectrums of valence [11, 12, 14]. Emotional valence differentiates positive and negative emotions and can be detected using facial micro-expressions [10]. The participant's facial expressions were therefore recorded during the experience. Finally, cognitive data was collected using electroencephalography (EEG). The EEG signal was recorded using 32 electrodes with a sampling rate of 1000 Hz and analyzed with EEGLab (San Diego, USA) and Brainvision (Morrisville, USA). It was filtered with IIR filters with a low cut-off at 1 Hz and a high cut-off at 40 Hz, then cleaned using continuous ASR in Matlab (Natick, USA) and re-referenced to the common average reference. EEG frequency activity was extracted for three bands for the sum of the *Cz*, *Pz*, *P3* and *P4* electrodes: alpha (8–13 Hz), beta (13–22 Hz), and theta (4–8 Hz). This method has been used by Pope et al. [15] to calculate engagement by dividing beta by the sum of alpha and theta. Data was synchronized using techniques previously described in [16, 17].

Building on this method and previous research [18, 19], we have extracted the dimensions of the engagement index and have used [20]'s classification for interpreting results. Alpha rhythms are associated with quieted states, beta rhythms with focused, active states, while theta rhythms are associated with quiet focus states such as meditation [20].

For each type of variable, the following features were calculated: mean, 10th and 90th percentiles as presented in Table 1. Since these measures are recorded continuously, a representative scalar value for each measure is needed to describe the subject's psychophysiological state in each movie period. To integrate and analyze multiple psychophysiological data for a subject in a given condition, the mean per movie per subject was used. The 10th and 90th percentiles were also considered for the following reasons: during a movie, the seat can be either moving or still. If a measure tends to have higher values while the seat is moving, then its 90th percentile may catch the effect of movement better than mean does. In the opposite case, if a measure tends to have lower values in conditions with seat movements, then its 10th percentile may be more sensitive to the movement.

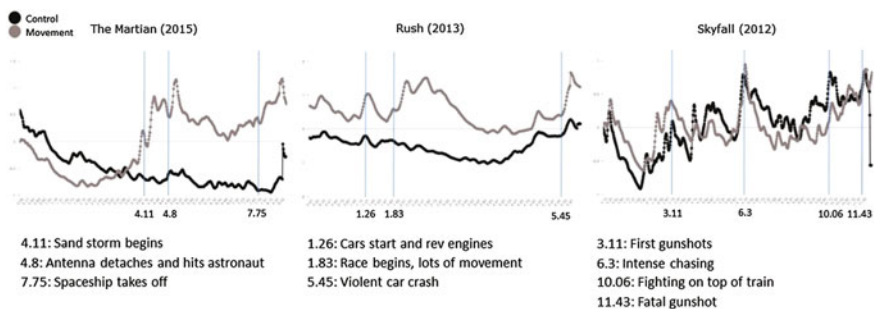


**Table 1** Operationalization of the variables

Types of measure	Instruments	Variables	Features
Emotional states	Recorded with MediaRecorder 2 (Wageningen, Netherlands) Analyzed with: Facereader (Wageningen, Netherlands)	Valence neutral happy sad Angry surprised scared disgusted	P10: 10th percentile for a subject in a given condition
Emotional arousal	Biopac (Goleta, USA)	Normalized eda	P90: 90th percentile for a subject in a given condition
Cognitive states	Brainvision (Morrisville, USA)	Alpha Beta Theta	Mean: mean of value for a subject in a given condition

## 5 Analysis and Results

T-tests were performed using a mixed linear approach for analyzing the continuous dependent variables which naturally contained repeated measures for each participant. Due to space constraints, we shall only report significant results. Results suggest that the D-BOX vibro-kinetic seat experience magnifies the movie experience. First, normalized electrodermal activity is significantly higher with the D-BOX seat activated on all three variations of the measure (10th and 90th percentiles and mean). Mean value for eda\_mean with D-BOX activated is 8.61 versus 7.57 for the control condition ( $p$ -value is 0.01). Mean value for eda\_p90 with D-BOX activated is 10.06 versus 9.00 for the control condition ( $p$ -value is 0.01). Mean value for eda\_p10 with D-BOX activated is 5.30 versus 4.61 for the control condition ( $p$ -value is 0.02). Figure 1 graphically illustrates the physiological difference between the conditions. To produce this figure, each raw EDA signal point is converted into its  $z$ -score.



**Fig. 1** EDA over time for all three movies and important events minutes

We then calculated the mean of all participants' z-score for a given time and applied a 5-s moving average window as to take into account the delayed physiological response to a certain stimulus. The sequence of points represent EDA variations through time. Major dramatic events are highlighted on both graphs. We see that the effect is especially important for *Rush* and *The Martian*, where the movie events produce significant movement.

Second, the results show that the subjects appear to experience more positive emotions. With the movement activated, the spectator's positive facial emotions (happy mean) are significantly amplified. Mean value for happy\_mean with D-BOX activated is 0.06 versus 0.03 for the control condition ( $p$ -value is 0.09). Negative facial emotions (angry and scared, 10th percentile) are significantly less present. Mean value for angry\_p10 with D-BOX activated is 0.003 versus 0.01 for the control condition ( $p$ -value is 0.03). Mean value for scared\_p10 with D-BOX activated is 1.21E-06 versus 4.85E-06 for the control condition ( $p$ -value is 0.05).

Finally, preliminary EEG results show a more relaxed cognitive state in the vibro-kinetic condition. Specifically, the moving condition generates less beta activity. It appears that the moving condition triggers less cognitive activity, which would be compatible with a more immersive cinematic experience. Mean value for beta\_p90 with D-BOX activated is 6.34 versus 6.70 for the control condition ( $p$ -value is 0.07).

## 6 Concluding Comments

A theoretical contribution is made by filling the literature gap on the effects of a vibro-kinetic cinema seat on a spectator's psychophysiological response. The motion-enabled recliner chair produces an enhanced cinema experience by rendering an artistically designed vibro-kinetic stimulation in sync with the movie's events.

We have demonstrated that there is a clear difference between a traditional cinema viewing experience and one with this movement enhancing seat. The results thus show that motion-enabled seats produce a heightened experience for spectators. Additional research will be conducted to better interpret EEG data. Further research will also be conducted to determine if specific movements on an XYZ axis produce a certain psychophysiological response when in sync with a movie scene. This would allow the targeting of specific emotional responses which a filmmaker could choose to enhance using a vibro-kinetic movement seat.

## References

1. Galloso, I., Palacios, J.F., Feijoo, C., Santamaria, A.: On the influence of individual characteristics and personality traits on the user experience with multi-sensorial media: an experimental insight. *Multimedia Tools Appl.* 1–44 (2016)
2. Jackman, A.H.: 3-D cinema: immersive media technology. *GeoJournal* **80**(6), 853–866 (2015)
3. McMahan, A.: Immersion, engagement and presence. In: *The Video Game, Theory Reader*, vol. 67 (2003)
4. Recuber, T.: Immersion cinema: the rationalization and reenchantment of cinematic space. *Space Cult.* **10**(3), 315–330 (2007)
5. Visch, V.T., Tan, E.S., Molenaar, D.: The emotional and cognitive effect of immersion in film viewing. *Cogn. Emot.* **24**(8), 1439–1445 (2010)
6. Riedl, R., Léger, P.M.: *Fundamentals of NeuroIS*. Springer, Berlin (2016)
7. Kaye, J.J.: Making Scents: aromatic output for HCI. *Interactions* **11**(1), 48–61 (2004)
8. Carù, A., Cova, B.: How to facilitate immersion in a consumption experience: appropriation operations and service elements. *J. Consum. Behav.* **5**(1), 4–14 (2006)
9. Waltl, M., Timmerer, C., Hellwagner, H.: Improving the quality of multimedia experience through sensory effects. In: *2010 Second International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE (2010)
10. Abadi, M.K., Staiano, J., Cappelletti, A., Zancanaro, M., Sebe, N.: Multimodal engagement classification for affective cinema. In: *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE. (2013)
11. Bos, M.G.N., Jentgens, P., Beckers, T., Kindt, M.: Psychophysiological response patterns to affective film stimuli. *PLoS ONE* **8**(4), e62661 (2013)
12. Fleureau, J., Guillotel, P., Orlac, I.: Affective benchmarking of movies based on the physiological responses of a real audience. In: *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE (2013)
13. Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O.: Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* **30**, 261–273 (1993)
14. Rothwell, S., Lehane, B., Chan, C.H., Smeaton, A.F., O'Connor, N.E., Jones, G.J.F., Diamond D.: The CDVplex biometric cinema: Sensing physiological responses to emotional stimuli in film. In: *The 4th International Conference on Pervasive Computing* (2006)
15. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* **40**(1), 187–195 (1995)
16. Charland, P., Léger, P.M., Sénécal, S., Courtemanche, F., Mercier, J., Skelling, Y., Labonté-Lemoyne, E.: Assessing the multiple dimensions of engagement to characterize learning: a neurophysiological perspective. In *JoVE J. Visualized Exp.* **101**, (2015)
17. Léger, P.M., Sénécal, S., Courtemanche, F., Ortiz de Guinea, A., Titah, R., Fredette, M., Labonté-LeMoyne, E.: Precision is in the eye of the beholder: Application of eye fixation-related potentials to information systems research. *J. Assoc. Inf. Syst.* **15**(10), 651 (2014)
18. Léger, P.M., Davis, F.D., Cronan, T.P., Perret, J.: Neurophysiological correlates of cognitive absorption in an enactive training context. *Comput. Hum. Behav.* **34**, 273–283 (2014)
19. de OrtizGuinea, A., Titah, R., Léger, P.M.: Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research. *Comput. Hum. Behav.* **29**(3), 833–844 (2013)
20. Baars, B.J., Gage, N.M.: *Fundamentals of cognitive neuroscience: A beginner's guide*. Academic Press (2012)

# Reinforcement Sensitivity and Engagement in Proactive Recommendations: Experimental Evidence

Laurens Rook, Adem Sabic and Markus Zanker

**Abstract** We drew on revised Reinforcement Sensitivity Theory to claim that users with an anxiety-related behavioral inhibition would experience proactively delivered recommendations as potential threats. Such users would display higher user engagement especially when they were interrupted by inaccurate (vs. accurate) recommendations, because they ruminate about them. This prediction was tested and confirmed in a controlled experiment that exposed participants to proactive recommendations on their smartphone. Results highlight the need to gain more knowledge on the neural correlates of anxiety, and to apply such insights to human–computer interaction design for recommender systems.

**Keywords** Behavioral inhibition • Fight-flight-freeze system • Recommendation delivery • Proactivity • Human–computer interaction

## 1 Introduction

Recommender systems (RS) are automated decision support tools designed to provide custom-made advice on items to facilitate people’s navigation in large information spaces [1]. RS provide personalized suggestions based on presumed preferences and needs of a user and other people’s behavior. They help people overcome information overload either by providing accurate recommendations on

---

L. Rook (✉)

Delft University of Technology, Delft, The Netherlands  
e-mail: l.rook@tudelft.nl

A. Sabic

Alpen-Adria-University Klagenfurt, Klagenfurt, Austria  
e-mail: adem.sabic@gmail.com

M. Zanker

Free University of Bozen-Bolzano, Bolzano, Italy  
e-mail: Markus.Zanker@unibz.it

request or by delivering them proactively. In addition to making the most accurate predictions to a user, also alternative measures for recommendation quality such as novelty, diversity, and unexpectedness of recommended items are increasingly explored [2]. In the present study, we take a NeuroIS approach to RS [3] by drawing on revised Reinforcement Sensitivity Theory (RST), a biopsychological theory from cognitive neuroscience [4], to posit that users with sensitivity towards behavioral inhibition will experience unexpected proactive recommendations as potential threats, try to cope with them via error-related rumination, and that this will lead them to display higher user engagement under the right circumstances.

## 2 Theory

### 2.1 *Reinforcement Sensitivity Theory and Prediction*

Originally, RST [5] explained personality as grounded in a general behavioral inhibition system (a brain system related to anxiety triggered by novel stimuli) and a behavioral approach system (a brain system triggered by reward and non-punishment). Revised RST [4] split the behavioral inhibition system into primary anxiety (BIS-Anxiety) when someone is confronted with conflicting novel stimuli, and secondary anxiety initiated by fight-flight-freeze responses to fear (FFFS-Fear). Neurologically, such aversive behaviors are located in the hippocampus, partially mediated by the prefrontal and right inferior frontal anterior cingulate cortex, and right inferior frontal gyrus [6, 7]. Interestingly, in laboratory settings, people prone to BIS-Anxiety have been found to be more sensitive to goal-conflict (e.g., novel information not making sense from an appetitive-aversive point of view). They easily detect such errors, and usually invest cognitive effort in correcting them—turning into anxious ruminators whenever necessary. Consequently, people with BIS-Anxiety typically perform well in intellectual tasks. They tend to excel in educational setting, and seem to flourish in intellectually demanding jobs—especially, when those people are above average in intelligence, and hold desk-based (vs. hazardous) positions in industry [8]. It should be emphasized that—however sparse the applications of RST on industrial and organizational psychology—especially the resolution of goal-conflict of people with BIS-Anxiety is regarded beneficial to workplace behavior, as it may facilitate problem solving at work [9].

Combining RS and revised RST literatures, we claim that proactive recommendations can be understood as novel and unexpected—but potentially threatening—stimuli, which users receive when browsing for information on their computer devices. RS may facilitate search activity in tune with individual preferences [1, 2] for people not qualified by primary anxiety, but trigger goal conflict in users high on BIS-Anxiety—especially when RS appear inaccurate, leading such users to anxiously ruminate to trace and correct the error—which, eventually, leads

to higher user engagement. This novel prediction, inspired by prior work on neural correlates of technology acceptance [10], was put to the test in an experimentally controlled user study.

### 3 Method

#### 3.1 *Participants and Design*

Participants from a Dutch university enrolled in an applied statistics course were randomly assigned to the experimental conditions of a Recommendation Accuracy (high, low) factorial design on User Engagement, to which BIS-Anxiety and FFFS-Fear scores were added as covariates. An initial sample of 156 participants (87 men and 69 women;  $M$  age = 21.17 years,  $SD$  = 1.49) completed the study, but excluded from analyses were the data of those ( $N$  = 25; 16.03%) who reportedly did not receive proactive recommendations while working on the experiment (see below). This resulted in a final sample of 131 participants (71 men and 60 women;  $M$  age = 21.21 years,  $SD$  = 1.49), which was used for the analyses reported below.

#### 3.2 *Materials and Procedure*

Per email, participants were invited to take part in a study on responsible e-tourism. They first completed an online pre-survey assessing the BIS/BAS scales [11] (see below) and demographics. Next, they received instructions to download a smartphone application from the Google Play or Apple App Store, depending on their mobile operating system. In the application environment, participants were asked to find a way to travel from Delft, the Netherlands, to London, UK on a specified date to attend a major event and to find a place to stay during the event. Precisely 60 s after their first exploration of these tourism-related challenges, they began to receive proactive recommendations, depending on experimental condition (see below). On completion, participants received the link to an online post-survey assessing manipulation checks and user experiences with the application. Participants were debriefed in class as to the purposes of the study.

#### 3.3 *Measures*

**Manipulation of Recommendation Accuracy.** To induce the recommendation accuracy manipulation, participants received five proactive recommendations that were either highly or marginally relevant. Participants in the high recommendation

accuracy condition received RS directly useful for solving the issues of mode and means of travel and accommodation. Participants in the low recommendation accuracy condition received RS relevant for a stay in London, but too generic to solve the posed trip planning challenges.

**Behavioral Inhibition System.** The original BIS/BAS scales [11] and their Dutch translation [12] were used to measure individual differences in BIS-Anxiety and FFFS-Fear (following instructions by [13]):

*BIS-Anxiety.* Four items measured BIS-Anxiety: *Criticism or scolding hurts me quite a bit, I feel pretty worried or upset when I think or know somebody is angry at me, I feel worried when I think I have done poorly at something important, and I worry about making mistakes*, on a 4-point scale anchored at 1 (*very true for me*) and 4 (*very false for me*; Cronbach  $\alpha = 0.71$ ).

*FFFS-Fear.* Three items measured FFFS-Fear: *Even if something bad is about to happen to me I rarely experience fear or nervousness* (reverse coded), *If I think something unpleasant is going to happen I usually get pretty worked up*, and *I have very few fears compared to my friends* (reverse coded), on a 4-point scale anchored at 1 (*very true for me*) and 4 (*very false for me*; Cronbach  $\alpha = 0.57$ ), which occurs more often [14], and is accredited to the small number of items in the scale.

**User Engagement.** Engaged users do not submit minimal results, but invest reasonable energy in solving challenging tasks properly [15], also in computer-mediated settings [16]. Consistent with this, we operationalized user engagement in the present study as the total number of characters that a participant submitted as solution for the series of travel tasks they had worked on.

*Manipulation Check.* Four items were used to assess whether participants had experienced the accuracy of the recommendations they received in line with experimental conditions. Example items were: *The travel solutions I produced for the e-Tourism Challenge were of good quality* and *The recommended set of links for the e-Tourism Challenge enabled me to submit high-quality travel solutions* on a 7-point scale anchored at 1 (*not at all true for me*) and 7 (*very true for me*; Cronbach  $\alpha = 0.83$ ).

## 4 Results

### 4.1 Manipulation Check

A linear regression analysis on the four items of the manipulation check for recommendation accuracy showed a significant main effect of low versus high recommendation accuracy,  $\beta = 0.48$ ,  $t(130) = 2.14$ ,  $p < 0.04$ , which indicated that participants indeed had experienced the quality of the proactive recommendations they received in line with the experimental condition they had been randomly assigned to. This confirmed that the manipulation had been effective.

### 4.2 Correlations

Table 1 shows the descriptive statistics and the bivariate correlations. Recommendation accuracy was not in itself significantly correlated with BIS-Anxiety and FFFS-Fear or with User Engagement. Consistent with prior studies, BIS-Anxiety and FFFS-Fear were significantly and positively correlated [13, 14, 17]. FFFS-Fear was significantly and negatively correlated with User Engagement.

### 4.3 User Engagement

We conducted a series of negative binomial regression analyses to explore the main effects of Recommendation Accuracy, BIS-Anxiety, FFFS-Fear, as well as their interactions, on User Engagement. The likelihood ratio for the full negative binomial model was  $\chi^2(7) = 22.02, p < 0.01$ , which showed that our model was significant. Negative binomial regression analysis revealed a significant interaction effect of Recommendation Accuracy and BIS-Anxiety on User Engagement; people high on BIS-Anxiety displayed more user engagement in the face of inaccurate (vs. accurate) recommendations. Further test of the simple interactions yielded a significant main effect of BIS-Anxiety on User Engagement when Recommendation Accuracy was low, Wald  $\chi^2(1) = 1.00, p < 0.05$ , but not when Recommendation Accuracy was high, Wald  $\chi^2(1) = 1.00, ns$  (see Fig. 1). Probing of a significant three-way interaction indicated that this interaction effect between Recommendation Accuracy and BIS-Anxiety existed only for people low on FFFS-Fear, Wald  $\chi^2(1) = 1.00, p < 0.0001$ , but not high on FFFS-Fear, Wald  $\chi^2(1) = 0.00, p = 0.93$ .

**Table 1** The descriptive statistics and the bivariate correlations

	<i>M</i>	<i>SD</i>	1	2	3	4
1. Recommendation accuracy	0.52	0.5	–			
2. BIS-anxiety	11.45	2.31	0.08	–		
3. FFFS	8.08	1.69	0.08	0.36**	–	
4. User engagement	139.60	125.68	–0.09	–0.14	–0.17*	–

Note *N* = 131; \**p* < 0.05 level, \*\**p* < 0.01 level; two-tailed



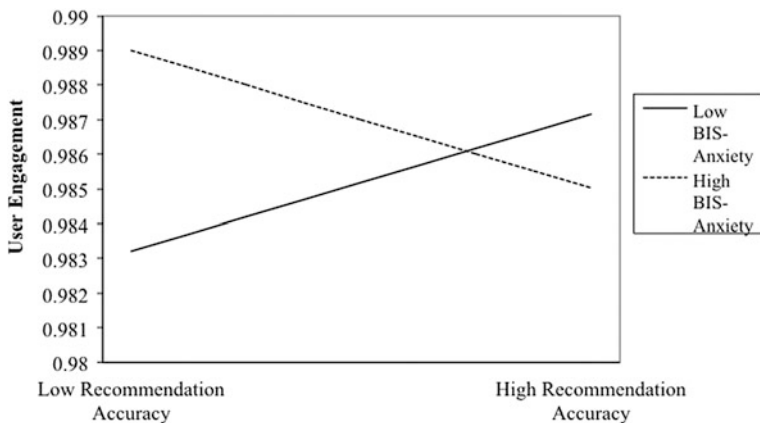


Fig. 1 Recommendation accuracy  $\times$  BIS-anxiety on user engagement

## 5 Discussion and Conclusion

The present study reported first evidence that people high on BIS-Anxiety display higher levels of user engagement when exposed to inaccurate recommendations, because they, apparently, put more cognitive effort in finding and restoring inaccuracy errors. This confirms extant theorizing in RS that accuracy alone is not enough [18], and makes clear that consideration of alternative quality measures like novelty, diversity, and unexpectedness, indeed, holds major potential for yielding heightened user engagement [2]. Importantly, though, the key contribution of the present study lies in its recognition, derived from the neuroscience research framework of RST, that proactivity must be understood as trigger for constructive behavioral inhibition—that is, when RS are designed such that they refrain from invoking fear. Given the documented association between behavioral inhibition and the stress hormone cortisol—also in disruptive human–computer interactions [19]—it makes perfect sense for future study in this direction to propose a NeuroIS [20] research agenda aimed at explicitly establishing this cortisol-inhibition linkage also for proactive RS. In addition, traditional electroencephalography (EEG) procedures for testing cortisol-inhibition linkages [7] could in future work be adapted to our experimental paradigm to explore when proactivity turns into a stressor invoking fear rather than anxiety. For anxiety-prone users, this would make the difference between vigilant versus unproductive, unhealthy or no user engagement, whatsoever. In conclusion, the present study, therefore, shows the viability of adopting a human–computer interaction perspective on RS in general, and on taking a neurobiological perspective to the study of proactive recommender systems in particular.

## References

1. Ricci, F., Rokach, L., Shapira, B.: *Recommender Systems Handbook*. Springer, New York (2015)
2. Jannach, D., Resnick, P., Tuzhilin, A., Zanker, M.: Recommender systems: beyond matrix completion. *Comm. ACM*. **59**, 94–102 (2016)
3. Dimoka, A., Pavlou, P.A., Davis, F.: NeuroIS: the potential of cognitive neuroscience for information systems research. *Inform. Syst. Res.* **22**, 687–702 (2011)
4. Gray, J.A., McNaughton, N.: *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford University Press, Oxford (2000)
5. Gray, J.A.: *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-Hippocampal System*. Oxford University Press, Oxford (1982)
6. Corr, P.J.: Anxiety: splitting the phenomenological atom. *Pers. Individ. Differ.* **50**, 889–897 (2010)
7. Tops, M., Boksem, M.A.S.: Cortisol involvement in mechanisms of behavioral inhibition. *Psychophysiology* **48**, 723–732 (2011)
8. Perkins, A.M., Corr, P.J.: Anxiety as an adaptive emotion. In: Gerrod Parrott, W. (ed.) *The Positive Side of Negative Emotions*, pp. 37–54. The Guilford Press, New York (2014)
9. Corr, P.J., McNaughton, N., Wilson, M.R., Hutchison, A., Burch, G., Poropat, A.: Neuroscience of motivation and organizational behavior: putting the reinforcement sensitivity theory (RST) to work. In: Kim, S., Reeve, J.M., Bong, M. (eds.) *Recent Developments in Neuroscience Research on Human Motivation: Advances in Motivation and Achievement*, vol. 19, pp. 65–92. Emerald Group Publishing, Bingley (2017)
10. Dimoka, A.: What does the brain tell us about trust and distrust? evidence from a functional neuroimaging study. *MIS Q.* **34**, 373–396 (2010)
11. Carver, C.S., White, T.L.: Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J. Pers. Soc. Psychol.* **67**, 319–333 (1994)
12. Franken, I.H.A., Muris, P., Rassin, E.: Psychometric properties of the Dutch BIS/BAS scales. *J. Psychopathol. Behav.* **27**, 25–30 (2005)
13. Heym, N., Ferguson, E., Lawrence, C.: An evaluation of the relationship between gray's revised RST and Eysenck's PEN: distinguishing BIS and FFFS in Carver and White's BIS/BAS scales. *Pers. Individ. Differ.* **45**, 709–715 (2010)
14. Keiser, H.N., Ross, S.R.: Carver and Whites' BIS/FFFS/BAS scales and domains and facets of the five factor model of personality. *Pers. Individ. Differ.* **51**, 39–44 (2011)
15. Rich, B.L., LePine, J.A., Crawford, E.R.: Job engagement: antecedents and effects on job performance. *Acad. Manage. J.* **31**, 599–627 (2010)
16. Ray, S., Kim, S.S., Morris, J.G.: The central role of engagement in online communities. *Inform. Syst. Res.* **25**, 528–546 (2014)
17. Walker, B.R., Jackson, C.J.: How the five factor model and revised reinforcement sensitivity theory predict divergent thinking. *Pers. Individ. Differ.* **57**, 54–58 (2014)
18. McNee, S.M., Riedl, J., Konstan, J.A.: Being accurate is not enough: how accuracy metrics have hurt recommender systems. In: *Proceedings of the CHI'06 Conference on Human Factors in Computing Systems*, pp. 1097–1101. ACM, New York (2006)
19. Riedl, R., Kindermann, H., Auinger, A., Javor, A.: Technostress from a neurobiological perspective: system breakdown increases the stress hormone cortisol in computer users. *Bus. Inf. Syst. Eng.* **2**, 61–69 (2012)
20. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P.H., Pavlou, P.A., Müller-Putz, G., Riedl, R., vom Brocke, J., Weber, B.: On the use of neurophysiological tools in IS research: developing a research agenda for NeuroIS. *MIS Q.* **36**, 679–702 (2012)

# The Choice Is Yours: The Role of Cognitive Processes for IT-Supported Idea Selection

Isabella Seeber, Barbara Weber, Ronald Maier  
and Gert-Jan de Vreede

**Abstract** The selection of good ideas out of hundreds or even thousands has proven to be the next big challenge for organizations that conduct open idea contests for innovation. Cognitive load and attention loss hinder crowds to effectively run their idea selection process. Facilitation techniques for the reduction and clarification of ideas could help with such problems, but have not yet been researched in crowd settings that are prevalent in idea contests. This research-in-progress paper aims to contribute to this research gap by investigating IT-supported selection techniques that differ in terms of selection direction and selection type. A laboratory experiment using eye-tracking will investigate variations in selection type and selection direction. Moreover, the experiment will test the effects on the decision-making process and the number and quality of ideas in a filtered set. Findings will provide explanations why certain mechanisms work for idea selection. Potential implications for theory and practice are discussed.

**Keywords** Idea contest · Idea quality · Idea selection · Open innovation  
Screening rules

---

I. Seeber (✉) · R. Maier  
Department of Information Systems, Production and Logistics Management,  
University of Innsbruck, Innsbruck, Austria  
e-mail: isabella.seeber@uibk.ac.at

B. Weber  
Department of Applied Mathematics and Computer Science,  
Technical University of Denmark, Copenhagen, Denmark  
e-mail: bweb@dtu.dk

G.-J. de Vreede  
Information Systems and Decision Sciences Department, University of South Florida,  
Florida, USA  
e-mail: gdevreede@usf.edu

## 1 Introduction

An innovation contest is a “(web-based) competition of innovators who use their skills, experience and creativity to provide a solution for a particular contest challenge defined by an organizer” [1]. IBM’s Innovation Jam [2], Dell’s Ideastorm [3] or Cisco’s iPrize [4] are examples where organizations successfully tapped into the creative power of an external crowd to foster their firm’s innovativeness [5]. During an idea contest, individuals generate hundreds or even thousands of ideas [2]. While more ideas increase the likelihood of more good ideas [6], contest organizers face challenges when it comes to the selection of the best ideas. It has been established that individuals often fail in discerning the best ideas [7], mostly for reasons related to the overwhelming amount of information that needs to be processed [8]. Moreover, research shows that idea selection processes differ from platform to platform or from organization to organization [2, 3, 9]. Patterns of successful idea selection processes have yet to be discovered. It has been suggested that one way to address the idea selection challenge is to employ a crowd during the initial screening of ideas using specific assessment criteria [10]. To outsource idea selection to the crowd, appropriate IT support is needed where crowd-based selection mechanisms guide individual members through idea selection. Thus far, it is, however, unclear how to best design such selection mechanisms so that they enable the crowd to select the best ideas.

This paper contributes to a better understanding of crowd-based idea selection mechanisms. We conceptualize two screening rules (conjunctive and disjunctive) to prompt an individual member of the crowd to select good ideas or to drop bad ideas. We refer to this feature as the *selection direction* mechanism. In addition, we conceptualize the *selection type* mechanism to guide an individual to either make a single or multiple choice selection. We test our exploratory theory using eye tracking to investigate the underlying decision-making processes of individuals during idea selection.

## 2 Background and Hypotheses Development

Idea selection in idea contests can be conceptualized as multi-attribute decision-making. In such a decision-making process, an individual is confronted with multiple, sometimes conflicting criteria [11]. For example, many idea contests specify a list of evaluation criteria on their website that a jury (of experts) will use to determine the winning ideas [9]. Common evaluation criteria relate to idea quality dimensions such as novelty (the idea is original because nobody has expressed it before), workability (the idea can be implemented), relevance (the idea has a purpose and satisfies the problem seeker’s goals), and elaborateness (the idea is thoroughly worked out in detail) [6]. The preference of the crowd or an idea’s popularity is an additional attribute that is often considered during idea selection

[10]. Ratings, votes, and comments are therefore additional points of consideration during idea selection. Hence, individuals that are specifically tasked to select the good or drop bad ideas need to consider multiple attributes. Processing multiple and potentially conflicting attributes demands information elaboration [12] with the risk to encounter cognitive overload [13]. According to the elaboration likelihood model (ELM) [14], information needs to catch one's attention and even if information is attended, not all information processing takes the central route in the brain leading to high elaboration of information. It may also take the peripheral route resulting in low elaboration. This paper builds on ELM and investigates IT-supported selection techniques and the way they influence information processing.

## ***2.1 Selection Direction and Screening Rules***

Extreme value logic assumes that “a group can discern good ideas from bad ideas” [7]. This implies that individuals are good in determining the best and worst ideas, but have problems with ideas in the middle. Recent research shows that idea selection may work best when individuals need to select the worst ideas, that is lemons, instead of the best ideas, that is stars [15]. In the information-processing literature, this behavior is discussed under the term screening rules [16]. Screening rules are decision heuristics that explain how individuals choose from multi-attribute products [17]. Conjunctive and disjunctive decision rules represent two non-compensatory screening rules. The conjunctive decision rule adopts an elimination-by-aspects approach and describes that an individual would only choose an alternative (e.g., an idea) if all relevant attributes have acceptable levels. Thus, an individual would only select an idea that meets the minimum threshold for each quality criterion and has received acceptable crowd evaluations. In contrast, the disjunctive decision-rule mandates that an idea must achieve acceptable levels for at least one of the attributes. When returning to extreme value logic, screening rules might help to explain how individuals focus on attributes when selecting good ideas (stars) versus bad ideas (lemons). In the case of stars, it is assumed that individuals adopt the conjunctive decision rule, that is, an idea is only considered as a very good idea if all quality criteria and crowd evaluation criteria meet chosen thresholds. In the case of lemons, it is assumed that individuals adopt the disjunctive decision rule. Hence, they would drop an idea from further consideration as soon as one attribute performs poorly or below a certain threshold.

In a setup where each relevant idea attribute is visualized at a unique area of a screen, eye tracking [18] could help to explore whether the star approach is indeed related to the conjunctive decision rule and the lemon approach to the disjunctive decision rule. Individuals should display different eye tracking patterns depending on the selection direction (i.e., stars vs. lemons). Specifically, we should observe that individuals consider all attributes in the star approach and consider attributes only until a poor attribute value has been found in the lemon approach. Thus, we

should have fewer attributes attended in the lemon approach and as a result fewer eye fixations [18]:

H1: Individuals that eliminate bad ideas (lemons) will attend fewer attributes (operationalized as fixation count) than individuals that select good ideas (stars).

If the logic related to H1 holds, individuals that select bad ideas should experience less cognitive load and should have more cognitive resources available to vigilantly attend the idea description. Assuming that individuals deem the task of finding the best ideas sufficiently important to allocate these cognitive resources available, this should allow them to decide more accurately on the inherent quality of the idea. Consequently, we expect individuals that adopt the disjunctive decision-rule fostered by prompting them to select bad ideas, to have more ideas of high quality in the final list referred to as the set of filtered ideas. Therefore,

H2: Individuals that eliminate bad ideas (lemons) will achieve higher idea quality in the set of filtered ideas than individuals that select good ideas (stars).

## ***2.2 Selection Type and Multiple Choice***

Research on IT-supported facilitation shows that attention loss and cognitive load are key problems during idea filtering [19]. One useful practice to address this challenge is to lower the number of ideas in the set of ideas to consider [20]. As a consequence, individuals need to process fewer options (ideas) and attributes (idea characteristics) which thus lowers the likelihood of cognitive load and decreases choice quality [13]. Particularly in early stages of idea filtering, individuals tend to minimize cognitive effort by adopting attribute-based processing of information [21]. This means that an individual can keep cognitive load lower when prompted to select or drop just a single idea. If multiple ideas (stars or lemons) have to be selected, this likely forces an individual to switch to alternative-based processing [21]. Such processing not only requires to keep track of (un)favorable attributes of a single idea, but of multiple ideas that are above or below an acceptable threshold [22]. This presumably induces higher cognitive load as the increased number of choices will also increase the demand on working memory [22]. While the intrinsic cognitive load (i.e. the complexity inherent to the task) is the same for both selection types, we expect differences in the extrinsic load that stem from differences in the instructions (i.e. prompting subjects to choose one versus multiple ideas) [23, 24]. Thus, we hypothesize:

H3: Individuals that must make a single choice from a small set of ideas to consider will have lower cognitive load (operationalized as pupil dilation and heart rate variability) when compared to individuals that must make multiple choices from the same idea set.

Regardless of selection direction (i.e. stars or lemons) we assume that individuals decide by comparing ideas and by integrating attended information. We expect that the selection of a single idea requires information integration to a smaller extent compared to the selection of multiple ideas, because a single choice allows the individual to abort the selection as soon as a lemon or star is found. In the case of multiple choices, an individual needs to decide for each of the ideas in the set of ideas to consider. We operationalize the information integration by treating each idea as a distinct Area of Interest (AOI) and by counting the run count for each AOI (i.e. the number of times the eyes of a participant left an AOI and moved back). Therefore, we hypothesize:

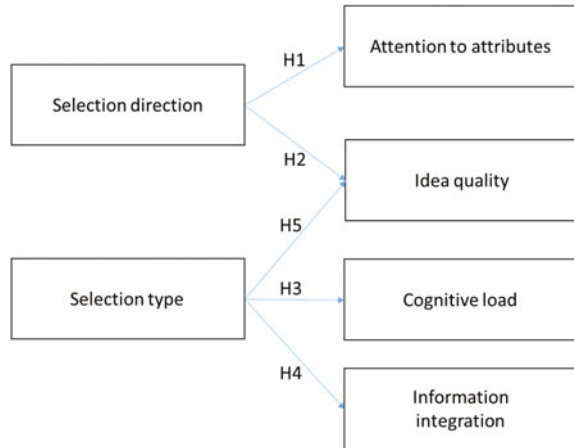
H4: Individuals that must make a single choice from a small set of ideas to consider will exhibit less information integration than individuals that must make multiple choices from the same idea set.

If we find support for H3 and H4, it will suggest that individuals experience lower cognitive load and less information integration when making a single choice and hence require less cognitive effort. Consequently, they can use their cognitive resources for assessing the quality of the idea as part of the idea description. The availability of additional cognitive resources may lead to a more thorough assessment of ideas, which we expect to result in a higher number of high-quality ideas in the set of filtered ideas. Therefore, we hypothesize:

H5: Individuals that must make a single choice from a small set of ideas to consider will achieve higher idea quality in the filtered set than individuals that must make multiple choices from the same set of ideas to consider.

### 3 Methods and Expected Contributions

We will test our hypotheses in a laboratory experiment with a  $2 \times 2$  between-subjects design (selection direction: lemons vs. stars; selection type: single vs. multiple choice). Figure 1 shows our research model. A web-based user interface will visualize a randomly selected set of four ideas to consider at a time. Each idea will be represented with its preview of the idea description (upon click the whole idea description will open), star rating, number of likes, and number of comments. The task description will prompt the user to either select or drop an idea depending on the treatment. For example, the prompt for the lemon and multiple choice treatment will be “Please choose from the set of ideas one or more ideas you deem worst”. Previous research suggests that users can recall up to 85 ideas for a similar task [25]. The innovation contest under investigation has generated more than 400 ideas. Therefore, users will process subsets of 80 ideas in multiple rounds. IS Master students from a European University will function as subjects and receive course credit as compensation. In addition to the constructs that we introduced

**Fig. 1** Research model

above and measure with eye tracking and heart rate monitoring equipment, we will assess idea quality as judged by external raters and the number of ideas in the final sets. Perceived need for cognition [21], perceived cognitive effort [26], perceived task importance, perceived decision confidence, perceived decision quality, and perceived satisfaction with product will be collected through surveys and help to control for individual effects. Additional control variables and demographics will also be collected.

We expect our findings to contribute to research on idea selection in several ways. First, we will provide in-depth insights into decision-making processes through our neuro-physical measurements. This will provide empirical evidence to the challenges of idea selection that have so far been collected from field observations and interviews. Second, our results on the effects of selection type and direction will provide theoretical underpinnings to our understanding of idea selection. Assuming that both theorized selection mechanisms have effects on the way people make decisions and on the idea quality in the filtered set, research can use these results and assess the effects on additional relevant outcomes such as process effectiveness. Moreover, our results will contribute to practice by showcasing an implemented prototype that supports processing of large idea sets. This would provide a novel approach to organizations that want to execute idea selections with organizational teams or outsource such selection processes to the crowd.

**Acknowledgements** The research was partially funded by the Austrian Science Foundation (FWF): P 29765-GBL.



## References

1. Bullinger, A.C., Neyer, A.K., Rass, M., Moeslein, K.M.: Community-based innovation contests: where competition meets cooperation. *Creativity Innov. Manage.* **19**, 290–303 (2010)
2. Bjelland, O.M., Wood, R.C.: An inside view of IBM's Innovation Jam. *MIT Sloan Manage. Rev.* **50**, 32–40 (2008)
3. Bayus, B.L.: Crowdsourcing new product ideas over time: an analysis of the Dell IdeaStorm community. *Manage. Sci.* **59**, 226–244 (2013)
4. Jouret, G.: Inside cisco's search for the next big idea. *Harvard Bus. Rev.* (2009)
5. Boudreau, K.J., Lakhani, K.R.: Using the crowd as an innovation partner. *Harvard Bus. Rev.* **91**, 60–69 (2013)
6. Dean, D.L., Hender, J.M., Rodgers, T.L., Santanen, E.L.: Identifying quality, novel, and creative ideas: constructs and scales for idea evaluation. *J. Assoc. Inf. Syst.* **7**, 30 (2006)
7. Girotra, K., Terwiesch, C., Ulrich, K.T.: Idea generation and the quality of the best idea. *Manage. Sci.* **56**, 591–605 (2010)
8. Velamuri, V.K., Schneckenberg, D., Haller, J., Moeslein, K.M.: Open evaluation of new product concepts at the front end of innovation: objectives and contingency factors. *R&D Manag.* (2015)
9. Merz, A., Seeber, I., Maier, R., Richter, A., Schimpf, R., Füller, J., Schwabe, G.: Exploring the effects of contest mechanisms on idea shortlisting in an open idea competition. In: 37th International Conference on Information Systems, Dublin, Ireland (2016)
10. Magnusson, P.R., Wästlund, E., Netz, J.: Exploring users' appropriateness as a proxy for experts when screening new product/service ideas. *J. Prod. Innov. Manag.* **33**, 4–18 (2016)
11. Yoon, P.K., Hwang, C.-L.: Multiple attribute decision making: an introduction. Thousand Oaks, CA: Sage (2011)
12. Fadel, K.J., Meservy, T.O., Jensen, M.L.: Exploring knowledge filtering processes in electronic networks of practice. *J. Manage. Inf. Syst.* **31**, 158–181 (2015)
13. Pilli, L.E., Mazzon, J.A.: Information overload, choice deferral, and moderating role of need for cognition: empirical evidence. *Rev. Administração* **51**, 36–55 (2016)
14. Petty, R.E., Cacioppo, J.T.: The elaboration likelihood model of persuasion. *Adv. Exp. Soc. Psychol.* **19**, 123–205 (1986)
15. Klein, M., Garcia, A.C.B.: High-speed idea filtering with the bag of lemons. *Decis. Support Syst.* **78**, 39–50 (2015)
16. Svenson, O.: Process descriptions of decision making. *Organ. Behav. Hum. Perform.* **23**, 86–112 (1979)
17. Gilbride, T.J., Allenby, G.M.: A choice model with conjunctive, disjunctive, and compensatory screening rules. *Mark. Sci.* **23**, 391–406 (2004)
18. Schulte-Mecklenbeck, M., Kühberger, A., Ranyard, R.: A handbook of process tracing methods for decision research: a critical review and user's guide. New York, NY: Taylor & Francis (2011)
19. Kolfschoten, G.L., Brazier, F.M.: Cognitive load in collaboration: convergence. *Group Decis. Negot.* **22**, 975–996 (2013)
20. De Vreede, G.J., Briggs, R.O., Massey, A.P.: Collaboration engineering: foundations and opportunities: editorial to the special issue on the journal of the association of information systems. *J. Assoc. Inf. Syst.* **10**, 121–137 (2009)

21. Levine, J.M., Moreland, R.L.: A history of small group research. In: Kruglanski A.W., Stroebe, W. (eds) *Handbook of the history of social psychology*, New York, NY: Psychology Press (2012)
22. Stibel, J.M., Dror, I.E., Ben-Zeev, T.: The collapsing choice theory: dissociating choice and judgment in decision making. *Theor. Decis.* **66**, 149–179 (2009)
23. Chandler, P., Sweller, J.: Cognitive load theory and the format of instruction. *Cogn. Instr.* **8**, 293–332 (1991)
24. Sweller, J., Van Merriënboer, J.J., Paas, F.G.: Cognitive architecture and instructional design. *Educ. Psychol. Rev.* **10**, 251–296 (1998)
25. Kornish, L.J., Ulrich, K.T.: Opportunity spaces in innovation: empirical analysis of large samples of ideas. *Manage. Sci.* **57**, 107–128 (2011)
26. Wang, W., Benbasat, I.: Interactive decision aids for consumer decision making in e-commerce: the influence of perceived strategy restrictiveness. *MISQ* **33**, 293 (2009)

# Blood Pressure Measurement: A Classic of Stress Measurement and Its Role in Technostress Research

Thomas Fischer, Gerhard Halmerbauer, Eva Meyr and René Riedl

**Abstract** In this paper, we present blood pressure measurement as an additional data collection method for technostress research. Considering that blood pressure is an important stress indicator and that, to the best of our knowledge, no prior Information Systems (IS) paper had an explicit focus on blood pressure measurement, the present paper is urgently needed, in particular from a technostress measurement perspective. We briefly describe the best practice in blood pressure measurement. Based on this foundation, we present a review of 15 empirical technostress studies that used blood pressure as a stress indicator. We find significant application variety in the extant literature, signifying the potential of blood pressure measurement for longitudinal technostress research. Yet, researchers should more explicitly adhere to international guidelines for the application of blood pressure measurement in future research, thereby securing data collection and data analysis quality.

**Keywords** Technostress · Stress · Blood pressure · Self-measurement  
NeuroIS

---

T. Fischer (✉) · G. Halmerbauer · E. Meyr · R. Riedl  
University of Applied Sciences Upper Austria, Steyr, Austria  
e-mail: thomas.fischer@fh-steyr.at

G. Halmerbauer  
e-mail: gerhard.halmerbauer@fh-steyr.at

E. Meyr  
e-mail: Eva-Maria.Meyr@students.fh-steyr.at

R. Riedl  
e-mail: Rene.Riedl@fh-steyr.at

R. Riedl  
University of Linz, Linz, Austria

## 1 Introduction

Technostress is a phenomenon that arises from “[d]irect human interaction with ICT [information and communication technologies], as well as perceptions, emotions, and thoughts regarding the implementation of ICT in organizations and its pervasiveness in society in general” [1, p. 18]. Technostress has become an established topic in Information Systems (IS) research. Evidence indicates a steadily increasing number of publications during the past years [2].

To advance research in this area, we reviewed technostress research before and called for several methodological adaptations. First, we called for more frequent measurement outside of research laboratories, mainly in order to create more externally valid findings [3]. Second, we recently highlighted that there is an overreliance on self-report measures and advocated the more frequent use of multi-method designs [2], predominantly because such an approach can explain additional variance in outcome variables that can hardly be explained through the use of single data collection methods [4]. Third, we provided an overview of previous organizational technostress research and the use of neurophysiological measures in this context [5]. In the last of these reviews we showed that only few studies applied neurophysiological measures (e.g., [6]). However, we also revealed that in those rare cases in which neurophysiological measures were used, measurement of cardiovascular indicators of stress (e.g., heart rate, heart rate variability and blood pressure) was most popular.

Despite the fact that heart rate and heart rate variability would come more easily to a researcher’s mind when thinking about longitudinal stress or technostress measurement in the field (e.g., using chest belts common in sports applications), in this paper we highlight why blood pressure should become an important method in IS technostress research. In the next section, we briefly summarize fundamental knowledge on blood pressure and its measurement, followed by a review of empirical technostress studies that used blood pressure as a stress indicator. We close this paper by providing insight on future research directions.

## 2 Blood Pressure Measurement

Blood pressure refers to the pressure that blood is exerting on the walls of blood vessels, typically in large arteries of the systemic circulation (e.g., brachial artery). The actual pressure is usually estimated through measurements of systolic and diastolic pressure levels on the outside of the vessel. Systolic blood pressure (SBP) represents the maximal force of the blood against vessel walls when the left ventricle of the heart is contracting (“systole”), while diastolic blood pressure

(DBP) represents the minimal force when the left ventricle is relaxed (“diastole”) [7]. Normal blood pressure levels are usually defined as a maximum of 120 mm Hg<sup>1</sup> SBP and a maximum of 80 mm Hg DBP [8].

In the past, blood pressure has been referred to as “the commonest measurement made in clinical practice” [9, p. 23]. Considering that hypertension (i.e., elevated blood pressure levels with SBP above 135/140 mm Hg and DBP above 85/90 mm Hg, [10, 11]) is prevalent in about one third of the population in Western countries such as Germany [12] or the USA [13], this statement is not surprising. Importantly, as shown in a review by Juster et al. [7], SBP and DBP have been used frequently in studies focusing on the effects of chronic stress, far more frequently than any other cardiovascular or respiratory indicator of stress.

In clinical practice, the auscultatory method is still frequently applied to measure blood pressure. Here, a trained person places a cuff on the upper arm (on the level of the brachial artery), inflates it above the level of systolic blood pressure and checks for specific sound patterns during deflation which indicate systolic and diastolic blood pressure (see [11] for more details, particularly on the phases and the varying sounds that are used as indicators). As this method requires a well-trained observer and is influenced particularly by noise in the environment, it was criticized in the literature [10, 11].

Mostly used in automated self-measurement devices, the oscillometric method uses oscillations of the blood vessel, instead of sounds, to estimate systolic and diastolic blood pressure [11]. This method is less susceptible to noise and also the cuff placement is less of an issue, though systolic and diastolic blood pressure are only estimated by algorithms of the involved devices, and not directly measured. It follows that measurements made with different devices should only be compared with caution [14]. However, evidence indicates high correlations between measurements using the auscultatory and oscillometric methods [11]. Thus, the use of self-measurement devices to determine blood pressure in IS research settings is an important measurement option.

Devices used for self-measurement of blood pressure (SBPM) are mostly offered for placement on the upper arm, wrist, or finger [10]. Due to the importance of measurement on heart level and the distance from the heart, upper arm devices are usually recommended, while finger monitors are less reliable, or even unreliable [10, 11]. Measurements on the wrist have some advantages. As an example, it is not susceptible to the circumference of the arm (which can have detrimental influence on the measurement with upper arm devices if a false cuff size has been chosen). However, measurements on the wrist are more susceptible to the right positioning of the arms, which requires thorough patient education to ensure measurement on heart level [14].

In the next section, we present the results of a review on the types of blood measurement methods and devices that have been applied in previous technostress

---

<sup>1</sup>“mm Hg” or “millimeter of mercury” (in a mercury sphygmomanometer) is a unit used to define the pressure of bodily fluids, with 1 mm Hg = 0.00133 bar.

research, followed by a section on further topics which have also already been investigated based on blood pressure measurement; also we briefly discuss the relationship of blood pressure with other neurophysiological measures.

### 3 Blood Pressure in Technostress Research

In order to identify relevant papers for our review, we used twelve papers drawn from previous reviews of technostress research (i.e., [15–26] taken from [1, 2, 5]) as the basis for a forward search in Google scholar (02/21/2017 to 02/24/2017). We opted for a forward search as we were interested in additional technostress research that used blood pressure as a stress indicator since the publications by the research groups of Werner Kuhmann and Wolfram Boucsein in the 1980s and 1990s.

Nine of these papers were drawn from previous reviews, though, as we focused on empirical research, we did not include the review paper by Boucsein [27]. Instead, four of the six papers that constituted the basis for the review paper by Boucsein [27] were included (i.e., two studies were excluded as they are only available in German). Based on a review of title and abstract of query results, we found three additional papers, thus resulting in a selection of fifteen papers for this review.

In Table 1, we summarized key features of these studies, including the main characteristics of used samples, setting of the study (i.e., laboratory or field research), measurement location (i.e., “arm” for measurement on the upper arm, finger, or wrist), measurement method (i.e., auscultatory, oscillometric, or other, if specified), and used devices.

Importantly, of the reviewed studies, only six were published in IS outlets (highlighted with an \* in Table 1), while the remaining nine publications were published in non-IS outlets (e.g., six publications in Ergonomics). Hence, in total we found only six studies which have applied blood pressure measurement in three decades of IS technostress research. In addition, though we explicitly conducted a forward research in order to identify more recent publications, we only identified two technostress studies after 2000 that included blood pressure measurement.

As a showcase for the potential of blood pressure measurement in longitudinal research designs, we found several studies that collected data over several days or even weeks. Schleifer and Okogbaa [25] collected data at three points during the day over four consecutive days. Johansson and Aronsson [19] collected blood pressure at five points during a work day (approximately every 2 h) over three non-consecutive days. Lundberg et al. [26] measured blood pressure every 10 min during their test phases (either 90 or 60 min) over three consecutive days. Wastell and Newman [17, 18] measured blood pressure every hour (from 9 a.m. to 1 p.m.) on work days over a period of twelve weeks (six weeks before and after system implementation). Most other studies, mainly conducted in laboratory settings, collected blood pressure during a baseline condition (first measurement) and then again at the end of their test (second measurement).

**Table 1** Overview of technostress studies applying blood pressure measurement

Study and samples	Setting	Location	Method/device
Johansson and Aronsson [19] <i>Sweden; insurance company; 21 f 6 f (system breakdown— improvised)</i>	Field	Arm	Auscultatory <i>Cardy 8 mini mercury sphygmomanometer</i>
Kuhmann et al. [21] <i>Germany; students; 22 f/46 m</i>	Laboratory	Arm	Auscultatory <i>Boso BC 40</i>
Kuhmann [22] <i>Germany; students; 10 f/38 m</i>	Laboratory	Arm	Auscultatory <i>Boso BC 40</i>
Schleifer and Okogba [25] <i>USA; clerical-secretarial agency; 45 f</i>	Laboratory	Arm	Other (infrasound) <i>Sphygmetrics, Model SR-2</i>
*Emurian [15] <i>USA; students; 10 m</i>	Laboratory	Arm	Auscultatory <i>PARAMED monitor</i>
*Emurian [16] <i>USA; students; 16 f/16 m</i>	Laboratory	Arm	Auscultatory <i>PARAMED monitor</i>
Lundberg et al. [26] <i>Sweden; students; 30 m</i>	Laboratory	Arm	Auscultatory <i>n/a (manually)</i>
Harada et al. [28] <i>Japan; students; 6 f/6 m</i>	Laboratory	Finger	Other (volume clamp) <i>Ohmeda 2300 (Finapres)</i>
Thum et al. [24] <i>Germany; students; 20 f/20 m</i>	Laboratory	Arm	Auscultatory <i>IBS SD-700A</i>
*Wastell and Newman [17] <i>UK; ambulance service; 18 n/a</i>	Field	Finger	Oscillometric <i>OMRON HEM-815F</i>
*Wastell and Newman [18] <i>UK; ambulance service; 18 n/a</i>	Field	Finger	Oscillometric <i>OMRON HEM-815F</i>
Kohlisch and Kuhmann [23] <i>Germany; students; 15 f/27 m</i>	Laboratory	Arm	Auscultatory <i>Boso BC 40</i>
*Henderson et al. [29] <i>Australia; students; 21 f/11 m</i>	Laboratory	Finger	Other (volume clamp) <i>n/a (Finapres)</i>
Hjortskov et al. [20] <i>Denmark; students; 12 f</i>	Laboratory	Arm/Finger	Oscillometric/Other <i>OMRON 705 CP Ohmeda 2300 (Finapres)</i>
*Clayton [30] <i>USA; students; 29 f/11 m</i>	Laboratory	Arm	Oscillometric <i>iHealth Lab model BP5</i>

Amongst the reviewed studies, measurement on the upper arm using the auscultatory method was most common. Some studies (i.e., [20, 28, 29]) measured blood pressure continuously using a volume clap on the finger [11], though this method can restrict the mobility of participants (due to constant connection to a monitor).

It is important to note that in none of the reviewed studies participants were advised to make self-measurements of their blood pressure, even in those studies in

which self-measurement devices based on the oscillometric method were used (i.e., [17, 18, 20, 30]). Hence, one of the main advantages of self-measurement devices, that is the reduction of the so-called “white-coat effect” (i.e., elevated blood pressure levels in the presence of medical professionals conducting blood pressure measurements, [31]) was not derived in these studies.

## 4 Blood Pressure Measurement in IS Technostress Research and in Other IS Domains

Our results indicate that blood pressure measurement was mainly used in laboratory studies. Moreover, we identified the examination of the blood pressure effects of duration and variability of system response times (SRT) as the most prevalent topic (i.e., 7 out of 15 papers [15, 16, 21–24, 28]). Importantly, research results have been mixed. Considering that the individual studies used different types of stressors (e.g., different implementations of “slow” and “fast” response times, or presence or absence of time pressure to perform a task), it is likely that the mixed results are a consequence of differences in stimuli.

Other stressors that have been investigated include the general impact of computer-based work on individual well-being [17–19], the effects of system breakdowns [19], levels of monotony of computer-based tasks and physiological consequences [26], task monitoring and related perceptions of performance pressure [29], lack of social support or even an unfriendly social environment and physiological consequences [20], and the physiological effects of separation from mobile devices [30].

Based on a database search,<sup>2</sup> we identified additional recent applications of blood pressure measurement that could be interesting to IS researchers and that are not directly related to stress. Turel et al. [32] investigated the effect of videogame addiction on blood pressure and other cardio-metabolic indicators (mediated by sleep patterns and level of obesity). Their study revealed a positive relationship between obesity and blood pressure and they concluded that elevated blood pressure is an important addiction-related health risks. Stafford et al. [33] investigated the acceptance of conversational robots by older people. Among other tasks, participants had to draw a representation of the conversational robots before interaction with the robot, while physiological parameters including blood pressure were measured. The study found that larger drawings were related to higher SBP after interacting with the robot. Finally, Why and Johnston [34] investigated the relationship between cynicism, state anger, and cardiovascular reactivity outside of social interaction, involving a computer-based task where the mouse was

---

<sup>2</sup>Search in ISI—Web of Science on 04/06/2017 using the query: Topic: “Blood pressure” AND Topic: “information technology” OR “information system” OR “human-computer interaction”, which resulted in 209 publications.



manipulated to arouse anger. They found that state anger moderated the positive relationship between individual cynicism and blood pressure (i.e., cynicism was only positively related to blood pressure when state anger was high).

## 5 Blood Pressure and Its Physiological Correlates

In the fifteen reviewed technostress studies, blood pressure has frequently been paired with other neurophysiological measures. In particular, further cardiovascular indicators (heart rate, heart rate variability) were applied in all studies, while measures of electrodermal activity (e.g., SCL, SCR) were applied in five studies, electromyography in three studies, and stress hormones were measured in two studies.

A number of the reviewed studies found similar correlation patterns for cardiovascular indicators (e.g., blood pressure positively correlates with heart rate, and negatively correlates with heart rate variability, [19, 24, 25]), though there have also been studies which found different correlation patterns. For example, a number of studies that investigated changes in workload (e.g., due to varying SRT) found that higher workload positively affects SBP, while no change in heart rate could be observed [16–18, 21]. Kuhmann et al. [21] argue that such differences might be caused by workload type. In essence, they argue that blood pressure is more closely related to physical workload, while heart rate is more closely related to mental workload. In a further study, Kohlisch and Kuhmann [23] showed in the context of a data entry task that low motor demands (i.e., a small number of keystrokes per minute) may also result in elevated blood pressure. It follows that the moderating effect of workload type is not well established.

Another potential explanation for differences between blood pressure reactivity and heart rate reactivity was provided by Hjortskov et al. [20]. They argued that blood pressure frequently stays elevated because it is influenced by local mechanisms (e.g., muscle activity) and is therefore not as sensitive to changes in mental load as HRV. They showed that though HRV (LF/HF ratio) returned to baseline after elimination of an experimentally induced stressor, blood pressure stayed high throughout the experiment and DBP even further increased during control sessions. It follows that HRV could be a good indicator of the presence of a stressor, BP could be a good indicator for the level of individual relaxation. In the context of general job stress, this argumentation has been established by Steptoe et al. [35] who showed that elevated blood pressure levels were still present after work.

For electrodermal activity, Thum et al. [24] reported that higher workload (due to short SRT) led to elevated blood pressure. Yet, electrodermal activity increased when individuals were confronted with a long SRT, which can be a sign of emotional strain, as was also reflected in self-ratings of the emotional state (i.e., short SRT was rated positively, while long SRT was rated negatively). In the same study a positive correlation was found for blood pressure and frontalis EMG activity.

Regarding stress hormones and blood pressure, we refer the reader to a study by Johansson and Aronsson [19] who reported on an improvised study during an unforeseen computer breakdown. They found that the breakdown led to elevated adrenaline excretion, which was accompanied by increased DBP.

As in the study by Thum et al. [24], other studies also reported deviations of blood pressure from individual self-reports. For example, while Kuhmann [22] found that participants rated a short SRT more positively, this was not reflected in any physiological changes, eventually caused by a lack of time pressure during task execution. Kohlisch and Kuhmann [23], in contrast, found differences in physiological states due to changes in SRT, which were not accompanied by significant changes in self-reported states. Replicating the results of Thum et al. [24], Harada et al. [28] also found that physiological activation was highest when SRT was fast, but self-reports on emotional states showed an opposite pattern. Clayton et al. [30] found that self-reports on emotional states (unpleasantness and anxiety) reflected physiological responses.

Against the background of the discussion in this section, it is important to consider the specific context of a study to understand potential increases, or decreases, of blood pressure. Overall, we see the application potential of blood pressure as a complement to other cardiovascular measures (e.g., heart rate), predominantly because it can be a good indicator of individual relaxation after stress onset. A general review of blood pressure and its regulation in the human body can be found in [36]. IS researchers are advised to consider the insights provided in this and similar reviews in their study design.

## 6 Conclusion and Further Directions

Blood pressure measurement has not played a significant role in IS technostress research so far, though we identified and reviewed a number of studies that could inform future studies. What is striking is the variety of study designs and different procedures (e.g., number of measurements, frequency, timing, methods, measurement location) that is observable in those few studies alone. To foster the application of blood pressure measurement in IS research in general, and specifically in technostress studies, we therefore recommend that researchers refer to the guidelines that are provided and regularly updated by international health organizations (e.g., [10, 11, 14, 31]). These guidelines also list potential confounders (see Table 2) that should be taken into account, and which have been controlled for in some of the reviewed studies (e.g., smoking, coffee or alcohol consumption, arm position), but, importantly, not in all studies (e.g., importance of uncrossed legs for blood pressure measurement).

An important point that is reiterated in all of these guidelines, and not only valid for blood pressure measurement, is that device selection should be made carefully. As a consequence of the prevalence of blood pressure-related health issues in human society, researchers are in the advantageous situation that numerous

**Table 2** Overview of important confounders in blood pressure measurement

Individual	Situation	Measurement
Caffeine consumption	Room temperature	Cuff size
Nicotine consumption	Background noise	Arm position (on heart level)
Alcohol consumption	Physical activity (exercise level)	Seating position (e.g., no crossed legs)
General health status (e.g., pregnancy, current medication, except aspirin)	Extent of talking	Relaxation before measurement

organizations exist worldwide which continuously pay attention to the validation of new blood pressure measurement devices (on the general importance of measurement issues in NeuroIS research, see [37]). An overview of validated devices and related studies can be found online (<http://www.dableducational.org/>). By using this list, we found, for example, that devices used in the reviewed studies [20] and [30] have been validated in accordance with international standards (i.e., [9]).

As there were no technostress studies in our review that actually applied self-measurement of blood pressure and due to the lack of studies that used wrist devices for this purpose, our own research group is currently concerned with the comparison of seven corresponding devices (i.e., OMRON RS8, RS6, RS3; BEURER BC40, BC57; BOSO Medistar+, Medilife PC3). Among other reasons, such comparison studies are important because recent research indicates that blood pressure could be an important stress indicator in stress-sensitive adaptive enterprise systems [38].

In conclusion, we hope that this review paper provides a useful overview of previous IS technostress studies and helps to establish blood pressure measurement as an extension to the current measurement toolset of technostress researchers. It should be noted though that the interpretation of blood pressure levels should be made with caution, because it is affected by many factors, all of which are potential confounders in scientific research. However, if blood pressure measurement is used as a complement to other neurophysiological measures (e.g., [1, 2, 4, 5, 39]), including measurement of brain activity (e.g., EEG; for details see Müller-Putz et al. [40]), then blood pressure will likely become a valuable extension in the IS researchers' toolset.

## References

1. Riedl, R.: On the biology of technostress: literature review and research agenda. *DATA BASE Adv. Inf. Syst.* **44**, 18–55 (2013)
2. Fischer, T., Riedl, R.: Technostress research: a nurturing ground for measurement pluralism? *Commun. Assoc. Inf. Syst.* **40**, 17 (2017)

3. Fischer, T., Riedl, R.: NeuroIS in situ: on the need for neuroIS research in the field to study organizational phenomena. In: Liang, T.-P., Yen, N.-S. (eds.) *Workshop on Information and Neural Decision Sciences*, pp. 20–21. National Chengchi University, Taipei, Taiwan (2014)
4. Tams, S., Hill, K., de Guinea, A.O., Thatcher, J., Grover, V.: NeuroIS—alternative or complement to existing methods? illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *J. Assoc. Inf. Syst.* **15**, 723–753 (2014)
5. Fischer, T., Riedl, R.: The status quo of neurophysiology in organizational technostress research: a review of studies published from 1978 to 2015. In: Davis, F.D., Riedl, R., Vom Brocke, J., Léger, P.-M., Randolph, A.B. (eds.) *Information Systems and Neuroscience, Lecture Notes in Information Systems and Organisation*, vol. 10, pp. 9–17. Springer International Publishing, Cham (2015)
6. Riedl, R., Kindermann, H., Auinger, A., Javor, A.: Technostress from a neurobiological perspective—system breakdown increases the stress hormone cortisol in computer users. *Bus. Inf. Syst. Eng.* **4**, 61–69 (2012)
7. Juster, R.-P., McEwen, B.S., Lupien, S.J.: Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neurosci. Biobehav. Rev.* **35**, 2–16 (2010)
8. Chobanian, A.V., Bakris, G.L., Black, H.R., Cushman, W.C., Green, L.A., Izzo, J.L., JR, Jones, D.W., Materson, B.J., Oparil, S., Wright, J.T., JR, et al.: The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. *JAMA* **289**, 2560–2572 (2003)
9. O'Brien, E., Atkins, N., Stergiou, G., Karpettas, N., Parati, G., Asmar, R., Imai, Y., Wang, J., Mengden, T., Shennan, A.: European society of hypertension international protocol revision 2010 for the validation of blood pressure measuring devices in adults. *Blood Pressure Monit.* **15**, 23–38 (2010)
10. O'Brien, E., Asmar, R., Beilin, L., Imai, Y., Mallion, J.-M., Mancina, G., Mengden, T., Myers, M., Padfield, P., Palatini, P., et al.: European society of hypertension recommendations for conventional, ambulatory and home blood pressure measurement. *J. Hypertens.* **21**, 821–848 (2003)
11. Pickering, T.G., Hall, J.E., Appel, L.J., Falkner, B.E., Graves, J., Hill, M.N., Jones, D.W., Kurtz, T., Sheps, S.G., Roccella, E.J.: Recommendations for blood pressure measurement in humans and experimental animals. Part 1: blood pressure measurement in humans. a statement for professionals from the subcommittee of professional and public education of the american heart association council on high blood pressure research. *Circulation* **111**, 697–716 (2005)
12. Neuhauser, H.K., Adler, C., Rosario, A.S., Diederichs, C., Ellert, U.: Hypertension prevalence, awareness, treatment and control in Germany 1998 and 2008–11. *J. Hum. Hypertens.* **29**, 247–253 (2015)
13. National Center for Health Statistics (US): *Health, United States: With Special Feature on Racial and Ethnic Health Disparities*. Hyattsville, MD, USA (2015)
14. Ogedegbe, G., Pickering, T.: Principles and techniques of blood pressure measurement. *Cardiol. Clin.* **28**, 571–586 (2010)
15. Emurian, H.H.: Physiological responses during data retrieval: comparison of constant and variable system response times. *Comput. Hum. Behav.* **7**, 291–310 (1991)
16. Emurian, H.H.: Cardiovascular and electromyograph effects of low and high density work on an interactive information system. *Comput. Hum. Behav.* **9**, 353–370 (1993)
17. Wastell, D., Newman, M.: Information system design, stress and organisational change in the ambulance services: a tale of two cities. *Acc. Manage. Inf. Technol.* **6**, 283–300 (1996)
18. Wastell, D., Newman, M.: Stress, control and computer system design: a psychophysiological field study. *Behav. Inf. Technol.* **15**, 183–192 (1996)
19. Johansson, G., Aronsson, G.: Stress reactions in computerized administrative work. *J. Organ. Behav.* **5**, 159–181 (1984)
20. Hjortskov, N., Rissén, D., Blangsted, A.K., Fallentin, N., Lundberg, U., Søgaard, K.: The effect of mental stress on heart rate variability and blood pressure during computer work. *Eur. J. Appl. Physiol.* **92**, 84–89 (2004)

21. Kuhmann, W., Boucsein, W., Schaefer, F., Alexander, J.: Experimental investigation of psychophysiological stress-reactions induced by different system response times in human-computer interaction\*. *Ergonomics* **30**, 933–943 (1987)
22. Kuhmann, W.: experimental investigation of stress-inducing properties of system response times. *Ergonomics* **32**, 271–280 (1989)
23. Kohlisch, O., Kuhmann, W.: System response time and readiness for task execution the optimum duration of inter-task delays. *Ergonomics* **40**, 265–280 (1997)
24. Thum, M., Boucsein, W., Kuhmann, W., Ray, W.J.: Standardized task strain and system response times in human-computer interaction. *Ergonomics* **38**, 1342–1351 (1995)
25. Schleifer, L.M., Okogbaa, G.O.: System response time and method of pay: cardiovascular stress effects in computer-based tasks\*. *Ergonomics* **33**, 1495–1509 (1990)
26. Lundberg, U., Melin, B., Evans, G.W., Holmberg, L.: Physiological deactivation after two contrasting tasks at a video display terminal: learning vs repetitive data entry. *Ergonomics* **36**, 601–611 (1993)
27. Boucsein, W.: Forty years of research on system response times—what did we learn from it? In: Schlick, C.M. (ed.) *Industrial Engineering and Ergonomics*, pp. 575–593. Springer, Berlin (2009)
28. Harada, H., Okabe, K., Katsuura, T., Kikuchi, Y.: Effects of time stress on psychophysiological responses during data entry tasks. *Appl. Hum. Sci. J. Physiol. Anthropol.* **14**, 279–285 (1995)
29. Henderson, R., Mahar, D., Salba, A., Deane, F., Napier, R.: Electronic monitoring systems: an examination of physiological activity and task performance within a simulated keystroke security and electronic performance monitoring system. *Int. J. Hum. Comput. Stud.* **48**, 143–157 (1998)
30. Clayton, R.B., Leshner, G., Almond, A.: The extended iSelf: the impact of iPhone separation on cognition, emotion, and physiology. *J. Comput. Mediated Commun.* **20**, 119–135 (2015)
31. Parati, G., Stergiou, G.S., Asmar, R., Bilo, G., Leeuw, P. de, Imai, Y., Kario, K., Lurbe, E., Manolis, A., Mengden, T., et al.: European society of hypertension guidelines for blood pressure monitoring at home: a summary report of the second international consensus conference on home blood pressure monitoring. *J. Hypertens.* **26**, 1505–1526 (2008)
32. Turel, O., Romashkin, A., Morrison, K.M.: Health outcomes of information system use lifestyles among adolescents: videogame addiction, sleep curtailment and cardio-metabolic deficiencies. *PLoS ONE* **11**, e0154764 (2016)
33. Stafford, R.Q., MacDonald, B.A., Li, X., Broadbent, E.: Older people’s prior robot attitudes influence evaluations of a conversational robot. *Int. J. Soc. Robot.* **6**, 281–297 (2014)
34. Why, Y.P., Johnston, D.W.: Cynicism, anger and cardiovascular reactivity during anger recall and human-computer interaction. *Int. J. Psychophysiol.* **68**, 219–227 (2008)
35. Steptoe, A., Cropley, M., Joeques, K.: Job strain, blood pressure and response to uncontrollable stress. *J. Hypertens.* **17**, 193–200 (1999)
36. Guyenet, P.G.: The sympathetic control of blood pressure. *Nat. Rev. Neurosci.* **7**, 335–346 (2006)
37. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**, i–xxxv (2014)
38. Adam, M.T.P., Gimpel, H., Maedche, A., Riedl, R.: Design blueprint for stress-sensitive adaptive enterprise systems. *Bus. Inf. Syst. Eng.* **59**(4), 277–291 (2017)
39. Schaaff, K., Degen, R., Adler, N., Adam, Marc T. P.: Measuring affect using a standard mouse device. *Biomed. Eng. (Biomedizinische Technik)* **57**(Suppl 1), 761–764 (2012)
40. Müller-Putz, G.R., Riedl, R., Wriessnegger, S.C.: Electroencephalography (EEG) as a research tool in the information systems discipline: foundations, measurement, and applications. *Commun. Assoc. Inf. Syst.* **37**, 46 (2015)

# On the Role of Users' Cognitive-Affective States for User Assistance Invocation

Celina Friemel, Stefan Morana, Jella Pfeiffer and Alexander Maedche

**Abstract** User assistance systems are often invoked automatically based on simple triggers (e.g., the assistant pops up after the user has been idle for some time) or they require users to invoke them manually. Both invocation modes have their weaknesses. Therefore, we argue that, ideally, the assistance should be invoked intelligently based on the users' actual need for assistance. In this paper, we propose a research project investigating the role of users' cognitive-affective states when providing assistance using NeuroIS measurements. Drawing on the theoretical foundations of the Attentional Control Theory, we propose an experiment that helps to understand how cognitive-affective states can serve as indicators for the best point of time for the invocation of user assistance systems. The research described in this paper will ultimately help to design intelligent invocation of user assistance systems.

**Keywords** Assistance · Invocation · NeuroIS · Attentional control theory  
Cognitive-affective user states · Affect · Mental effort

---

C. Friemel (✉) · S. Morana · J. Pfeiffer · A. Maedche  
Institute of Information Systems and Marketing (IISM),  
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany  
e-mail: celina.friemel@kit.edu

S. Morana  
e-mail: stefan.morana@kit.edu

J. Pfeiffer  
e-mail: jella.pfeiffer@kit.edu

A. Maedche  
e-mail: alexander.maedche@kit.edu

## 1 Introduction

Digital assistants like Siri or Alexa, chatbots like the ones on WeChat [1] and other forms of user assistance strongly developed over the last years and the trend towards providing advanced user assistance in digital services is even growing [2, 3]. The common idea of user assistance systems is to support users to perform their tasks better [4]. One of the early attempts to create such an assistant was Microsoft's Clippy. Yet, Clippy is a famous and regularly trending example for the dismal failure of such user assistance [5]. One of Clippy's most severe design mistakes was its proactive invocation mode. Proactively offering assistance at the right point can be a helpful feature in order to relieve the user's effort, ensure successful task performance and avoid errors [2, 6]. Certainly, Clippy appeared in the most inappropriate moments and interrupted users when not required. This led to Clippy's rapid downfall [5], which demonstrates the importance of a careful invocation design of assistance. The example shows that the communication via assistants needs to be well designed and adapted to the user in order to enhance trust and usage, and ultimately performance [7, 8]. Thus, the right timing of assistance invocation is an important design aspect [9, 10]. Invocation design in this context describes how assistance is activated. Some researchers [10] suggest a more advanced invocation which is provided by the system that "monitors the user in some way" (p. 504). However, existing approaches of assistance invocation are mainly dominated by either automatic activation or manual user requests. The automatic provisioning is often designed with static predefined rules by for example applying explicit user modelling, that incorporates the users' goals, needs, or other preferences to detect their need for assistance [11]. Both modes have been proven to not be entirely sufficient [6, 12], possibly because users are not always aware of when they need help and likewise frequently do not know how to use assistance effectively [13–15]. Furthermore, the "right time to intervene is [still] difficult to predict" [16] for the systems and consequently users get annoyed or out of flow when being interrupted at the wrong moments [5, 17]. As user assistance serves to relieve users' mental working memory [10] the system should not additionally burden the user with interruptions at the wrong time.

To address this research gap of providing intelligent invocation of user assistance [18], we argue for taking into account the cognitive-affective states of the user in real-time. With the term cognitive-affective states we refer to user states that involve both affective as well as cognitive activity [19]. These states heavily impact the interaction between humans and technology, users' need for assistance and consequently their task performance [20–23]. User states that influence users' interaction with IT are, in particular, task-dependent negative cognitive-affective states, such as frustration or anxiety [6] as well as high mental effort [24]. Thus, we assume that these user states correspondingly influence users' need for assistance [21, 22]. Drawing on the theoretical assumptions of the Attentional Control Theory [20], we argue that the assessment of the users' negative cognitive-affective states with neurophysiological data is an important design aspect to further improve user

assistance invocation [6, 25]. Ultimately, systems can automatically adjust assistance invocation to sensed user states to increase the users' efficiency, performance, and satisfaction [6]. In our research we follow a NeuroIS approach [8, 26]. One major advantage with regard to the outlined problem is the opportunity to observe latent variables, such as the users' need for assistance, "directly from body signals" [27]. Thus, the research question guiding this work is:

Can we identify users' need for assistance by unobtrusive and real-time measurement and analysis of cognitive-affective states of the user?

With this we want to expand and add value to NeuroIS literature as well as user assistance research by investigating psychophysiological correlates that reliably and timely detect the users' need for assistance and the IT-related behavior of assistance usage [26]. The research described in this paper will ultimately help to design intelligent invocation of user assistance systems.

## 2 Conceptual and Theoretical Foundations

### 2.1 Assistance and Invocation Modes

Assistance systems are provided in order to support users to perform their tasks better [4]. Assistance tends to become more and more tailored to the users' needs in order to increase performance at the right time and in the right context [2, 4]. Moreover, varying in their degree of system intelligence (e.g. provision of context-aware assistance) and interaction enabled by the system (e.g. offering highly sophisticated dialog interfaces), assistance systems can exhibit different maturity levels in terms of sensing the users' current environment and activities [4].

One critical design aspect when providing adequate user assistance is to determine when a user actually wants or needs assistance [16]. The right point of interrupting users has been studied extensively in the context of notifications [13, 28, 29]. Badly timed interruptions can cause deteriorated performance and decision-making, negative user states (like annoyance, frustration, cognitive overload) and ultimately distrust in the systems' competency and usage [28–31]. Correspondingly, the timing or invocation of assistance is crucial for the assistance's success. Recent research examining the right time to provide assistance agrees that it should be guided by the users' characteristics, needs, and context [6, 16, 25, 32]. First approaches of such user modeling [11, 16] did not provide accurate determinants for assistance invocation [12]. In line with other research [6, 33–35], we argue that users' cognitive-affective states determine the need for assistance. Modeling approaches on this new determinants exists [6, 35]. Yet, they focus on user modeling, or use manually pre-determined thresholds for assistance invocation. This reveals the lack of efficient timing to intervene with assistance that acts on the sensing of users' cognitive-affective states in real-time.



## 2.2 *Theorizing on Invocation Determinants: A NeuroIS Perspective*

**Assistance and NeuroIS.** In order to gain a better theoretical understanding of human behavior, NeuroIS research [36] is eager to find psychophysiological and neural correlates for already established IT constructs [26]. Especially in the context of offering user assistance, the traditional IS research methods like surveys and interviews encounter difficulties of reliably predicting users' need of assistance [15].

As users have been found to either not always be aware of when they need help, underestimate their need for assistance, or occasionally do not want to admit that they need assistance [13, 15, 37], a NeuroIS research approach can potentially shed light on this issue. However, objective measurement methods for users' assistance need in an objective and unobtrusive way are still absent. Neurophysiological tools offer great potential for new insights on user states by measuring direct responses to stimuli from the human body [27]. Applying this approach enables to capture unconscious processes that users might not be able to introspect or to gain insights on determinants of behavior that users are uncomfortable to report on [26]. Moreover, the possibility of obtaining real-time as well as continuous data enables analysis of temporal aspects and the measurement of simultaneous processes of constructs [26]. Consequently, this helps to design adaptive IT artifacts that consider user states which determine IT behavior [38].

With regard to designing assistance invocation, this integration of psychophysiological determinants for users' assistance needs is absent in existing research. First attempts to include affective user states into invocation determination of user assistance exist [6, 35, 39, 40]. Yet, these studies primarily address this issue only from a theoretical view. Particularly, to our knowledge, there exists no published research on determining assistance invocation by empirically integrating psychophysiological measurements in order to monitor users' states. However, not only from a NeuroIS perspective but also from a psychological view this approach can unlock great potential for reliably detecting users' need for assistance.

**Attentional Control Theory (ACT).** For decades now, researchers agree that affective as well as cognitive states have motivational properties that lead to observable behavior in IS [38, 41, 42] as well as non-IS contexts [20, 43].

In their psychological theory, Eysenck et al. [20, 44] offer valuable insights into the effects of affect and cognition on users' need for assistance. It describes the influence of especially negative affective states on people's task performance and related behavior, in particular, their coping strategies. In order to prevent a performance loss due to experienced negative affect and increased cognitive effort, people adjust their behavior with, for instance, searching for assistance. Eysenck et al. revealed that provoked anxiety impairs peoples' processing efficiency when working on a goal-directed task because people shift their attentional focus from the current task to the threatening stimulus. This increases their cognitive resource utilization. When responding to this change, people need additional resources

(internal or external) to cope with the situation in order to not experience a loss in performance effectiveness. However, this leads to a decrease in processing efficiency. One possibility to prevent people from such an efficiency loss caused by negative affective states is the provision of auxiliary processing resources [20], for example by offering assistance.

**Cognitive-Affective States.** The ACT mainly focuses on anxiety as negative affective state which impairs attentional control on a current task and ultimately efficiency and resulting performance [20]. Nevertheless, the theoretical implications of ACT have already been applied in the IS context and expanded to negative affective states, in general, that evidently influence IT-related behavior [42, 45]. By definition, an affective state arises from an individual's reaction to an event and influences cognitive, physiological, as well as behavioral components [6, 46]. Especially in the context of human-computer interaction, affective states play an important role when explaining user behavior [19, 21, 47]. Moreover, user states of high cognitive activity are often related to emotional responses; either in parallel or as interacting occurrences [26]. Monitoring user states that involve both affective as well as cognitive activity can reveal new insights on the users and their needs [26]. Baker et al. [19] refer to the latter as cognitive-affective user states. Within the context of assistance invocation especially negative cognitive-affective states are assumed to reveal important insights [23]. They are characterized by a negative affective valence, which can be assessed with the help of facial electromyography tools or facial expression analysis [48]. Examples of negative cognitive-affective user states are frustration or boredom [49]. Likewise, anxiety can be categorized as such a cognitive-affective state [50]. As antecedents of negative cognitive-affective user states certain stressors such as time pressure caused by dropped network connections and high task complexity have been found to increase the need for assistance [37].

### 3 Research Propositions

Drawing on ACT, we assume that negative cognitive-affective user states influence the related users' behavior in general and specifically the usage of assistance. This assumption is based on the influence of negative cognitive-affective user state on the users' attention focus. As this focus will shift from executing the task to coping with the affect-evoking stimuli the user has to invest more cognitive resources [20]. This increase in resource utilization is represented by users' mental effort, respectively the amount of cognitive resources that is required to manage the workload demanded by a task [51]. As users experience a negative cognitive-affective state, we therefore assume that their level of mental effort increases accordingly:

*Proposition P1: A negative cognitive-affective state increases users' mental effort.*

This increase in mental effort leads to potential loss in users' efficiency and ultimately task performance [20]. In order to prevent them from this undesirable outcome users are hypothesized to utilize coping strategies to decrease mental effort, respectively to increase efficiency and performance [20, 52, 53]. In the case of our research project, this is represented by the usage of a user assistance system that offers additional information [54, 55]:

*Proposition P2: Increased mental effort increases assistance usage.*

## 4 Proposed Methodology

In a first step, we propose to examine the effect of negative cognitive-affective user states on the users' IT-related behavior (in this case the usage of the offered user assistance). To test whether the theoretically derived propositions proof to be valid we plan to conduct a laboratory experiment.

We measure the users' cognitive-affective state with a combination of psychophysiological tools. We assess users' emotional valence via facial expressions with webcams [48] and users' mental effort via heart rate with ECG [8, 38]. ECG is one possibility to assess mental effort among others and has been found to be a reliable predictor for peoples' mental effort [56]. Furthermore, compared to other methods, such as EEG [57], it constitutes a minimally invasive measurement method [26]. Together with existing technical restrictions, this led to our decision of approximating mental effort via ECG measures. As the experimental context, we chose a travel booking scenario and formulate this as a goal-directed task according to the ACT [20] by providing participants incentives for a successful task execution. The experimental task will compromise the configuration of a travel with specific constraints with respect to budget and time. The participants' task is to configure the optimal travel in order to fulfill the experiments' objective. A virtual travel agency will offer the assistance that the user can consult manually, if needed. The participants will be randomly assigned to one of two treatment groups and a control group. The treatment structure will be composed of a *low negative affect* condition and a *high negative affect* condition. The treatments differ with respect to the amount of task features that stimulate negative cognitive-affective states. We then observe when they use the assistance and how this depends on the treatment.

## 5 Expected Contribution and Future Work

In the next steps, we will finalize the experimental design and carry out the experiment. Conducting the experiment and subsequently evaluating the experimental data will reveal valuable insights on the determinants of users' assistance needs. With this, we will gain first design knowledge on the appropriate invocation

timing of user assistance systems. We aim at validating the two suggested constructs of user states as neurophysiological correlates for assistance need in order to design neuro-adaptive invocation of user assistance in later stages of this research [26]. The final objective is to test and evaluate the resulting derived design knowledge in a follow-up experiment. Thereby, we contribute to collecting initial design knowledge towards finding the right moments for user assistance invocation. With ultimately testing the effects of such an invocation on the user, we will further contribute to research on user assistance in general. The hypothesized positive outcomes by applying timely user assistance [6] as well as possible negative effects caused by interrupting the user with the assistance itself [58] will be identified. Conceivably, the expected results will identify the optimal time to offer user assistance even before the user experiences any negative cognitive-affective state. User assistance in the future can then be designed to detect if the user is trending towards a negative cognitive-affective state in order to prevent any associated performance loss by offering timely user assistance.

Furthermore, the proposed experiment has some limitations that open up opportunities for future work on the results. As we will use an ECG measurement approach for assessing participants' mental effort, future research on the topic could evaluate other measurement methodologies in comparison. The approach of Eye-Fixation Related Potential (EFRP) proposed by Léger et al. [59] could offer further insights into participants' mental effort during task execution and help to identify when to offer user assistance. Moreover, we are aware of the fact that not only users' negative cognitive-affective states might constitute a need for user assistance. Other factors apart from these should be investigated in future research on the topic. Positive affective states have been found to influence task performance, too [60]. Together with examining the role of users' attention, this could complement the proposed research of finding the optimal timing for offering user assistance.

## References

1. Mou, Y., Xu, K.: The media inequality: comparing the initial human-human and human-AI social interactions. *Comput. Hum. Behav.* **72**, 432–440 (2017)
2. Sarikaya, R.: The technology behind personal digital assistants: an overview of the system architecture and key components. *IEEE Sig. Process. Mag.* **34**, 67–81 (2017)
3. Baig, E.C.: Personal digital assistants are on the rise (and they want to talk). <http://www.usatoday.com/story/tech/columnist/baig/2016/05/08/personal-digital-assistants-rise-and-they-want-talk/83715794/>
4. Maedche, A., Morana, S., Schacht, S., Werth, D., Krumeich, J.: Advanced user assistance systems. *Bus. Inf. Syst. Eng.* **58**, 2–5 (2016)
5. Veletsianos, G.: Cognitive and affective benefits of an animated pedagogical agent: considering contextual relevance and aesthetics. *J. Educ. Comput. Res.* **36**, 373–377 (2007)
6. Liao, W., Zhang, W., Zhu, Z., Ji, Q., Gray, W.D.: Toward a decision-theoretic framework for affect recognition and user assistance. *Int. J. Hum. Comput. Stud.* **64**, 847–873 (2006)

7. Parasuraman, R., Wilson, G.F.: Putting the brain to work: neuroergonomics past, present, and future. *Hum. Factors* **50**, 468–474 (2008)
8. Vom Brocke, J., Riedl, R., Léger, P.-M.: Application strategies for neuroscience in information systems design science research. *J. Comput. Inf. Syst.* **53**, 1–13 (2013)
9. Silver, M.S.: On the design features of decision support systems: the role of system restrictiveness and decisional guidance. In: *Handbook on Decision Support Systems 2: Variations*. pp. 261–293. Springer, Berlin, Heidelberg (2008)
10. Gregor, S., Benbasat, I.: Explanations from intelligent systems: theoretical foundations and implications for practice. *MIS Q.* **23**, 497–530 (1999)
11. Horvitz, E., Breese, J., Heckerman, D., Hovel, D., Rommelse, K.: The lumiere project: bayesian user modeling for inferring the goals and needs of software users. In: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, pp. 256–265 (1998)
12. Antwarg, L., Lavie, T., Rokach, L., Shapira, B., Meyer, J.: Highlighting items as means of adaptive assistance. *Behav. Inf. Technol.* **32**, 1–17 (2012)
13. Schiaffino, S., Amandi, A.: Polite personal agents. *IEEE Intell. Syst.* **21**, 12–19 (2006)
14. Babin, L.M., Tricot, A., Mariné, C.: Seeking and providing assistance while learning to use information systems. *Comput. Educ.* **53**, 1029–1039 (2009)
15. Novick, D.G., Elizalde, E., Bean, N.: Toward a more accurate view of when and how people seek help with computer applications. In: *ACM 25th International Conference on Design of Communication*, pp. 95–102 (2007)
16. Ginon, B., Stumpf, S., Jean-Daubias, S.: Towards the right assistance at the right time for using complex interfaces. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces—AVI '16*. pp. 240–43. ACM Press, New York (2016)
17. McFarlane, D., Latorella, K.: The scope and importance of human interruption in human-computer interaction design. *Hum. Comput. Interact.* **17**, 1–61 (2002)
18. Morana, S., Schacht, S., Scherp, A., Maedche, A.: A review of the nature and effects of guidance design features. *Decis. Support Syst.* **97**, 31–42 (2017)
19. Baker, R.S.J., D’Mello, S.K., Rodrigo, M.M.T., Graesser, A.C.: Better to be frustrated than bored: the incidence, persistence, and impact of learners’ cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum. Comput. Stud.* **68**, 223–241 (2010)
20. Eysenck, M.W., Derakshan, N., Santos, R., Calvo, M.G.: Anxiety and cognitive performance: attentional control theory. *Emotion* **7**, 336–353 (2007)
21. Klein, J., Moon, Y., Picard, R.W.: This computer responds to user frustration: theory, design, and results. *Interact. Comput.* **14**, 119–140 (2002)
22. Picard, R.W.: Affective computing: challenges. *Int. J. Hum. Comput. Stud.* **59**, 55–64 (2003)
23. Fairclough, S.H.: Fundamentals of physiological computing. *Interact. Comput.* **21**, 133–145 (2009)
24. Fehrenbacher, D.D., Djamshbi, S.: Information systems and task demand: an exploratory pupillometry study of computerized decision making. *Decis. Support Syst.* **97**, 1–11 (2017)
25. Yorke-Smith, N., Saadati, S., Myers, K.L., Morley, D.N.: The design of a proactive personal agent for task management. *Int. J. Artif. Intell. Tools* **21**, 1250004-1-30 (2012)
26. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Pavlou, P.A., Müller-Putz, G.R., Riedl, R., vom Brocke, J., Weber, B.: On the use of neurophysiological tools in is research: developing a research agenda for NeuroIS. *MIS Q.* **36**, 679–702 (2012)
27. Liang, T.-P., vom Brocke, J.: Special issue: neuroscience in information systems research. *J. Manage. Inf. Syst.* **30**, 7–12 (2014)
28. Bailey, B.P., Konstan, J.A.: On the need for attention-aware systems: measuring effects of interruption on task performance, error rate, and affective state. *Comput. Hum. Behav.* **22**, 685–708 (2006)
29. Adamczyk, P.D., Bailey, B.P.: If not now when? the effects of interruption at different moments within task execution. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, vol. 6, pp. 271–278 (2004)

30. Speier, C., Valacich, J.S., Vessey, I.: The influence of task interruption on individual decision making: an information overload perspective. *Decis. Sci.* **30**, 337–360 (1999)
31. LeeTirman, S., Cutrell, E.B., Czerwinski, M., Hoffman, H.: Effective notification systems depend on user trust. In: *Human-Computer Interaction—INTERACT 2001 Conference Proceedings*, pp. 684–685 (2001)
32. Wang, J., Wang, Y., Wang, Y.: CAPFF: A context-aware assistant for paper form filling. *IEEE Trans. Hum. Mach. Syst.* 1–6 (2016)
33. Hudlicka, E.: To feel or not to feel: the role of affect in human–computer interaction. *Int. J. Hum. Comput. Stud.* **59**, 1–32 (2003)
34. Liao, W., Zhang, W., Zhu, Z., Ji, Q.: A decision theoretic model for stress recognition and user assistance. In: *Proceedings of Twentieth National Conference on Artificial Intelligence. Seventeenth Innovative Applications of Artificial Intelligence Conference*, pp. 529–534 (2005)
35. Li, X., Ji, Q.: Active affective state detection and user assistance with dynamic bayesian networks. *IEEE Trans. Syst. Man Cybern. Syst. Hum.* **35**, 93–105 (2005)
36. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**, i–xxxv (2014)
37. Ellis, S., Tyre, M.J.: Helping relations between technology users and developers: a vignette study. *IEEE Trans. Eng. Manage.* **48**, 56–69 (2001)
38. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS—Information Systems and the Brain*, Springer, Berlin (2016)
39. Dong, J., Yang, H.I., Oyama, K., Chang, C.K.: Human desire inference process based on affective computing. In: *Proceedings of International Computer Software and Applications Conference*, pp. 347–350 (2010)
40. Frank, K., Robertson, P., Gross, M., Wiesner, K.: Sensor-based identification of human stress levels. In: *2013 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops*, pp. 127–32 (2013)
41. Gregor, S., Lin, A.C.H., Gedeon, T., Riaz, A., Zhu, D.: Neuroscience and a nomological network for the understanding and assessment of emotions in information systems research. *J. Manage. Inf. Syst.* **30**, 13–48 (2014)
42. Hibbeln, M., Jenkins, J.L., Schneider, C., Valacich, J.S., Weinmann, M.: How is your user feeling? Inferring emotion through human-computer interaction devices. *MIS Q.* **41**, 1–22 (2017)
43. Brown, J.S., Farber, I.E.: Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.* **48**, 465–495 (1951)
44. Eysenck, M.W., Derakshan, N.: New perspectives in attentional control theory. *Pers. Individ. Differ.* **50**, 955–960 (2011)
45. Anderson, B.B., Vance, A., Kirwan, C.B., Eargle, D., Jenkins, J.L.: How users perceive and respond to security messages: a NeuroIS research agenda and empirical study. *Eur. J. Inf. Syst.* **2016**, 1–27 (2016)
46. Nass, C., Brave, S.: Emotion in human–computer interaction. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 77–92 (2007)
47. Scheirer, J., Fernandez, R., Klein, J., Picard, R.W.: Frustrating the user on purpose: a step toward building an affective computer. *Interact. Comput.* **14**, 93–118 (2002)
48. Ekman, P.: Facial expression and emotion. *Am. Psychol.* **48**, 384–392 (1993)
49. D’Mello, S., Graesser, A.: The half-life of cognitive-affective states during complex learning. *Cogn. Emot.* **25**, 1299–1308 (2011)
50. Schlenker, B.R., Leary, M.R.: Social anxiety and self-presentation: a conceptualization and model. *Psychol. Bull.* **92**, 641–669 (1982)
51. Sweller, J., van Merriënboer, J., Paas, F.: Cognitive architecture and instructional design. *Educ. Psychol. Rev.* **10**, 251–296 (1998)
52. Peavler, S.W.: Pupil size, information overload and performance differences. *Psychophysiology* **11**, 559–566 (1974)

53. Paas, F., Van Merriënboer, J.J.G.: The efficiency of instructional conditions: an approach to combine mental effort and performance measures. *Hum. Factors* **35**, 737–743 (1993)
54. Folkman, S., Moskowitz, J.T.: Positive affect and the other side of coping. *Am. Psychol.* **55**, 647–654 (2000)
55. Beaudry, A., Pinsonneault, A.: Understanding user responses to information technology: a coping model of user adaptation. *MIS Q.* **29**, 493–524 (2005)
56. Haapalainen, E., Kim, S., Forlizzi, J.F., Dey, A.K.: Psycho-physiological measures for assessing cognitive load. In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pp. 301–310 (2010)
57. Mirhoseini, S.M.M., Léger, P.-M., Sénécal, S.: The influence of task characteristics on multiple objective and subjective cognitive load measures. Presented at the information systems and neuroscience, Springer, Berlin (2017)
58. Jenkins, J.L., Anderson, B.B., Vance, A., Kirwan, C.B., Eargle, D.: More harm than good? how messages that interrupt can make us vulnerable. *Inf. Syst. Res.* **27**(4), 880–896 (2016)
59. Léger, P.M., Titah, R., Sénécal, S., Fredette, M., Courtemanche, F., Labonte-Lemoyne, É., De Guinea, A.O.: Precision is in the eye of the beholder: application of eye fixation-related potentials to information systems research. *J. Assoc. Inf. Syst.* **15**, 651–678 (2014)
60. Jung, N., Wranke, C., Hamburger, K., Knauff, M.: How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety. *Front. Psychol.* **5** (2014)

# Measuring and Explaining Cognitive Load During Design Activities: A Fine-Grained Approach

Barbara Weber, Manuel Neurauter, Andrea Burattin,  
Jakob Pinggera and Christopher Davis

**Abstract** Recent advances in neuro-physiological measurements resulted in reliable and objective measures of Cognitive Load (CL), for example, using pupillary responses. However, continuous measurement of CL in software design activities, for example, conceptual modeling, has received little attention. In this paper, we present the progress of our work intended to close this gap by continuously measuring cognitive load during design activities. This work aims at advancing our understanding of WHEN and WHY designers face challenges. For this, we attempt to explore and explain the occurrence of CL using fine-granular units of analysis (e.g., type of subtasks, evolution of design artifact's quality, and manner of technology use). We expect implications for the future development of intelligent software systems, which are aware WHEN a particular designer experiences challenges, but also WHY challenges occur.

**Keywords** Business process management · Process modeling · Process model creation · Eye tracking · Cognitive load

## 1 Introduction

Contemporary software engineering practice differs fundamentally, as cloud-based apps and services do, from the monolithic mainframes of the 1980s. This presents a challenge, since timing and duration of software engineering design activities today is diffuse when compared to projects managed using structured methods and tools such as CASE (Computer Aided Software Engineering) and Integrated

---

B. Weber (✉)

Technical University of Denmark, Kongens Lyngby, Denmark  
e-mail: bweb@dtu.dk

B. Weber · M. Neurauter · A. Burattin · J. Pinggera  
University of Innsbruck, Innsbruck, Austria

C. Davis  
University of South Florida, Tampa, USA



Development Environments (IDEs). It has become increasingly difficult to assess WHEN software engineers (designers in the following) experience challenges while conducting a design activity (e.g., creating a conceptual model or programming) and to explain WHY these challenges occur.

The cognitive demands imposed on the designer are commonly described as cognitive load (CL) [1]. Recent advances in neuro-physiological measurements resulted in reliable and objective measures of continuous CL [1]. However, continuous measurement of CL in software design activities up to now has received little attention. Moreover, there is no comprehensive understanding of the factors influencing a designers CL while conducting a software design activity.

Qualitative approaches in the context of software engineering typically identify challenges designers face through manual analysis and subsequent coding of data (e.g., [2, 3]) rather than the usage of neuro-physiological measurements of CL. Quantitative studies to CL measurements, in turn, are typically either perception-based and not continuous (e.g., using the NASA-TLX instrument for measuring CL [4]) or conducted as stimuli-response experiments where markers are available (i.e., it is known when the stimulus occurs) that can be related to the responses (i.e., change in pupil dilation), for example, [5]. In contrast to stimulus-response settings, where changes in cognitive load in response to a stimulus induced by experimenters are evaluated, investigating CL during design activities is less structured and inherent to how a designer's individual design process unfolds.

We intend to close this gap by measuring CL through objective, neuro-physiological measures in a more realistic work setting where no markers exist. We aim to make software engineering processes and their cognitive demands more tractable, advancing our understanding of WHEN and WHY designers face challenges. For this, we attempt to explore and explain the occurrence of CL using more granular units of analysis (including several process-oriented factors such as the type of sub-tasks; the evolution of the quality of the design artifact, and the manner of technology use) derived from the designer's interactions with the design platform and eye fixations on the various parts of the design platform.

This paper focuses on one frequently re-occurring software design activity, that is, conceptual modeling, but eventually aims at broadening the range of design components to include program blocks. Our research is expected to contribute towards a better understanding of WHEN and WHY high CL occurs in design activities by gathering empirical data regarding variations and changes in CL and various process-oriented factors to potentially explain these changes.

## 2 Cognitive Load in Design Activities

Design activities, for example, conceptual modeling, involve the construction of a mental model of the domain from an informal requirements description and its externalization using the elements provided by the modeling notation [6] by using

the tools provided by a design platform, for example, the modeling editor [7]. During the externalization process, the designer evolves the design artifact, that is, conceptual model, through a series of interactions from an initial state through intermediate states to a final state reflecting the requirements of the domain. When performing a design activity, the designer exploits the malleability of their mental model to decompose cognitively ‘digestible’ sub-tasks, for example, a group of model elements. Recomposition maintains the integrity of the components and the intellectual control of the designer. This is the ‘dance’ of design that is choreographed using notations and design platforms [8]. The cognitive demands imposed on the designer are commonly described as CL, dependent on the task’s inherent complexity, the design platform, the designer’s domain knowledge, design expertise, and cognitive abilities. As a response to the cognitive challenges, the designer’s CL changes throughout the course of the design activity [9]. To objectively measure CL and to determine WHEN designers are challenged, neuro-physiological tools, for example, eye tracking [1], can be used. For a particular designer we assume designer-specific factors to remain stable during a design activity. Possible explanations for changes in CL probably stem from changes in task difficulty throughout a design activity and the designer’s interactions with design platform and design artifact. A messy intermediate design artifact could, for example, lead to higher CL when working on the artifact afterwards. To understand WHY challenges during the design process occur without having markers, we attempt to connect the CL data with data regarding the sub-tasks of the design process (to trace down differences in task difficulty throughout the design activity), the evolution of the design artifact and its quality, and the manner of technology use.

### 3 Research in Progress

Subsequently, we outline the current status of our work and provide details regarding our future endeavor.

#### 3.1 Step 1: Data Collection

To continuously assess CL we measure pupil dilation, which (under conditions of controlled illumination) reliably indicates CL [1]. Alternative load measures and the reasoning for choosing pupil dilation for our study are discussed in [9]. Process-oriented factors as possible explanations for CL are measured by collating interactions with the design platform using Cheetah Experimental Platform (CEP) [10] and eye movement data (e.g., fixations) using the Tobii-TX300 eye tracker. For synchronizing interactions, fixations, and pupillary response data and for performing data treatment, we rely on a dedicated platform extending the

capabilities of CEP towards analyzing CL [11, 12]. We collected data of 117 novice student modelers, who created a conceptual model using BPMN [9] after a training phase.

### 3.2 Step 2: Measuring Process-Oriented Factors

**Step 2.1: Measure sub-task specific CL:** For conceptual modeling, [13] showed the existence of the sub-tasks problem understanding, method finding, modeling, reconciliation, and validation. Since different types of sub-tasks involve different underlying cognitive processes, the changes in CL can stem from the type of sub-task the designer is currently engaged in. For this, we intend to automatically discover the different sub-tasks the user is engaged in at different periods of time when interacting with the design platform. For this, we rely on an existing task model [10, 13] and formulate the challenge of aligning the data coming from different modalities (i.e., neuro-physiological data, interactions with design platform) with the task model as a classification problem and solve it using supervised learning approaches, for example, Support Vector Machines [14]. For validation purposes the classifier will be compared with a previously defined gold standard. Initial results are promising [15, 16]. The alignment will be then used to slice the design activity into periods of time with start and end timestamps to calculate sub-task specific CL. Different measures such as the average CL, accumulated CL, or the instantaneous CL as suggested by Xie and Salvendy [17] can be calculated for the respective periods of time. For an overview of CL measurements in general and neuro-physiological measurements in particular please see [1].

**Step 2.2: Measurement of the design artifact's quality evolution:** When transforming a design artifact from one state into another, its quality (e.g., element alignment) can change and impact subsequent modifications. Put differently, a design artifact whose quality gradually degrades can make subsequent changes difficult by raising CL due to decreased readability. For this, quality will be operationalized as a set of properties [18, 19] (e.g., number of syntactical errors, alignment of elements). Values for each property can be calculated for each intermediate state. For this, we build upon infrastructure from the Austrian-funded ModErARe project [19].

**Step 2.3: Conceptualize and measure manner of technology use:** In addition, the manner of using the provided design platform can have an impact on CL, for example, tool features that are used effectively can lower CL, be ineffective, or even increase CL if used inappropriately. Here we plan to develop a rich conceptualization of technology use in line with [20] that goes beyond simple quantitative measures. Refactoring tools, for example, are frequently used as part of reconciliation sub-tasks with the goal to reduce CL of subsequent sub-tasks (e.g., by improving the understandability of the partial design artifact). When analyzing the potential impact of using refactoring tools, counting the number of its invocations is not sufficient, but it has to be considered whether its use was effective and led to

improvements of the partial design artifact. Moreover, the potential benefit of refactoring depends on when in the process it is applied and how much the quality of the design artifact is impaired at the moment of its application. Therefore, our conceptualization of the manner of technology use will capture how a particular designer uses the design platform to accomplish a certain (sub-)task considering the state of the intermediate design artifact. Such data can be measured using the designer's interactions with the design platform.

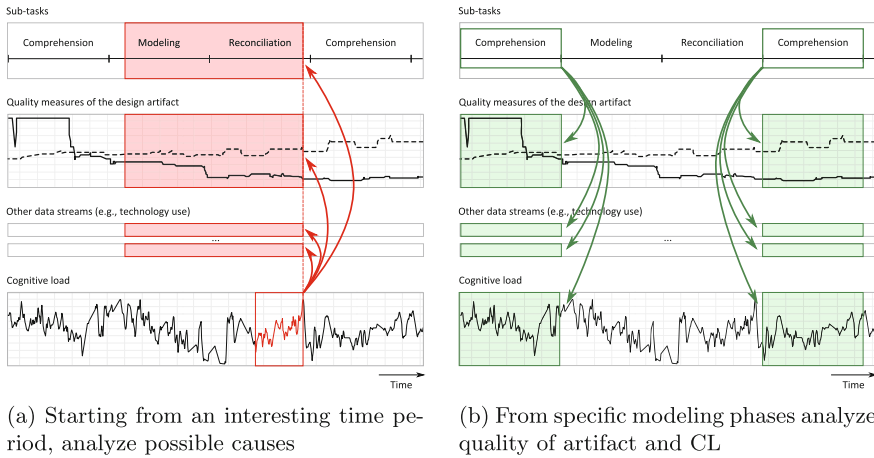
### **3.3 Step 3: Data Analysis**

The analysis of the collected data can be grouped in two families: intra- and inter-subject analysis. The former case considers data stemming from different modalities, taking a single subject into account. The latter groups subjects by different aspects and analyzes the relationships among those groups. For intra-subject analysis, two different analyses might be considered (cf. Fig. 1). In both cases, the analysis combines the information streams we extracted, for example, the sub-tasks, the quality measurements of the designed artifacts, the manner of technology use, and the cognitive load.

After running through a data cleaning procedure described in [12] the starting point for the first analysis scenario (cf. Fig. 1a) is the identification of interesting time periods regarding CL. Then, we examine the immediate history of the other streams for identifying possible causes for the high CL. The second analysis approach (cf. Fig. 1b), on the contrary, slices the design activity into periods of time by considering process-oriented factors (i.e., sub-tasks, evolution of quality of design artifact, and manner of technology use). For example, as illustrated in Fig. 1, the design activity is sliced into sub-tasks. Another possibility of slicing the design activity could be identifying periods with high quality of the design artifact and periods with low quality. The analysis then compares periods of time of the same type (e.g., the different sub-tasks in Fig. 1b) in terms of differences in CL. Considering inter-subject analysis, subjects are grouped by process-oriented factors and then tested for group differences. For example, subjects could be grouped based on their modeling behavior (e.g., subjects with reconciliation phases just at the end of the modeling session versus subjects with reconciliation phases throughout the modeling session) and groups could be tested for differences in their average CL.

## **4 Summary and Expected Impact**

In this paper, we investigate WHEN designers experience challenges by measuring CL and aim to explain changes in CL using different process-oriented factors. We have completed data collection and made substantial progress regarding the operationalization of process-oriented factors. Next, we plan to analyze the data as



**Fig. 1** Approaches for the intra-subject data analysis

outlined. If the approach proves viable we intend to broaden our scope and to address other software design activities like programming.

Assuming we are able to demonstrate the impact of process-oriented factors on CL, we expect our research to result in revised guidelines on how to investigate design activities and evaluate design artifacts. Future research will be advised to depart from a pure black-box approach and to increasingly consider process-oriented factors impacting CL (or other antecedents of task performance). We further expect implications for the development of neuro-adaptive systems that are not only aware WHEN a particular designer experiences challenges, but also WHY and can react with personalized feedback or adaptation.

## References

1. Chen, F., Zhou, J., Wang, Y., Yu, K., Arshad, S.Z., Khawaji, A., Conway, D.: Eye-Based Measures, pp. 75–85. Springer International Publishing, Cham (2016)
2. Hungerford, B.C., Hevner, A.R., Collins, R.W.: Reviewing software diagrams: a cognitive study. *IEEE Trans. Softw. Eng.* **30**, 82–96 (2004)
3. Haisjackl, C., Zugall, S., Soffer, P., Hadar, I., Reichert, M., Pinggera, J., Weber, B.: Making sense of declarative process models: common strategies and typical pitfalls. In: *Proceeding of Conference on BPMDS’13*, pp. 2–17 (2013)
4. Marculescu, B., Poulding, S.M., Feldt, R., Petersen, K., Torkar, R.: Tester interactivity makes a difference in search-based software testing: a controlled experiment. *Inf. Softw. Technol.* **78**, 66–82 (2015)
5. Knapen, T., de Gee, J.W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., Theeuwes, J.: Cognitive and ocular factors jointly determine pupil responses under equiluminance. *PLoS ONE* **11**, 1–13 (2016)
6. Recker, J.C., Safrudin, N., Rosemann, M.: How novices design business processes. *Inf. Syst.* **37**, 557–573 (2012)

7. Soffer, P., Kaner, M., Wand, Y.: Towards understanding the process of process modeling: theoretical and empirical considerations. In: Proceedings of conference on ER-BPM'11, pp. 357–69 (2012)
8. Caputo, K.: CMM Implementation Guide: Choreographing Software Process Improvement (Unisys Series). Addison-Wesley, Boston (1998)
9. Neurauder, M., Pinggera, J., Martini, M., Burattin, A., Furtner, M., Sachse, P., Weber, B.: The influence of cognitive abilities and cognitive load on business process models and their creation. In: Proceedings of Conference on NeuroIS'15, pp. 107–115 (2015)
10. Pinggera, J., Zugal, S., Weber, B.: Investigating the Process of Process Modeling with Cheetah Experimental Platform. In: Proceedings of Conference on ER-POIS'10, pp. 13–18 (2010)
11. Weber, B., Neurauder, M., Pinggera, J., Zugal, S., Furtner, M., Martini, M., Sachse, P.: Measuring cognitive load during process model creation. In: Proceedings of Conference on NeuroIS'15, pp. 129–136 (2015)
12. Zugal, S., Pinggera, J., Neurauder, M., Maran, T., Weber, B.: Cheetah experimental platform web 1.0: cleaning pupillary data. Technical report (arXiv.org)
13. Pinggera, J.: The process of process modeling. PhD thesis, University of Innsbruck, Department of Computer Science (2014)
14. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995)
15. Weber, B., Pinggera, J., Neurauder, M., Zugal, S., Martini, M., Furtner, M., Sachse, P., Schnitzer, D.: Fixation patterns during process model creation: initial steps toward neuro-adaptive process modeling environments. In: Proceedings of Conference on HICSS'16, pp. 600–609 (2016)
16. Burattin, A., Kaiser, M., Neurauder, M., Weber, B.: Eye tracking meets the process of process modeling: a visual analytic approach. In: Proceedings of Conference on TAProViz'16 (2016)
17. Xie, B., Salvendy, G.: Prediction of mental workload in single and multiple tasks environments. *Int. J. Cogn. Ergonom.* **4**, 213–242 (2000)
18. Burattin, A., Bernstein, V., Neurauder, M., Soffer, P., Weber, B.: Detection and quantification of flow consistency in business process models. *Softw. Syst. Model.*, pp. 1–22 (2017)
19. Haisjackl, C., Burattin, A., Soffer, P., Weber, B.: Visualization of the evolution of layout metrics for business process models. In: Proceedings of Conference on TAProViz'16 (2016)
20. Burton-Jones, A., Straub, D.W.: Reconceptualizing system usage: an approach and empirical test. *Inf. Syst. Res.* **17**, 228–246 (2006)

# How Product Decision Characteristics Interact to Influence Cognitive Load: An Exploratory Study

Sylvain Sénécal, Pierre-Majorique Léger, René Riedl  
and Fred D. Davis

**Abstract** The objective of this laboratory experiment was to explore how product decision characteristics interact to influence the decision-maker's cognitive load. A between-subject experiment with 23 participants was performed to test how four decision characteristics (Decision set size, Attribute value format, Display format, and Information sorting) interact to influence participants' cognitive load. Eye-tracking was used to assess cognitive load. Results indicate that the four product decision characteristics interact to influence cognitive load. We found, for example, that as the decision set size increased, the influence of attribute value format, display format, and information sorting on cognitive load varied. Theoretical contributions and managerial implications are discussed.

**Keywords** Decision characteristics · Decision-making · Eye-tracking  
Cognitive load · Information display

---

S. Sénécal (✉) · P.-M. Léger  
HEC Montréal, Montreal, Canada  
e-mail: ss@hec.ca

P.-M. Léger  
e-mail: pml@hec.ca

R. Riedl  
University of Applied Sciences Upper Austria, Steyr, Oberösterreich, Austria  
e-mail: Rene.Riedl@fh-steyr.at

R. Riedl  
University of Linz, Linz, Austria

F. D. Davis  
Texas Tech University, Lubbock, USA  
e-mail: fred.davis@ttu.edu

## 1 Introduction

Human decision making (e.g., choosing a product on a retail website) depends on a number of factors. Payne et al. [1] suggest that the characteristics of the decision-maker, the decision situation, and the social context may influence the decision strategy and thus the final decision. Prior research shows, for instance, that consumers use different decision-making strategies based on the number of alternatives (e.g., products) to choose from [1]. In the present study, we focus on decision characteristics. Specifically, we investigate how decision characteristics (e.g., number of alternatives) influence the decision-maker's instantaneous cognitive load. Furthermore, we investigate how decision characteristics interact (e.g., number of alternative and information sorting) in influencing cognitive load, which has not been investigated thus far to the best of our knowledge.

This research contributes to decision-making literature by investigating the interactions among decision characteristics and their influence on cognitive load. By gaining a better understanding of how these characteristics influence consumers' cognitive load, this research also has implications for practice, especially for online decision-making tools such as comparison matrices offered on many commercial websites.

## 2 Literature Review

The decision-making literature suggests that a decision is contingent on many characteristics (for a review, see [1]). Some characteristics are related to the social context of the decision. For instance, Bearden and Etzel [2] show that reference groups have varying influence on consumers' decision for public versus private and luxury versus necessity products. Other characteristics are related to the decision maker. For instance, the decision maker's prior knowledge and expertise may influence the way she processes the information [3]. Finally, some characteristics are related to the decision itself. For instance, traditional consumer information load research investigated how the number of alternatives and the number of attributes per alternative influence consumer decision-making [4].

Prior research has identified many characteristics affecting the decision outcome (for a review, see [1]), but limited research investigated interactions between the decision characteristics. Most research on decision characteristic interactions focused on the joint effect of the number of alternatives and attributes on the decision [5]. Considering this research deficit, the objective of the present study is to investigate a wider range of decision characteristic interactions. Also, to the best of our knowledge no research has investigated these interaction effects on the decision maker's actual (i.e., instantaneous) cognitive load during the decision-making process.

In this research, we focus on four characteristics: Decision set size, Attribute value format, Display format, and Information sorting. The decision set size, the number of alternatives and/or the number of attributes per alternative, has been



extensively studied. Results show that as the decision set size increases, the decision quality decreases (for a review, see [6]). The format in which attribute values are presented also influences the decision-making process. Prior research has compared numerical versus textual information formats [7] and has also compared simple versus complex numerical attribute formats [8] on decision outcomes. The display format of the information also plays a role in the decision-making process. In his classic study, Russo [9] shows that when information about unit price is presented in a convenient manner to consumers, they tend to spend less. Thus, information not only needs to be available, it needs to be processable. In this research, we investigate the display format (matrix vs. text) and the sorting of the information within these formats (sorted vs. not sorted).

Cognitive Load Theory (CLT) was first proposed in the field of education in the 1980s [10] and is now used in a variety of fields including human–computer interaction [11]. It postulates that humans have a limited working memory capacity which interacts with a virtually unlimited long-term memory capacity [10]. According to CLT, three types of cognitive load exist: (1) intrinsic load (interaction between task and one’s expertise), (2) extraneous load (additional load due to poor instructions), and (3) germane load (related to processes involved in encoding and retrieving elements in or from long-term memory). CLT suggests that these cognitive load types are additive [12] and should stay within working memory limits in order to avoid information overload and its adverse consequences. Thus, as task demand increases, the sum of cognitive load types (represented by one’s instantaneous cognitive load, i.e., cognitive load that fluctuates each moment someone works on a task) will eventually reach one’s cognitive capacity limit. In this research, we thus influence intrinsic load by manipulating various decision characteristics (Decision set size, Attribute value format, Display format, and Information sorting).

## 3 Method

### 3.1 *Experimental Design*

In order to test the interactions between decision characteristics, we used a  $3 \times 2 \times 2$  between-subject experimental design. The first factor is the decision set. It was composed of either 3, 5, or 7 alternatives and 3, 5, or 7 attributes per alternative (i.e.,  $3 \times 3$ ,  $5 \times 5$ , or  $7 \times 7$  decision set). The second factor is attribute value format. In one condition, attribute values were numerical without any unit of measure (values ranging from 1 to 10), while in the other condition, attribute values were presented with their respective unit of measure (\$, stars, etc.). The third factor is display format; the information was either presented in a matrix format or in a text format. Finally, the information was either presented in a sorted fashion (e.g., all brand prices were presented in the first row) or in an unsorted manner (e.g., the first row displayed price information for a brand, quality for another brand, and reputation for the third brand). Figure 1 provides an example of the manipulated factors.

Product A		Product B		Product C	
Price	\$169	Reputation	2.5 stars	Warranty	5 years
Reputation	3 stars	Warranty	4 years	Price	\$139
Warranty	5 years	Price	\$139	Reputation	3 stars

**Fig. 1** Example of an experimental stimulus (decision set size ( $3 \times 3$ ), attribute value format (unit of measure), Display format (matrix), and Information sorting (not sorted))

### 3.2 Sample and Procedure

Twenty-three participants were recruited among the student panel of HEC Montréal. After completing the consent form, participants were asked to perform a set of product choices. For each choice, participants had to process the product information displayed on a computer screen and then had to report their decision on an answering sheet. In total, participants had to perform 72 product choices (3 sets of 24 choices). In total, 1656 product choices were collected. Participants had a maximum of 20 s to perform each choice.

In order to generate one dominant alternative in each choice set, the following algorithm was used. It ensured that one alternative took the best value in more than one half of the attributes.

- Let  $K$  be the number of attributes (and the number of products). In our experiment  $K = 3, 5, \text{ or } 7$ .
- Let  $M = (K + 1)/2$  be the number of attributes where the dominant product has the highest value.
- Let  $A_{i,j}$  be the assigned value of the  $i$ -th attribute for the  $j$ -th product ( $i, j = 1, \dots, K$ ).
- Let  $(V_{i,1}, V_{i,2}, \dots, V_{i,10})$  the set of 10 values that can be taken by the  $i$ -th attribute, these values are ranked and the last value is always the dominant one (for example  $(1/10, 2/10, \dots, 10/10)$  or  $(200 \$, 190 \$, \dots, 100 \$)$ ).
- We randomly select  $D$ , a value from  $\{1, \dots, K\}$  to represent the dominant product.
- We randomly select  $(d_1, \dots, d_M)$  the  $M$  values from  $\{1, \dots, K\}$  to designate the attributes where product D will be the best choice.
- For each attribute  $i = 1, \dots, K$ , we randomly select  $B_i$  from [13] to represent the highest value that can be simulated for this attribute.
- We assign:
  - $A_{d_1,D} = V_{d_1,Bd_1}$
  - $A_{d_2,D} = V_{d_2,Bd_2}$

- ...
- $A_{dM,D} = V_{dM,BdM}$

For each of the selected attributes, the dominant product takes the highest value.

- For all other remaining  $A_{ij}$ 's, we randomly select a value from:
  - $\{V_{i,1}, V_{i,2}, \dots, V_{i,Bi}\}$  when the attribute  $i$  is not one of the selected attribute ( $d_1, \dots, d_M$ ).
  - $\{V_{i,1}, V_{i,2}, \dots, V_{i,Bi-1}\}$  when the attribute  $i$  is one of the selected attribute ( $d_1, \dots, d_M$ ).

### 3.3 Measured Variables and Apparatus

Most consumer research has used explicit measures to assess cognitive load. For instance, Aljukhadar et al. [14] used a self-reported measure of cognitive load to assess the influence of decision set size on consumers' cognitive load. However, only a few studies in consumer and information system research used implicit cognitive load measures (e.g., [15–17]). Because cognitive load may fluctuate rapidly during a decision-making task and that a decision maker may not be self-conscious of her cognitive load at all times, we used an implicit cognitive load measure. For this research, pupil dilation measured with an eye-tracker (SMI, Teltow, Germany) was used to assess cognitive load [18, 19].

Two additional variables were used as control variables. The first control variable is the decision quality. In order to assess decision quality, the selection of the dominant alternative (i.e., one alternative that was better on at least one attribute and at least equal on all other attributes) was assessed. Second, although participants had a time limit to make their decision, decision time was also used as a control variable.

## 4 Results

Table 1 presents the average cognitive load participants experienced in the twelve different experimental conditions.

In order to test the interactions among the four experimental factors, a multivariate regression model on the average pupil dilation per subject and per interface was used. The fixed effects were all 4 main effects and all 6 two-way interactions. In addition, the decision time and a dummy variable indicating if the participant correctly identified the dominant product (Decision quality) were added as control variables. In order to account for potential correlations between repeated measures, a random (Gaussian) effect was added for each participant. The decision time was

**Table 1** Average cognitive load per experimental condition (mm)

		Unit of measure		No unit of measure	
		Sorted	Not sorted	Sorted	Not sorted
3 × 3	Matrix	3.26	3.30	3.19	3.21
	Text	3.29	3.32	3.24	3.19
5 × 5	Matrix	3.32	3.40	3.26	3.19
	Text	3.36	3.32	3.17	3.20
7 × 7	Matrix	3.40	3.45	3.26	3.26
	Text	3.38	3.47	3.24	3.27

not significant, but the decision quality was found to be a significant control variable ( $F = 1.13, p < 0.04$ ). Interactions results are presented in Table 2.

Results suggest that 4 out of 6 double interactions were significant. The Decision set size \* Attribute value format interaction was significant ( $F = 23.37, p < 0.0001$ ). In the condition where participants saw the units of measure of the attributes, the cognitive load increased as the decision set got larger. However, when the unit of measure was not present, cognitive load only showed an increase when the decision set was the largest.

The Decision set size \* Display format interaction was also significant ( $F = 6.14, p = 0.0022$ ). In the condition where participants were exposed to text format, the cognitive load increased as the decision set got larger. However, in the matrix condition, cognitive load only showed an increase when the decision set was the largest. Thus, only up to a certain point the matrix format helped process information (Fig. 2).

The Decision set size \* Information sorting was also significant ( $F = 4.28, p = 0.0141$ ). In smaller decision sets, the information sorting had no effect on

**Table 2** Main effects and interactions results

Effect	Num DF	Den DF	F value	p value
Decision set size (3, 5, 7)	2	1594	87.83	<0.0001
Attribute value format (unit of measure)	1	1594	517.65	<0.0001
Decision set size * attribute value format	2	1594	23.37	<0.0001
Display format (matrix or text)	1	1594	0.05	0.8304
Decision set size * display format	2	1594	6.14	0.0022
Attribute value format * display format	1	1594	1.16	0.2810
Information sorting	1	1594	11.69	0.0006
Decision set size * information sorting	2	1594	4.28	0.0141
Attribute value format * information sorting	1	1594	25.64	<0.0001
Display format * information sorting	1	1594	0.35	0.5515
Decision quality	1	1594	4.12	0.0426
Decision times	1	1594	1.13	0.2889

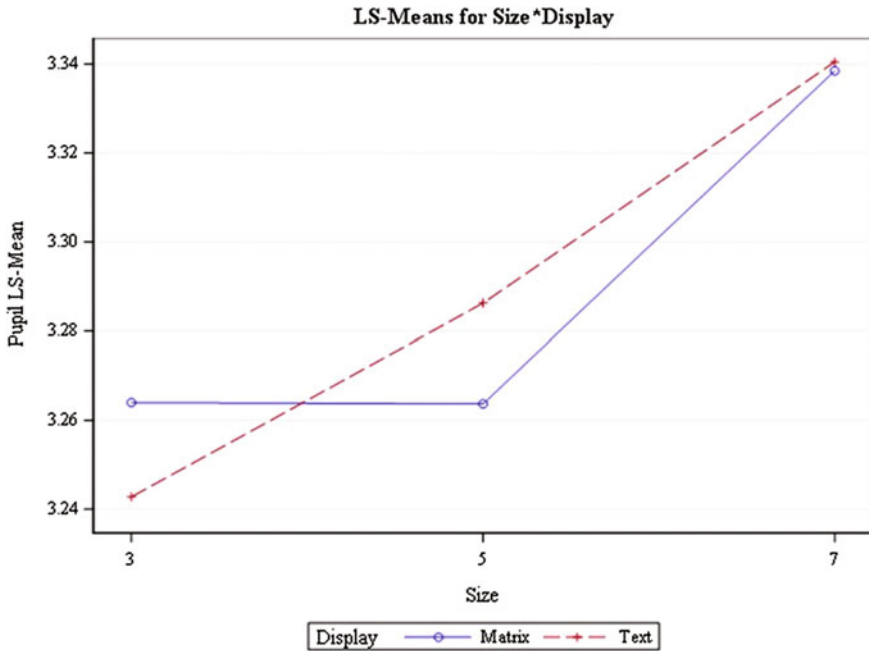


Fig. 2 Decision set size \* display format interaction

cognitive load. However, in the largest decision set, the cognitive load was greater for participants exposed to information not sorted.

Finally, the Attribute value format \* Information sorting interaction was significant ( $F = 25.64, p < 0.0001$ ). In the condition where participants were exposed to attributes' unit of measure, there was no difference between sorted or not sorted information. However, when participants were exposed to attributes with no unit of measure, participants had a greater cognitive load when exposed to information that was not sorted.

## 5 Discussion

Results suggest that decision characteristics influence decision makers' actual cognitive load. More importantly, they suggest that decision characteristics interact to influence instantaneous cognitive load. For instance, as the decision set size increases, the attribute value format, display format, and information sorting effects vary.

Prior research has mostly investigated the interaction between elements of the decision set (number of alternatives and number of attributes) [5]. Our findings

suggest that decision set size interacts with other decision characteristics in affecting the decision maker's cognitive load.

Our findings have implications for commercial website managers who want to present their products and services in a way that is easy for their customers to process. For instance, our results suggest that larger decision sets (i.e., 7 alternatives with 7 attributes) need to be presented with the attributes' respective unit of measure and sorted, but not necessarily in a matrix format to be easier to process. Intuitively, smaller decision sets (e.g., 3 alternatives with 3 attributes) are less sensitive to other decision characteristics, because they remain easy to process in any format.

As with any research endeavor, our research has limitations. First, we limited our investigation to four decision characteristics, but the impact of many additional characteristics could be investigated [1]. Second, in this research we assessed cognitive load during decision-making processes, but did not focus on participants' use of various decision-making strategies [20]. Path analysis using eye-tracking data [21] in conjunction with cognitive load data could be used to help better understand the interplay between cognitive demand across and within different decision-making strategies.

## References

1. Payne, J.W., Bettman, J.R., Johnson, E.J.: *The Adaptive Decision Maker*. Cambridge University Press (1993)
2. Bearden, W.O., Etzel, M.J.: Reference group influence on product and brand purchase decisions. *J. Consum. Res.* **9**(2), 183–194 (1982)
3. Alba, J.W., Hutchinson, J.W.: Dimensions of consumer expertise. *J. Consum. Res.* **13**(4), 411–454 (1987)
4. Malhotra, N.K.: Information load and consumer decision making. *J. Consum. Res.* **8**(4), 419–430 (1982)
5. Lee, B.K., Lee, W.N.: The effect of information overload on consumer choice quality in an on-line environment. *Psychol. Mark.* **21**(3), 159–183 (2004)
6. Eppler, M.J., Mengis, J.: The concept of information overload: a review of literature from organization science, accounting, marketing, MIS, and related disciplines. *Inf. Soc.* **20**(5), 325–344 (2004)
7. Stone, D.N., Schkade, D.A.: Numeric and linguistic information representation in multiattribute choice. *Organ. Behav. Hum. Decis. Process.* **49**(1), 42–59 (1991)
8. Johnson, E.J., Payne, J.W., Bettman, J.R.: Information displays and preference reversals. *Organ. Behav. Hum. Decis. Process.* **42**(1), 1–21 (1988)
9. Russo, J.E.: The value of unit price information. *J. Mark. Res.*, pp. 193–201 (1977)
10. Sweller, J., Van Merriënboer, J.J., Paas, F.G.: Cognitive architecture and instructional design. *Educ. Psychol. Rev.* **10**(3), 251–296 (1998)
11. Hollender, N., et al.: Integrating cognitive load theory and concepts of human–computer interaction. *Comput. Hum. Behav.* **26**(6), 1278–1288 (2010)
12. Paas, F., et al.: Cognitive load measurement as a means to advance cognitive load theory. *Educ. Psychol.* **38**(1), 63–71 (2003)

13. Smeesters, D., Mussweiler, T., Mandel, N.: The effects of thin and heavy media images on overweight and underweight consumers: social comparison processes and behavioral implications. *J. Consum. Res.* **36**(6), 930–949 (2009)
14. Aljukhadar, M., Senecal, S., Daoust, C.-E.: Using Recommendation Agents to Cope with Information Overload. *Int. J. Electron. Commer.* **17**(2), 41–70 (2012)
15. Sénécal, S., et al.: Consumers' cognitive lock-in on websites: evidence from a neurophysiological study. *J. Internet Commer.* **14**(3), 277–293 (2015)
16. Ortiz de Guinea, A., Titah, R., Léger, P.-M.: Explicit and implicit antecedents of users' behavioral beliefs in information systems: a neuropsychological investigation. *J. Manage. Inf. Syst.* **30**(4), 179–210 (2014)
17. Mirhoseini, S.M.M., Léger, P.-M., Sénécal, S.: The Influence of Task Characteristics on Multiple Objective and Subjective Cognitive Load Measures, in *Information Systems and Neuroscience*, pp. 149–156. Springer, Berlin (2017)
18. Laeng, B., Sirois, S., Gredebäck, G.: Pupillometry a window to the preconscious? *Perspect. Psychol. Sci.* **7**(1), 18–27 (2012)
19. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS*. Springer, Berlin (2016)
20. Riedl, R., Brandstätter, E., Roithmayr, F.: Identifying decision strategies: a process-and outcome-based classification method. *Behav. Res. Methods* **40**(3), 795–807 (2008)
21. Pfeiffer, J., et al.: On the influence of context-based complexity on information search patterns: An individual perspective. *J. Neurosci. Psychol. Econom.* **7**(2), 103 (2014)

# Why and How to Design Complementary NeuroIS and Behavioral Experiments

Anthony Vance, Jeffrey L. Jenkins, Bonnie Brinton Anderson,  
C. Brock Kirwan and Daniel Bjornn

**Abstract** Neurophysiological methods offer insights into human cognition that cannot be obtained using traditional methods. However, they are often limited by the artificiality of an experimental setting or the intrusiveness of the method. For these reasons, it is often advisable to complement a NeuroIS experiment with a behavioral experiment, either in a laboratory or field setting. The purpose of this paper is to discuss four guidelines for why and how to effectively design complementary behavioral and NeuroIS experiments. These include: (1) extend NeuroIS experiments with behavioral experiments using theory, rather than replicate; (2) select a behavioral study to enhance ecological and external validity; (3) use the results of each methodology to inform the other; and (4) use NeuroIS and behavioral studies in tandem to inform IT artifact design. By applying these points, researchers can more effectively design complementary NeuroIS and behavioral experiments that together provide richer insights into phenomena under study.

**Keywords** NeuroIS · Behavioral experiments · Ecological validity  
External validity · IT artifact design

---

A. Vance (✉) · J.L. Jenkins · B.B. Anderson · C.B. Kirwan · D. Bjornn  
Brigham Young University, Provo, UT, USA  
e-mail: anthony.vance@byu.edu

J.L. Jenkins  
e-mail: jeffrey\_jenkins@byu.edu

B.B. Anderson  
e-mail: bonnie\_anderson@byu.edu

C. Brock Kirwan  
e-mail: kirwan@byu.edu

D. Bjornn  
e-mail: dbjornn@byu.edu



## 1 Introduction

Neurophysiological methods offer insights into human cognition that cannot be obtained using traditional methods. However, these insights can be limited by the artificiality of an experimental task or setting, or the intrusiveness of the method. For these reasons, it is often advisable to complement a NeuroIS experiment with a behavioral experiment, either in a laboratory or field setting.

The purpose of this paper is to discuss why and how to effectively design a behavioral experiment to complement a NeuroIS experiment. As Dimoka et al. explain, multiple methods can triangulate results to obtain greater certainty [1]. Similarly, Tams et al. argue that multiple methods provide a more holistic view of constructs under investigation [2]. We add to these reasons by arguing that complementary behavioral studies can also: (1) enhance ecological and external validity, (2) observe how neural processes and phenomena are related to behavioral change, and (3) evaluate designs of IT artifacts based on NeuroIS findings.

In this paper, we discuss four guidelines for designing behavioral experiments to complement NeuroIS experiments:

1. Extend NeuroIS experiments with behavioral experiments using theory, rather than simply replicate.
2. Select a behavioral study to enhance ecological and external validity.
3. Use the results of each methodology to inform the other.
4. Use the NeuroIS and behavioral studies in tandem to inform IT artifact design.

For each of the above guidelines, we offer practical insights from our experience in conducting four studies that combined NeuroIS and behavioral experiments. By applying these points, NeuroIS researchers can more effectively design complementary NeuroIS and behavioral experiments that together provide greater confidence and richer insights into phenomena under study.

## 2 Reasons for Designing a Complementary Behavioral Study

### 2.1 *Provide Ecological and External Validity*

Often NeuroIS studies face challenges with ecological validity. Riedl et al. [3] identify three dimensions of intrusiveness: degree of invasiveness, degree of natural position, and degree of movement. For example, the fMRI methodology requires users to lie still in a supine position while being scanned. Realistic interaction with a computer is limited. Other neurophysiological tools may involve intrusive head gear or attachments to the face or limbs.

Similarly, the stimuli associated with various NeuroIS tools, and requirements for stimuli repetition, may reduce the external validity. For example, participants may

see screenshots of computer interactions rather than interacting with the computer. Dimoka et al. [1, pp. 682, 695] therefore recommend to “replicate [NeuroIS experiments] in a more traditional setting and compare the corresponding behavioral responses to test for external validity,” and that “the richness provided by multiple sources of measures can be used to enhance the ecological validity of IS studies”.

## ***2.2 Observe Whether Neural Processes Result in Behavioral Change***

Recently, there has been a renewed interest in the neuroscience literature in incorporating behavioral testing. In a recent commentary, Krakauer et al. [4] argued that simply understanding the parts of the brain involved in a certain behavior is not enough to explain that behavior. These authors quote the early computational neuroscientist David Marr in saying that “trying to understand perception by understanding neurons is like trying to understand a bird’s flight by studying only feathers. It just cannot be done” [5, p. 27]. We echo this sentiment here. Often, the ultimate aim of a NeuroIS study is a deeper understanding of human behavior in an IS context. While neurophysiology measures give more information about research participants that would otherwise be unobservable, they do not necessarily provide a complete explanation of participants’ behavior. Krakauer and colleagues cite several examples of neuroscience findings that were later revealed to be incomplete or incorrect without the corresponding naturalistic behaviors, including bradykinesia in Parkinson Disease [6], sound localization in barn owls [7], and electrolocation in weakly electric fish [8]. Here, we propose that the same principle holds in NeuroIS—neuroscience techniques in the absence of behavioral data may yield incomplete explanations of phenomenon.

## ***2.3 Evaluate the Design of IT Artifacts***

Together, NeuroIS and behavior studies offer powerful validation of IT artifact design. Using NeuroIS can help objectively explain *how* the design of an artifact influences the user’s neurology and thereby *why* an artifact may influence decisions and behaviors [9]. For example, in Jenkins et al. [10] we found that neural activation in the medial temporal lobe (MTL)—a brain region associated with declarative memory—is substantially reduced when showing computer warnings in the middle of other computing tasks as measured by the fMRI. Based on these results, we were able to empirically support that dual-task interference impacts people’s ability to process warnings, which explains why they disregard warnings during computing tasks. We also found that neural activation increased when showing the warning between computing tasks.

Behavioral studies can also evaluate the impact of the IT artifact design on relevant real-world outcomes. In the same paper [10], we conducted a complementary behavioral experiment in a realistic setting using the Chrome Cleanup Tool security message with over 1000 participants. We found that finessing the timing of warnings increased adherence to security messages by over 500%. By conducting both an fMRI and a behavioral field experiment, we were able to conclude (1) why the timing of warnings influences user behaviors, and (2) that the timing of warning had a substantial impact on warning adherence behavior in real-life settings.

### **3 How to Design a Complementary Behavioral Study**

#### ***3.1 Guideline 1: Extend NeuroIS Experiments with Behavioral Experiments Using Theory, Rather Than Simply Replicate***

There is clear value in replicating an experimental design for a variety of reasons [11]. Specifically for NeuroIS studies, Dimoka et al. argue that “no single neurophysiological measure is usually sufficient on its own, and it is advisable to use many data sources to triangulate across measures” [1, p. 694]. Further, because neurophysiological methods are limited along the dimensions of freedom of movement, naturalness of position, and invasiveness [3], a behavioral experiment that closely replicates a NeuroIS experiment will share these same limitations. This can be a missed opportunity to gain new insights that complement those already gained via a NeuroIS experiment.

Rather than pure replication, we recommend designing a behavioral experiment that extends a NeuroIS experiment in terms of method, context (e.g., using a more naturalistic task, setting, or mode of interaction), or both [12]. This can result in a behavioral experiment that is substantially different from its NeuroIS counterpart. In these cases, it is especially important that both experiments be linked by theory, such as the same theoretical explanation and related hypotheses. In this way, NeuroIS experiments can be augmented by testing their findings in realistic contexts and with larger sample sizes.

For example, in Anderson et al. [13] we conducted a fMRI experiment that examined how users habituate to security warnings. We also tested whether users habituate to polymorphic warnings—that is, warnings that change their appearance. Although the fMRI hypotheses were strongly supported, ecological validity was limited because of the invasiveness of the fMRI method, as well as the unnatural mode of interaction and the large number of warnings participants viewed, as compared to everyday life. To enhance the ecological validity of the study overall, we designed a behavioral laboratory experiment in which participants conducted a realistic task on their own laptops. However, the underlying theory was the same, and we tested a subset of the fMRI hypotheses via mouse cursor tracking, which unobtrusively measured attention.

### ***3.2 Guideline 2: Select a Behavioral Study to Enhance Ecological and External Validity***

The ideal complementary experiment is one in which the researchers observe behavior. While surveys are helpful in gathering participant perceptions and intentions, a behavioral study allows researchers to pair how people behave with the insights of *why* they behave that way as measured in the NeuroIS study.

In one project [14], we conducted a series of experiments that incorporated this guideline. At one stage, the participants completed the Iowa Gambling Task while we recorded EEG data. We were able to determine each participant's risk profile based on the profile of their neurophysiological reactions to penalties and rewards. During the next stage, in a new room and without the EEG net, participants completed an image classification computer task on their own laptops. The behavior in question for this study was related to computer security, which was unknown to the participants. We tracked their behavior in response to security messages that were displayed during the course of the primary task. We were able to pair the data collected during the EEG part of the experiment with the data from the image classification task. In this way, we were able to improve the ecological validity of the study overall and demonstrate why people did not behave the way they said they would. Rather, they behaved consistent with the pattern established in the EEG study.

### ***3.3 Guideline 3: Use the Results of Each Methodology to Inform the Other***

Krakauer et al. [4] note that behavioral experiments are often needed either before or after conducting the neural experiments in order to close a mutually beneficial "knowledge loop." In that way, the behavior under investigation can be better defined through pilot or preliminary testing. Similarly, behavioral testing can be informed by the results of the neural data. In our previous studies, we have used the results of both fMRI [13] and EEG [14] experiments to inform behavioral experiments.

For example, in a study on users' risk perception [14], we measured an implicit reaction to risk using EEG which predicted subsequent responses to risk in a computing setting. We also collected behavioral measures of participants' perceptions of risk in a way that participants did not obviously associate with the main experiment. In this way, we were able to compare both behavioral and neural measures of risk perception and use each to inform the other.

### ***3.4 Guideline 4: Use the NeuroIS and Behavioral Studies Together to Inform IT Artifact Design***

NeuroIS and behavioral studies can provide a more comprehensive evaluation of IT artifact design than either method alone. NeuroIS studies can be used to evaluate the design of IT artifacts at a level that may not be available through perceptual methods. Dimoka et al. [15, p. 700] suggested “the brain areas associated with the desired effects can be used as an objective dependent variable in which the IT artifacts will be designed to affect”. Measuring neural data can provide insight into the precursors to behavior, which help explain why a behavior occurs. In contrast, behavioral studies can be used to evaluate the influence of IT artifact design on relevant real-world outcomes. When evaluating an IT artifact design across NeuroIS and behavioral studies, one should strive to ensure the design manipulation is similar across studies.

For example, in Jenkins et al. [10], the timing of warnings was manipulated: whether a warning was shown between versus in the middle of a task. They found strong evidence that dual-task interference (DTI) was induced in the brain when a security message interrupted another task. They also found that DTI was reduced when a security message followed immediately after another task. This manipulation was consistent in both the behavioral and NeuroIS studies. This finding was used to inform the design of security message in Google Chrome. We designed the message to display at low- and high-DTI times, and found displaying the message at low-DTI times substantially improved behavior. By conducting fMRI and behavioral experiments together, we were able to incorporate the neural insights from an fMRI experiment into an IT artifact design that we tested in the field, leading to greater support for the IT artifact design.

## **4 Conclusion**

This paper discusses guidelines for how to design complementary NeuroIS and behavioral experiments in a single study. Although there are many benefits for this approach, we stress that we do not intend for this to become a standard in NeuroIS research. For example, NeuroIS designs can be unobtrusive and provide good ecological validity. In other cases, the research question may be amply addressed using a NeuroIS experiment alone. However, when appropriate, combining NeuroIS and behavioral experiments can yield insights that may be unobtainable using either approach alone.

**Acknowledgements** This research was funded by NSF Grant #CNS-1422831 and a Google Faculty Research Award.

## References

1. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., Lschebeck, A., Kenning, P.H., Pavlou, P.A., Muller-Putz, G., Riedl, R., Vom Brocke J., Weber, B.: On the use of neurophysiological tools in research: developing a research agenda for NeuroIS. *MIS Quart.* **36**(3), 679–702 (2012)
2. Tams, S., Hill, K., Ortiz De Guinea, A., Thatcher, J., Grover, V.: NeuroIS—alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *J. Assoc. Inf. Syst.* **15**(10), 1 (2014)
3. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**(10), i–xxxv (2014)
4. Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., Maciver, M.A., Poeppel, D.: Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93**(3), 480–490 (2017)
5. Marr, D.: *Vision: A Computational Approach*. MIT Press, Boston (1982)
6. Mazzoni, P., Hristova, A., Krakauer, J.W.: Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J. Neurosci.* **27**, 7105–7116 (2007)
7. Grothe, B.: New roles for synaptic inhibition in sound localization. *Nat. Rev. Neurosci.* **4**, 540–550 (2003)
8. Fortune, E.S.: The decoding of electrosensory systems. *Curr. Opin. Neurobiol.* **16**, 474–480
9. Vom Brocke, J., Riedl, R., Léger, P.-M.: Application strategies for neuroscience in information systems design science research. *J. Comput. Inf. Syst.* **53**(3), 1–13 (2013)
10. Jenkins, J.L., Anderson, B.B., Vance, A., Kirwan, C.B., Eargle, D.: More harm than good? How messages that interrupt can make us vulnerable. *Inf. Syst. Res.* **27**(4), 880–896 (2016)
11. Dennis, A.R., Valacich, J.S.: A replication manifesto. *AIS Trans. Replication Res.* **1**, Article 1 (2015)
12. Berthon, P., Pitt, L., Ewing, M., Carr, C.L.: Potential research space in MIS: a framework for envisioning and evaluating research replication, extension, and generation. *Inf. Syst. Res.* **13**(4), 416–427 (2002)
13. Anderson, B.B., Vance, A., Kirwan, C.B., Jenkins, J., Eargle, D.: From warnings to wallpaper: why the brain habituates to security warnings and what can be done about it. *J. Manage. Inf. Syst.* **33**(3), 713–743 (2016)
14. Vance, A., Anderson, B.B., Kirwan, C.B., Eargle, D.: Using measures of risk perception to predict information security behavior: insights from electroencephalography (EEG). *J. Assoc. Inf. Syst.* **15**(10), 679–722 (2014)
15. Dimoka, A., Pavlou, P.A., Davis, F.D.: Research commentary—NeuroIS: the potential of cognitive neuroscience for information systems research. *Inf. Syst. Res.* **22**(4), 687–702 (2011)

# The Impact of Age and Cognitive Style on E-Commerce Decisions: The Role of Cognitive Bias Susceptibility

Nour El Shamy and Khaled Hassanein

**Abstract** The aging associated declines in cognitive abilities could render older adults more susceptible to cognitive biases that are detrimental to their e-commerce decisions' quality. Additionally, certain cognitive styles can lead online consumers to rely on decision heuristics which makes them less meticulous and more prone to bias. In this research-in-progress paper we introduce cognitive bias susceptibility as a potential mediator between age and cognitive style on one end, and decisional outcomes on the other. An experimental design to validate our proposed model is outlined. Both psychometric and eye-tracking methodologies are utilized to achieve a more holistic understanding of the relationships in the proposed model. Potential contributions and implications for future research are outlined.

**Keywords** Aging · Older adults · Cognitive style · Cognitive bias  
Order bias · Vividness bias · Eye-tracking · Decision quality · Decision effort

## 1 Introduction

In a world driven by digital transformation, the importance of online shopping to retail consumers is ever growing. Unfortunately, not all consumers can benefit equally from e-commerce. Older adults, those who are 60 years or older, are the fastest growing population segment, both globally and in North America [1, 2]. Additionally, they are the fastest growing segment of Internet users [3] and the largest user group in North America, comprising 26.8% of Internet users in 2014 [4]. Older adults are generally more affluent and therefore lucrative for vendors as a consumer segment [5], and many features of e-commerce (e.g., convenience, lack of physical barriers) can be particularly useful to them [6]. Nonetheless, they suffer

---

N. El Shamy (✉) · K. Hassanein  
DeGroote School of Business, McMaster University, Hamilton, Canada  
e-mail: elshamyn@mcmaster.ca

K. Hassanein  
e-mail: hassank@mcmaster.ca

from a decline in several physiological abilities such as sensorimotor skills and useful field of vision (UFOV) [7–9]; as well as fluid cognitive abilities such as selective attention and working memory [10, 11]; which prevent them from reaping the full utility of information technologies and can be taxing to the quality of their e-commerce decisions.

E-commerce decision making is generally a complex process for all consumers regardless of age. Online consumers have access to virtually unlimited choices and information [12] that exceeds their cognitive capacity. As a result, consumers resort to suboptimal decision-making strategies such as satisficing [13] and heuristics [14]. By trading accuracy, at different individual quality thresholds [15], to conserve cognitive effort [16], decision makers render themselves prone to systematic decision biases [14, 17]. The degree to which individuals predominantly approach their decisions by either relying on heuristics and intuition on one extreme, or meticulous attention and consideration of evidence on the other; corresponds to their cognitive style [18–21], which impacts their information seeking behavior during e-commerce tasks [22]. This effect can be further exacerbated by the diminishing cognitive abilities that are associated with aging. The roles of these two individual difference factors (i.e., age, cognitive style) have not been rigorously investigated in Information Systems (IS), and there have been recent calls to investigate the impact of age [23] and cognitive style [22] on IS phenomena including online decision behavior.

This research attempts to address this gap by examining how cognitive style and the decline in selective attention that is associated with aging manifest in e-commerce decisions. These individual difference factors are expected to make decision makers more susceptible to certain cognitive biases that are elicited by the way information is presented. Specifically, we investigate the influence of two decision biases (i.e., order bias, and vividness bias) that pertain to the interplay between attention and information presentation [24–26], and are thereby particularly relevant for older adults' decision making in e-commerce. Building on the theory of cognitive biases [14, 27] and the effort/accuracy framework [16, 28], this study introduces the construct “cognitive bias susceptibility” as a potential mediator between age and cognitive style on one end, and decision quality and cognitive effort on the other. We utilize state-of-the-art eye-tracking technology to tap into the decision-makers' mental processes [29–31] and develop an objective measure [29, 31, 32] of bias susceptibility.

## 2 Theoretical Development and Research Model

### 2.1 *Decision Making and Cognitive Biases*

Research on decision making and decision support in e-commerce has been dominated [17, 33] by the theory of bounded rationality [13] and the effort/accuracy



framework [16]. These paradigms acknowledge individuals' cognitive limitations and their behavioral tendencies to conserve cognitive effort during decision making; being "cognitive misers" [34]. The implication of the effort/accuracy framework is that decision makers are not only concerned about their decision quality, but are also concerned about their perceived decision effort [28]. Another implication, from Cognitive Fit theory [35], is that a better fit between task, technology, and user conserves decision effort, which can be then reallocated towards enhancing accuracy [17].

**Cognitive Biases.** Cognitive biases are inherent and systematic prejudices that influence decision-makers' behaviors and reduce the quality of their decisions [24, 27]. They can manifest in different decision and cognitive processes and have been classified in the literature in different ways [24, 25]. Some of these classifications include "perceptual" [24] or "presentation" [27] biases, which relate to the information presentation format that influence the attentional process of the decision maker. The order and vividness of presented information force a bottom-up stimulus-driven attention capture and working memory encoding, interfering with top-down goal-driven attention control [26].

An order bias is the tendency of decision makers to gravitate towards, and assign more value to, information that is presented earlier in a set as a result of declining attention [6, 27, 36]. Evidence of this bias has been reported in e-commerce, impacting consumers formulation of vendor appraisal based on the order of which other users' ratings of a vendor is presented [37]. The attentional drift diffusion model (aDDM) posits that as decision makers gaze and shift their attention between alternatives, they accumulate evidence in their favor, and generally a "bias exists in favor of alternatives fixated on first because they have accumulated more evidence" [26]. Additionally, eye-tracking research has consistently shown that individuals tend to look more towards the top and left sides of the screen when browsing [26, 38], particularly in North America where the major official spoken languages are English, French, and Spanish, which all follow the same orthography.

A vividness bias is the tendency of decision makers to gravitate towards salient and visually stimulating alternatives because they attract more attention and are easier to recall [6, 24, 26, 27]. Theories of image saliency are established in cognitive psychology and consumer behavior [26, 39], and measures of saliency (e.g., contrast, color against background) are used extensively in advertising, including in e-commerce [40, 41]. The richness of a vivid alternative stimulates and drives visual attention in its favor, making it more likely to be chosen or weighted higher relative to others, and increasing working memory load has been found to increase this effect [26].

**Cognitive Bias Susceptibility.** We define cognitive bias susceptibility as the likelihood of an individual falling prey to one or more cognitive biases that could be detrimental to the quality of a decision they are making in a certain context. Given that cognitive biases are subtle cognitive prejudices that likely occur subconsciously, decision makers may not be aware of their susceptibility to these biases. They may also fail to recall their decision process retrospectively. Additionally, individuals might not accurately reflect their own bias susceptibility due to social

desirability bias [32] or bias blind spot [42], which influence individuals' assessment and self-reporting of their own susceptibility. This makes it difficult for researchers to study the mechanisms through which such biases affect decision making.

NeuroIS measures, such as eye-tracking, have been generally encouraged specifically for constructs that are amenable to subtle or unconscious cognitive or physiological processes (e.g., attention, stress, anxiety) [32, 43, 44]. The theory of reading and comprehension provides support that there is a strong relationship between selective attention and working memory, noting that "the eye-mind assumption posits that there is no appreciable lag between what is being fixated and what is being processed" [45]. This notion of eye-mind, enables researchers to trace and make objective inferences about the mental processes of users through eye-tracking methodologies [26, 29, 32]. Thus, eye-tracking can be particularly useful in tracing the cognitive processes of users in real time [30], without the need for the user to stop and think aloud, which might interrupt the natural process of their decision making. Additionally, objective eye gaze behavior is much less prone to self-presentation biases that might influence the user during recall, and is immune to failure to recall [32]. Hence, in this research, an objective measure of cognitive bias susceptibility factoring two ocular metrics will be developed. More specifically, the measure will assess the degree to which a decision maker exhibits asymmetry in attention in favor of bias-inducing decision alternatives, factoring the percentages of total number of fixations and total fixation duration eye gaze metrics.

## 2.2 *Individual Differences in Decision Making*

**Age.** While age has been historically studied in IS, Tams et al. [23] argue that not much is known about its theoretical "touch points" with IS phenomena, and they set out a research agenda calling to further scrutinize aging in IS research. It is well documented that aging is associated with concomitant natural decline in fluid cognitive abilities, such as attention and working memory capacity [8, 46, 47]. Attention is the selective attendance to particular perceptual sensory inputs and disregard of others [26, 48], while working memory is a limited resource capacity of the brain where information required for accomplishing an active task is temporarily stored [48]. These two faculties are imperative to individuals' decision making [26], and impairments in these areas can impact their information seeking behavior [11], bias their decision process, and detriment the quality of their decisions [49].

In goal-driven tasks, such as e-commerce tasks, individuals exert a top-down control of their selective attention, steering their visual focus to the stimuli that are most relevant to their task demands [26]. Age-related differences in information search effectiveness have been attributed to the decline in older adults' selective attention [11] rather than diminishing physical abilities such as UFOV [50]. Additionally, researchers have found similar results of reduced information search

effectiveness by introducing cognitive load in the form of website complexity [31]. Given the complexity of e-commerce environments due to the overabundance of choice and information overload [51], e-commerce decisions can be particularly taxing to older adults.

The aDDM posits that stimuli fixated earlier in a task will be more likely encoded in working memory than others. The order and vividness biases force a bottom-up attentional drift in favor of primal and salient alternatives [26]. Previous studies have demonstrated that by reducing the working memory capacity of decision makers, through working memory overload interventions, their information seeking behavior becomes impaired, driving them to utilize “fixations as an external memory space, thereby reducing demands on cognitive memory” [26]. Given the decline in fluid cognitive efficacy that is associated with age [8, 23], it is expected that cognitive fatigue for older adults will occur earlier in a given task compared to younger adults, *ceteris paribus*. Alternatives that are least salient and lower in order of presentation will be less attended to, which makes these biases more prominent for older users.

**Cognitive Style.** Cognitive styles are habituated approaches to decision making that individuals predominantly utilize when making decisions in particular contexts [18, 52, 53]. Studies generally conceptualize cognitive styles as two extremes on a satisficer—maximizer continuum [17, 22] building on Simon’s seminal theory of bounded rationality and satisficing [13].

Maximizers are perfectionists who engage in evidence-based decision making, considering all available information as much as possible to carefully assess alternatives and reach an ideal, or near-ideal, decision [21, 22]. Satisficers, on the other hand, are more concerned with the efficiency of their decision-making process, and reduce the associated cognitive cost by using decision heuristics as shortcuts [21, 22, 54]. Karimi et al. [22] found that satisficers tend to spend less time when making online decisions, and consider fewer alternatives and attributes compared to maximizers. While satisficing by utilizing heuristics can be beneficial [55–57] in certain contexts (e.g., tight time constraints), it is generally a suboptimal strategy [56, 57] that is more susceptible to cognitive biases [14, 56, 57].

Individuals on either end of the cognitive style continuum will likely gravitate towards primal and salient alternatives [26]. However, given the satisficers’ tendency to trade accuracy for effort conservation, they will be less inclined to exert effort to examine less salient or lower placed alternatives in a set. Thus, satisficers can be more susceptible to order and vividness biases relative to maximizers.

As users fall prey to the order and vividness biases, they will dwell on primal and salient alternatives more and gather more evidence in their favor, making primal and salient alternatives more likely to be chosen [26]. Salient and primal alternatives might not necessarily be the most rational choice; hence these biases can be detrimental to the quality of users’ decisions. Additionally, given that these biases manifest as a result of heuristics applied by the decision makers, whereas these heuristics are “shortcuts” to reduce cognitive effort [14], we expect that users who are susceptible to cognitive biases will perceive less cognitive effort.

Building on the foregoing discussion, we advance the following hypotheses and capture the relationships between the abovementioned constructs in Fig. 1.

**H1a,b:** Older adults are more susceptible to the (a) order and (b) vividness bias than younger adults.

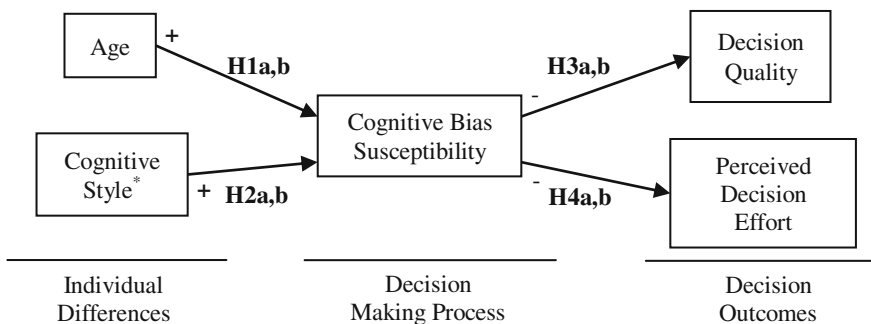
**H2a,b:** Satisficers are more susceptible to the (a) order and (b) vividness bias than maximizers.

**H3a,b:** Individuals susceptible to the (a) order (b) vividness bias will be less likely to select an optimal decision compared to those who aren't.

**H4a,b:** Individuals susceptible to the (a) order (b) vividness bias will experience lower perceptions of cognitive effort compared to those who aren't.

### 3 Methodology

A controlled e-commerce task experiment will be conducted using a  $2 \times 2 \times 2$  factorial design. The first factor, age, will be measured as a dichotomous variable following Tams [48] in which half of the participants will represent young adults (age < 30) and the other half older adults (age > 60). The second factor, cognitive style, will be measured using the maximization scale introduced by Schwartz et al. [21] in which participants will be categorized as either maximizers or satisficers. The interaction between age and cognitive style will be examined post hoc. The third factor (within-subjects, counterbalanced) will comprise a manipulation of the e-commerce task to induce the order and vividness biases. A “featured” icon will be overlaid on the pictures of some non-optimal alternatives to increase their saliency and to induce a vividness bias in one task. The order of the optimal alternative will be fixed near the bottom of the list in the order bias-induced task. To assess decision quality, we utilize the dominated versus non-dominated alternatives binary approach suggested by Häubl and Trifts [58] which provides an objective decision quality measure. A non-dominated alternative is the optimal choice in a set, and is superior to dominated alternatives in all attributes.



**Fig. 1** Proposed Research Model. \*Satisficing cognitive style is used for illustration, Maximizing is expected to have the opposite effect

To ensure the realism and sufficient complexity of the e-commerce tasks in our experiment and the effectiveness of our manipulations, a pilot study will be conducted to identify the optimal number of alternatives, attributes per alternative, and order of the non-dominated alternative in the order bias task. This is to avoid any possible confounding effects as a result of high task and webpage design complexity affecting cognitive load [31, 59]. Additionally, alternatives and attributes will be carefully selected to avoid any emotion-laden features, as these may induce an imaginability bias (a memory category bias [27]) which may contaminate the study results by taxing users' cognitive capacity [49, 60]. Extreme care will be taken to isolate and induce the two biases under investigation separately.

An objective measure for cognitive bias susceptibility will be developed utilizing two eye gaze behavior metrics; (i) the symmetry of the percentages of total number of fixations across alternatives and (ii) the symmetry of the percentages of total gaze durations across alternatives. By factoring these two ocular metrics, our proposed cognitive bias susceptibility construct will reflect the level of asymmetry in attention in favor of bias-inducing alternatives. The developed measure will be validated in our pilot study.

Utilizing state of the art eye-tracking equipment (Tobii Pro X2-60) measuring eye gaze at 60 Hz, we will be able to gain deep insights into users' decision-making processes including ocular behavior, with high temporal and spatial precision, and provide a more holistic understanding of cognitive bias susceptibility [30, 43, 44, 61, 62]. Areas of Interest (AOIs) will be utilized to categorize and analyze oculometric data. Eye gaze process tracing [63], ANOVA, and t-tests will be used to compare the decision outcomes and other oculometric metrics (e.g., gaze, fixations) between groups to test the hypotheses. To detect a medium effect size at a power of 0.8 and  $\alpha$  of 0.05, 30 participants will be required for each cell, thus a total of 120 participants will be recruited through the McMaster Digital Transformation Research Centre as well as the Gilbrea Centre for Studies in Aging which is also located at McMaster University.

## 4 Potential Contributions and Limitations

This research is expected to have important theoretical contributions. First, it will advance our understanding on how the attentional deficits associated with aging impact users' susceptibility to cognitive biases. Second, it will shed light on the understudied cognitive style construct in IS research and its impact on bias susceptibility. Third, it will advance our understanding on the role bias susceptibility plays within the effort/accuracy framework, and how it affects perceived effort and decision quality in e-commerce. In future research, we plan to build on the findings from this study and develop debiasing strategies for the order and vividness biases, which could be incorporated into standard e-commerce decision aids (e.g., recommendation agents).

This research is also expected to have practical implications. Poor consumer decision quality can be significantly costly to practitioners, especially considering the prevalence of free return and exchange policies and associated losses in retail. There was an estimated \$284 billion worth of returned products in 2014, and this problem has been further exacerbated by the proliferation of e-commerce [64]. By understanding how to debias consumers and improve their decision quality, retailers can improve customers' satisfaction with their decisions and reduce avoidable returns.

This research is not without limitations. Several individual difference factors were not investigated despite their potential relevance, (e.g., product knowledge) for pragmatic reasons related to sample size. Further, only two cognitive biases (i.e., order, vividness) are examined, and other biases can be detrimental to e-commerce decisions. These other constructs and biases can be investigated in future studies.

## References

1. Department of Economic and Social Affairs: Concise Report on the World Population Situation in 2014. New York (2014)
2. Department of Economic and Social Affairs: Profiles of Ageing (2015)
3. Wagner, N., Hassanein, K., Head, M.: Computer use by older adults: a multi-disciplinary review. *Comput. Hum. Behav.* **26**, 870–882 (2010)
4. Statista: Distribution of internet users in North America as of November 2014, by age group. North America: age distribution of internet users 2014 (2017)
5. Lian, J.-W., Yen, D.C.: Online shopping drivers and barriers for older adults: age and gender differences. *Comput. Hum. Behav.* **37**, 133–143 (2014)
6. El Shamy, N., Hassanein, K.: The influence of cognitive biases and decision making styles of older adults in e-commerce tasks: an exploratory study. In: Proceedings of the Fourteenth Pre-ICIS SIG-HCI Workshop, Fort Worth, TX (2015)
7. Prieto, T.E., Myklebust, J.B., Hoffmann, R.G., Lovett, E.G., Myklebust, B.M.: Measures of postural steadiness: differences between healthy young and elderly adults. *IEEE Trans. Biomed. Eng.* **43**, 956–966 (1996). doi:[10.1109/10.532130](https://doi.org/10.1109/10.532130)
8. Czaja, Sara J., Charness, Neil, Fisk, Arthur D., Hertzog, Christopher, Nair, Sankaran N., Rogers, Wendy A., Sharit, Joseph: Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychol. Aging* **21**, 333–352 (2006). doi:[10.1055/s-0029-1237430](https://doi.org/10.1055/s-0029-1237430). [Imprinting](#)
9. National Institute on Aging, and National Library of Medicine: Making Your Web Site Senior Friendly (2002)
10. Salthouse, Timothy A., Babcock, Renee L.: Decomposing adult age differences in working memory. *Dev. Psychol.* **27**, 763–776 (1991). doi:[10.1037/0012-1649.27.5.763](https://doi.org/10.1037/0012-1649.27.5.763)
11. Plude, D.J., Doussard-Roosevelt, J.A.: Aging, selective attention, and feature integration. *Psychol. Aging* **4**, 98–105 (1989). doi:[10.1037/0882-7974.4.1.98](https://doi.org/10.1037/0882-7974.4.1.98)
12. Wan, Y., Menon, S., Ramaprasad, A.: The paradoxical nature of electronic decision aids on comparison-shopping: the experiments and analysis. *J. Theor. Appl. Electron. Commer. Res.* **4**, 80–96 (2009). doi:[10.4067/S0718-18762009000300008](https://doi.org/10.4067/S0718-18762009000300008)
13. Simon, H.A.: A behavioral model of rational choice. *Q. J. Econ.* **69**, 99–118 (1955)
14. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. *Science* **185**, 1124–1131 (1974)

15. Chu, P.C., Spires, Eric E.: The joint effects of effort and quality on decision strategy choice with computerized decision aids. *Decis. Sci.* **31**, 259–292 (2000). doi:[10.1111/j.1540-5915.2000.tb01624.x](https://doi.org/10.1111/j.1540-5915.2000.tb01624.x)
16. Johnson, E.J., Payne, J.W.: Effort and accuracy in choice. *Manage. Sci.* **31**, 395–414 (1985). doi:[10.1287/mnsc.31.4.395](https://doi.org/10.1287/mnsc.31.4.395)
17. Davern, M., Shaft, T., Te'eni, D.: Cognition matters: enduring questions in cognitive IS research. *J. Assoc. Inf. Syst.* **13**, 273–314 (2012)
18. Sproles, George B., Kendall, Elizabeth L.: A Methodology for profiling consumers' decision-making styles. *J. Consum. Aff.* **20**, 267–279 (1986)
19. Carlson, John G.: Recent assessments of the Myers-Briggs type indicator. *J. Pers. Assess.* **49**, 356–365 (1985)
20. Barkhi, R.: Cognitive style may mitigate the impact of communication mode. *Inf. Manage.* **39**, 677–688. (2002). doi:[10.1016/S0378-7206\(01\)00114-8](https://doi.org/10.1016/S0378-7206(01)00114-8)
21. Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., Lehman, D.R.: Maximizing versus satisficing: happiness is a matter of choice. *J. Pers. Soc. Psychol.* **83**, 1178–1197 (2002). doi:[10.1037/0022-3514.83.5.1178](https://doi.org/10.1037/0022-3514.83.5.1178)
22. Karimi, S., Papamichail, K.N., Holland, C.P.: The effect of prior knowledge and decision-making style on the online purchase decision-making process: a typology of consumer shopping behaviour. *Decis. Support Syst.* **77**, 137–147 (2015)
23. Tams, S., Grover, V., Thatcher, J.: Modern information technology in an old workforce: toward a strategic research agenda. *J. Strateg. Inf. Syst.* **23**, 284–304. (2014). doi:[10.1016/j.jsis.2014.10.001](https://doi.org/10.1016/j.jsis.2014.10.001)
24. Fleischmann, M., Amirpur, M., Benlian, A., Hess, T.: Cognitive biases in information systems research: a scientometric analysis. In: *European Conference on Information Systems*, pp. 1–21. Tel Aviv (2014)
25. Arnott, D., Pervan, G.: Eight key issues for the decision support systems discipline. *Decis. Support Syst.* **44**, 657–672 (2008). doi:[10.1016/j.dss.2007.09.003](https://doi.org/10.1016/j.dss.2007.09.003)
26. Orquin, J.L., Loose, S.M.: Attention and choice: a review on eye movements in decision making. *Acta Psychol.* **144**, 190–206. (2013). doi:[10.1016/j.actpsy.2013.06.003](https://doi.org/10.1016/j.actpsy.2013.06.003)
27. Arnott, D.: Cognitive biases and decision support systems development: a design science approach. *Inf. Syst. J.* **16**, 55–78 (2006)
28. Todd, P., Benbasat, I.: The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. *MIS Q.* **16**, 373–393 (1992)
29. Duchowski, A.T.: *Eye Tracking Methodology: Theory and Practice*. Vasa. Second edn. (2007). doi:[10.1145/1117309.1117356](https://doi.org/10.1145/1117309.1117356)
30. Glaholt, M.G., Reingold, E.M.: Eye movement monitoring as a process tracing methodology in decision making research. *J. Neurosci. Psychol. Econ.* **4**, 125–146 (2011). doi:[10.1037/a0020692](https://doi.org/10.1037/a0020692)
31. Wang, Q., Yang, S., Liu, M., Cao, Z., Ma, Q.: An eye-tracking study of website complexity from cognitive load perspective. *Decis. Support Syst.* **62**, 1–10. (2014). doi:[10.1016/j.dss.2014.02.007](https://doi.org/10.1016/j.dss.2014.02.007)
32. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., et al.: On the use of neurophysiological tools in IS research: developing a research agenda for NeuroIS. *MIS Q.* **36**, 679–702 (2012)
33. Xiao, B., Benbasat, I.: E-commerce product recommendation agents: use, characteristics, and impact. *MIS Q.* **31**, 137–209 (2007)
34. Wang, W., Benbasat, I.: Interactive decision aids for consumer decision making in e-commerce: the influence of perceived strategy restrictiveness. *MIS Q.* **33**, 293–320. (2009). doi:Article
35. Vessey, I., Galletta, D.: Cognitive fit: an empirical study of information acquisition. *Inf. Syst. Res.* **2**, 63–84 (1991)
36. Yates, J.F., Curley, S.P.: Contingency judgement: primacy effects and attention decrement. *Acta Physiol. (Oxf)* **62**, 293–302 (1986)



37. Xu, Y.C., Kim, H.W.: Order effect and vendor inspection in online comparison shopping. *J. Retail.* **84**, 477–486 (2008). doi:[10.1016/j.jretai.2008.09.007](https://doi.org/10.1016/j.jretai.2008.09.007)
38. Pan, B., Hembrooke, H., Gay, G.K., Granka, L., Feusner, M.K., Newman, J.K.: The determinants of web page viewing behavior: an eye-tracking study. In: Proceedings of the ETRA '04 Symposium on Eye Tracking Research and Applications, vol. 1, pp. 147–154 (2004). doi:[10.1145/968363.968391](https://doi.org/10.1145/968363.968391)
39. Scott, L.M., Vargas, P.: Writing with pictures: toward a unifying theory of consumer response to images. *J. Consum. Res.* **34**, 341–356 (2007). doi:[10.1086/519145](https://doi.org/10.1086/519145)
40. Lim, K.H., Benbasat, I.: The effect of multimedia on perceived equivocality and perceived usefulness of information systems. *Mis Q.* **24**, 449–471 (2000). doi:[10.2307/3250969](https://doi.org/10.2307/3250969)
41. Cyr, D., Head, M., Larios, H., Pan, B.: Exploring human images in website design: a multi-method approach. *MIS Q.* **33**, 539–566 (2009)
42. Pronin, E., Lin, D.Y., Ross, L.: The bias blind spot: perceptions of bias in self versus others. *Pers. Soc. Psychol. Bull.* **28**, 369–381 (2002). doi:[10.1177/0146167202286008](https://doi.org/10.1177/0146167202286008)
43. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**, i–xxxv (2014)
44. Riedl, R., Léger, P.M.: *Fundamentals of NeuroIS. Studies in Neuroscience, Psychology and Behavioral Economics.* Springer, Berlin, Heidelberg (2016)
45. Just, M.A., Carpenter, P.A.: A theory of reading: from eye fixations to comprehension. *Psychol. Rev.* **87**, 329–354 (1980)
46. Ghisletta, P., Rabbitt, P., Lunn, M., Lindenberger, U.: Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. *Intelligence* **40**, 260–268 (2012). doi:[10.1016/j.intell.2012.02.008](https://doi.org/10.1016/j.intell.2012.02.008)
47. Finucane, M.L., Slovic, P., Hibbard, J.H., Peters, E., Mertz, C.K., Macgregor, D.G.: Aging and decision-making competence: an analysis of comprehension and consistency skills in older versus younger adults considering health-plan options. *J. Behav. Decis. Mak.* **15**, 141–164 (2002). doi:[10.1002/bdm.407](https://doi.org/10.1002/bdm.407)
48. Tams, S.: A refined examination of worker age and stress: explaining how, and why, older workers are especially techno-stressed in the interruption age. In: Davis, F.D., Riedl, R., vom Brocke, J., Léger, P.M., Randolph, A.B. (eds.) *Information Systems and Neuroscience*, pp. 175–183 (2017). doi:[10.1007/978-3-319-41402-7\\_22](https://doi.org/10.1007/978-3-319-41402-7_22)
49. Peters, E., Hess, T.M., Västfjäll, D., Auman, C.: Adult age differences in dual information processes: implications for the role of affective and deliberative processes in older adults' decision making. *Perspect. Psychol. Sci. J. Assoc. Psychol. Sci.* **2**, 1–23 (2007). doi:[10.1111/j.1745-6916.2007.00025.x](https://doi.org/10.1111/j.1745-6916.2007.00025.x)
50. Bergstrom, R., Jennifer, C., Olmsted-Hawala, E.L., Jans, M.E.: Age-related differences in eye tracking and usability performance: web site usability for older adults. *Int. J. Hum. Comput. Interact.* **29**, 541–548 (2013). doi:[10.1080/10447318.2012.728493](https://doi.org/10.1080/10447318.2012.728493)
51. Gudigantala, N., Song, J., Jones, D.R.: Transforming consumer decision making in e-commerce: a case for compensatory decision aids. In: Lee (ed.) *Transforming E-Business Practices and Applications: Emerging Technologies and Concepts: Emerging Technologies and Concepts*, pp. 72–88 (2010). doi:[10.4018/978-1-60566-910-6.ch005](https://doi.org/10.4018/978-1-60566-910-6.ch005)
52. Thunholm, P.: Decision-making style: habit, style or both? *Pers. Individ. Differ.* **36**, 931–944 (2004). doi:[10.1016/S0191-8869\(03\)00162-4](https://doi.org/10.1016/S0191-8869(03)00162-4)
53. Allinson, C.W., Hayes, J.: The cognitive style index: a measure of intuition-analysis for organizational research. *J. Manage. Stud.* **33**, 119–135 (1996)
54. Gilovich, T., Griffin, D., Kahneman, D.: *Heuristics and Biases: The Psychology of Intuitive Judgment.* Cambridge University Press, New York (2002)
55. Gigerenzer, G., Todd, P.M., The ABC Research Group: *Simple Heuristics That Make Us Smart.* Oxford University Press, New York (2014)
56. Kahneman, D.: *Thinking, Fast and Slow.* Farrar, Straus and Giroux, New York (2011)
57. Bazerman, M.H., Moore, D.: *Judgment in Managerial Decision Making*, 7th edn. John Wiley & Sons, Inc, USA (2009)



58. Häubl, G., Trifts, V.: Consumer decision making in online shopping environments: the effects of interactive decision aids. *Mark. Sci.* **19**, 4–21 (2000)
59. Buettner, R.: The Relationship between visual website complexity and a user's mental workload: a NeuroIS perspective. In: Davis, F., Riedl, R., vom Brocke, J., Léger, P.M., Randolph, A.B. (eds.) *Information Systems and Neuroscience*, pp. 107–113. Springer, Berlin (2017). doi:[10.1007/978-3-319-41402-7\\_14](https://doi.org/10.1007/978-3-319-41402-7_14)
60. Drolet, A., Luce, M.F.: The rationalizing effects of cognitive load on emotion based tradeoff avoidance. *J. Consum. Res.* **31**, 63–77 (2004). doi:[10.1086/383424](https://doi.org/10.1086/383424)
61. Dimoka, A., Pavlou, P., Davis, F.D.: Neuro IS: the potential of cognitive neuroscience for information systems research. *Inf. Syst. Res.* **22**, 687–702 (2011)
62. Tams, S., Hill, K., Thatcher, J.: Neuro-IS—Alternative or complement to existing methods? illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *J. Assoc. Inf. Syst.* **15**, 723–753 (2014)
63. Mao, J.Y., Benbasat, I.: The use of explanations in knowledge-based systems: cognitive perspectives and a process-tracing analysis. *J. Manage. Inf. Syst.* **17**, 153–179 (2000). doi:[10.1080/07421222.2000.11045646](https://doi.org/10.1080/07421222.2000.11045646)
64. DiChristopher, T.: Your holiday gift returns cost retailers billions. CNBC (2015)

# Expertise as a Mediating Factor in Conceptual Modeling

Christopher J. Davis, Alan R. Hevner, Élise Labonte-LeMoyne  
and Pierre-Majorique Léger

**Abstract** We use eye-tracking to better understand the notion of expertise in conceptual modeling of complex systems. This research-in-progress paper describes an ongoing experiment to exploit the capacity of eye-tracking to explore the significance of expertise as a mediating factor in conceptual modeling. The proposed methodology highlights the applicability, validity, and potential of well-established eye-tracking methods to measure the effects of expertise. By identifying the differences in the strategies that novices and experts use to search, detect, and diagnose errors, we anticipate being able to help define training curricula appropriate for each level to improve performance and model result quality.

**Keywords** Eye-tracking · Conceptual modeling · Expertise

## 1 Introduction

Our research explores differences between experts and novices in the context of conceptual modeling. Conceptual modeling is central to the design process of IT artifacts and is widely acknowledged to be conceptually complex [1, 2]. Critical to the effectiveness and utility of conceptual models are the notations used to present them in forms that can be shared with others. Such notations allow easier sharing of

---

C.J. Davis (✉)

University of South Florida, Saint Petersburg, FL, USA

e-mail: davis@mail.usf.edu

A.R. Hevner

University of South Florida, Tampa, FL, USA

e-mail: ahevner@usf.edu

É. Labonte-LeMoyne · P.-M. Léger

HEC Montréal, Montreal, Québec, Canada

e-mail: elise.labonte-lemoyne@hec.ca

P.-M. Léger

e-mail: pierre-majorique.leger@hec.ca

© Springer International Publishing AG 2018

F.D. Davis et al. (eds.), *Information Systems and Neuroscience*,

Lecture Notes in Information Systems and Organisation 25,

[https://doi.org/10.1007/978-3-319-67431-5\\_10](https://doi.org/10.1007/978-3-319-67431-5_10)

information between collaborators and prevent misunderstandings. They also ensure that models can be retrospectively understood even when the original designer is not present.

In this research-in-progress paper, we use eye-tracking to better understand the notion of expertise in business process modeling. Specifically, using a laboratory experiment, we monitor the visual attention of novices and experts as they search for and identify semantic and syntactic errors in conceptual models. Moody [3] argues that the cognitive effectiveness of visual notations has been under-researched, particularly as regards their contribution to design. Here we consider how expertise relates to cognitive effectiveness and how it might be operationalized in an experimental design. This research extends that of Recker et al. [4] to consider—in the context of process modeling—what constitutes expertise. Brain changes in the development of expertise are important [5]. Our research will extend the range of neurophysiological metrics used to identify and differentiate the skill-based adaptations that differentiate novice and expert designers. This, in turn, will enable better training guidelines for those in the workplace to be developed. It will also provide the basis to offer differentiated modes of teaching for students [4] and to enhance overall curriculum development.

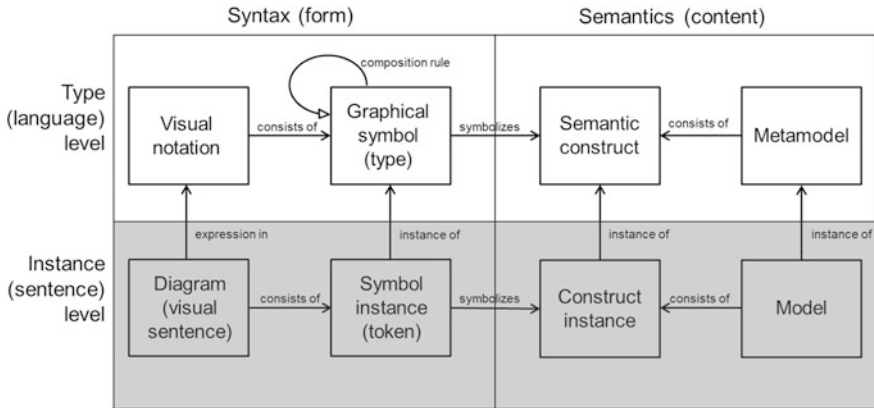
In Sect. 2 we review visual notations and their syntactic and semantic implications for design, as well as the literature on expertise, setting out the motivation for this phase of our research. In Sect. 3 we present a research protocol to answer our research question and detail the pilot study currently under way. Section 4 concludes with a discussion of research directions and implications. We offer some conjectures on the neurophysiological artifacts pertinent to this study and their potential to differentiate novice and expert designers.

## 2 Prior Research

### 2.1 *Conceptual Modeling, Notation, Syntax and Semantics*

Visual notations used in the development of software-intensive systems such as UML models and BPMN are oriented to human communication [6]: their sole purpose is to facilitate the communication and problem-solving activities central to design. Visual notations comprise visual syntax—composed of a symbolic vocabulary and grammar—and visual semantics—elements that give meaning to each symbol and symbol relationship. Together these enable the designer to “off-load” memory and information processing [4], promoting discovery and inferences about the process at hand. Figure 1 relates visual syntax and semantics with usage levels of type (language) and instance (sentence).

Errors of visual syntax include the use of invalid symbols (e.g. the use of a BPMN ‘gateway’ symbol to model an ‘activity’). Such symbol instance errors are relatively easy to detect given the standardization of the BPMN notation. Syntactic



**Fig. 1** Visual notation: syntax and semantics (from Moody, 2009)

errors equate to programming code the compiler does not understand (e.g. an instruction to multiply a string with an integer in C). The compiler will detect them, because it cannot compile them. Semantic errors are more subtle, offering valid symbol use but presenting an unintended or ambiguous design construct—for example, if the sequence of activities in a retail checkout scenario were mis-ordered. Such a logical error requires scrutiny of the model content rather than its form to detect.

Prior research (e.g. [2]) shows that appreciation for and measurement of cognitive effectiveness of notational form (syntax) is particularly lacking in understanding of how diagrams support the design process. The use and qualities of BPMN and other (external) representations provide a medium through which the formation of cognitive (internal) frameworks during design activity can be explored.

However, research to date has tended to focus on the semantic content of BPMN and similar diagrams, largely neglecting the effects of visual syntax. This offers a significant opportunity, since the power of graphical images stems from their capacity to “tap” the highly parallel human visual and cognitive systems [6]. We address this opportunity by exploring a range of neurophysiological artifacts that might be used to assess cognitive processing of symbol and construct “instances”. Using the “visual sentence” as the unit of analysis to accommodate syntactic and semantic errors, our research protocol operationalizes the elements in the (shaded) lower half of Fig. 1.

Moody [3] points out that cognitive effectiveness is not an intrinsic property of visual representations. Cognitive effectiveness is something that must be designed into them [6]. This is critical to our thesis: the utility, effectiveness, and other aspects of “goodness” of the design are functions of both the notation—its completeness in terms of capacity to represent the realm of problem and solution spaces plus its ease-of-use—and the competence of the designer. The malleability of visual representations also offers the opportunity to “design out” effectiveness through

error seeding and interference [7, 8]. This aspect of our research extends prior studies of business process design (e.g. [4]) and offers the potential to explore the characteristics of expertise that mediate modeling competence and effectiveness [9].

## 2.2 *Expertise and Error Diagnosis*

Differentiating novices from experts within any given profession can be challenging, given the myriad ways expertise has been defined in the literature. Experts know a great deal about a particular domain and understand how their discipline is organized: they differ from novices in terms of their knowledge, effort, recognition, analysis, strategy use and monitoring [5]. This includes an ability to comprehend and contribute to the language (including both its syntax and semantics) and methodology of the discipline (including notations and other tools). As expertise develops, performance becomes more intuitive and automatic and knowledge more tacit [10]. At this level of mastery, an individual immediately understands the critical aspects of a given situation and does not focus on the less significant attributes.

Novices are individuals who have limited or no experience in situations characteristic of their domain. A novice's understanding of the discipline is based largely on rules. At this level, novices can be seen as learners who rely on facts and features of the domain to guide their behavior and choices. Thus:

- Experts notice features and meaningful patterns of information that are not noticed by novices [11].
- Experts have acquired a great deal of content knowledge that is organized in ways that reflect a deep understanding of their subject matter [12].
- Experts' knowledge is tacit and cannot be reduced to sets of isolated facts or propositions. Rather, it reflects contexts of applicability: that is, the knowledge is conditioned on(to) a set of circumstances—in this case, the business scenarios presented during our experiments.
- Experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort: it is “ready to mind” [13].

These characteristics show that expertise is a significant mediator of cognitive effectiveness. Prior studies show that learning and skilled performance produce changes in brain activation [5]. Since the understanding of brain mechanisms is synergistic with understanding of behavioral mechanisms, eye-tracking offers a potent means to explore skill-based adaptations. Measures such as saccades and fixations have been shown to provide very good predictions of expertise in a range of contexts including surgery [14–16]. The interdependence of the (internal) mental model and (external) representational model provides the potential to use eye-tracking to characterize expertise more empirically.

This methodology should be well suited to study expertise in conceptual modeling which generally presents a significant challenge due to its inherent complexity: it has been argued that conceptual modeling is one of the most cognitively complex undertakings known in our field. This may explain the paucity of research: while eye-tracking has been used to assess model comprehension (e.g. [17]) and expertise as a moderator of information sourcing [9], we are aware of no such exploration of the impact of expertise on model comprehension.

### 3 Proposed Research Design and Protocols

To investigate our research question, we designed an experiment to exploit the capacity of eye-tracking artifacts to differentiate between novices and experts. In this within-subject experiment, novice and expert subjects are instructed to find errors in conceptual models. The cognitive effectiveness, accuracy, and confidence of the subjects will be compared. Initially, we propose to use years of experience and training as surrogate indicators of expertise.

Design-build iterations can be imitated using a variety of design candidate sentences (see Fig. 1) presented in BPMN [18]. We propose an experiment using blocks of models: each block will represent the same scenario (e.g. retail banking transactions or fast food ordering). Each block will comprise sets of near-identical models (i.e. identical in terms of scope and the number of BPMN symbols). Each set will contain three versions of a candidate design: one with no known errors; one with a known semantic error and one with a known syntactic error. The sets will be presented to the subject in a random order and the subject will see a series of 15 design candidates and be asked to identify if an error is present, what type of error it may be, and the extent to which they are confident about their assessment. Thus, each subject will evaluate 60 candidates in total (four blocks of 15 sets).

Figure 2 illustrates an example of the retail-banking scenario developed for the pilot study. This is one of four “sentences” in the experimental block. The models shown contain no known errors (a), a seeded semantic error (b), and a seeded syntactic error (c). Based on our pre-test, we anticipate that subjects would take between 15 and 20 min to inspect and judge a block of 15 models at this scale: the estimated total time for the experiment will be 1 h.

Data will be gathered using eye-tracking (Red 250, SensoMotoric Instruments GmbH, Tetlow, Germany). Building on guidelines from Léger et al. [19] and Riedl and Léger [20], the data will be analyzed to measure how much time was spent looking at the model before making a decision, time spent and number of visits in the area containing an error (area of interest), and time to first fixation to that area. Reaction times are also of interest in the form of time spent on each model image before answering whether or not an error is identified. Confidence will be measured with a one-item scale of 1–5 (from low confidence to high confidence) after each set.

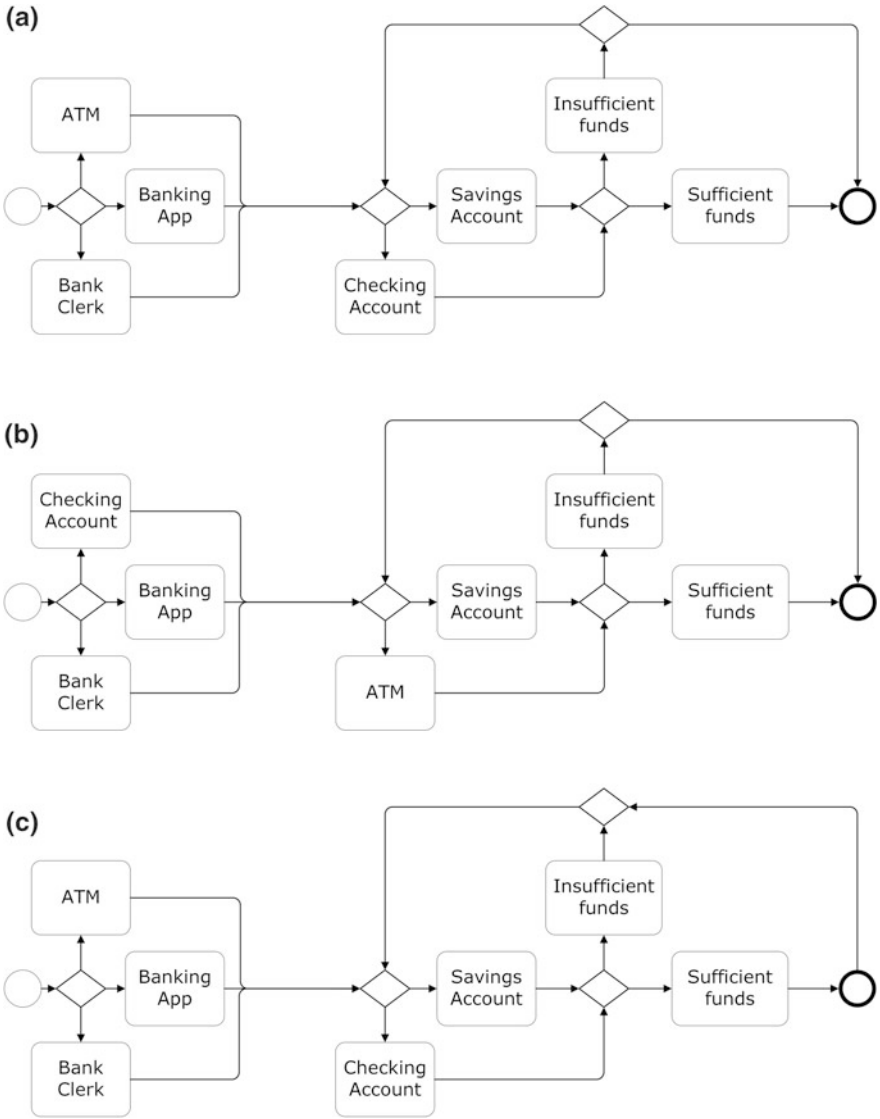


Fig. 2 Retail-banking scenario “sentences”

### 4 Discussion and Next Steps

The study is currently underway and initial results are being analyzed. The anticipated benefits and contributions of this study are substantial. Eye-tracking is particularly well suited to exploration of differences between how novices and experts use their knowledge, skills, and abilities to perform tasks. This provides

insight into the efficiency of search patterns, error detection, and decision-making in the context of conceptual modeling.

By identifying the differences in the strategies novices and experts use to search, detect, and diagnose, we anticipate being able to define training programs appropriate for each level to improve performance and result quality. Use of randomized blocks of similar but different “sentences” provides opportunities to “tune” the ET data capture by optimizing the model scale to capture data from the areas of interest. A further benefit of our proposed methodology is the potential to partition the models using a static (non-tracked) screen during which the location and type of error can be recorded. We anticipate the use of touchscreen or mouse-controlled annotation for data capture.

The experimental design allied with data capture and analysis at the symbol instance level offers discriminant validity (intra-subject) for our experiments. Further authority is added through the identification of multiple AOIs (error site(s); whole page) which provide units of analysis that are more finely grained than those that simply differentiate overall performance between subjects.

Future studies will be conducted with the addition of electroencephalography to identify cerebral activity underlying the error detection process. More specifically, we will use EFRP (eye-fixation related potentials) [19] to isolate the participant’s reaction to finding errors. Error-related potentials (ErrP) [2] are likely to be present at the moment of first fixation of the area of the model containing the error. We can then understand more by comparing the ErrP of experts and novices as well as the ErrP for semantic and syntactic errors. Such measures will enable us to broaden our understanding of process model design by operationalizing the skill-based adaptations that characterize design expertise [5], thus enabling us to differentiate them from prior studies of novice designers [4]. They will also offer insight into the choices of (external) representation [4, 21]. In turn, such insights will inform the development of decision support systems for model designers and enhanced curriculum development.

## References

1. Brocke, V.J., Riedl, R., Léger, P.: Application strategies for neuroscience in information systems design science research. *J. Comput. Inf. Syst.* **53**(3), 1–13 (2013)
2. Yeung, N., Botvinick, M.M., Cohen, J.D.: The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* **111**(4), 931 (2004)
3. Moody, D.: The “physics” of notations: towards a scientific basis for constructing visual notations in software engineering. *IEEE Trans. Softw. Eng.* **35**(5), 756–778 (2009)
4. Recker, J., Safrudin, N., Rosemann, M.: How novices design business processes. *Inf. Syst.* **37**, 557–573 (2012)
5. Hill, N., Schneider, W.: Brain changes in the development of expertise: neuroanatomical and neurophysiological evidence about skills-based adaptations. In: Ericsson, K., Charness, N., Feltovich, P., Hoffman, R. (eds.) *The Cambridge Handbook of Expertise and Expert Performance*, pp. 653–682. CUP, Cambridge, UK (2006)



6. Larkin, J., Simon, H.: Why a diagram is (sometimes) worth ten thousand words. *Cogn. Sci.* **11**(1) (1987)
7. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J.: *Eye tracking: a comprehensive guide to methods and measures*. OUP Oxford, UK (2011)
8. Hungerford, B., Hevner, A., Collins, R.: Reviewing software diagrams: a cognitive study. *IEEE Trans. Softw. Eng.* **30**(2), 82–96 (2004)
9. Léger, P.M., Riedl, R., vom Brocke, J.: Emotions and ERP information sourcing: the moderating role of expertise. *Ind. Manag. Data Syst.* **114**(3), 456–471 (2014)
10. Bruer, J.: *The Mind's Journey from Novice to Expert*. James S. McDonnell Foundation, Saint Louis, MO (2010)
11. Davis, C., Hufnagel, E.: Through the eyes of experts: a socio-cognitive perspective on the automation of fingerprint work. *MIS Q.* **31**(4), 681–703 (2007)
12. Masson, S., Potvin, P., Riopel, M., Brault Foisy, L.-M.: Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity. *Mind Brain Educ.* **8**(1), 44–55 (2014)
13. Bánáthy, B.: An information typology for the understanding of social systems. *Syst. Res. Behav. Sci.* **16**(6), 479–494 (1997)
14. Gegenfurtner, A., Lehtinen, E., Säljö, R.: Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educ. Psychol. Rev.* **23**(4), 523–552 (2011)
15. Jarodzka, H., Scheiter, K., Gerjets, P., Van Gog, T.: In the eyes of the beholder: how experts and novices interpret dynamic stimuli. *Learn. Instr.* **20**(2), 146–154 (2010)
16. Law, B., Atkins, M., Kirkpatrick, A., Lomax, A.: Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In: *Proceedings of the 2004 symposium on Eye tracking research & applications*, pp. 41–48. ACM (2004)
17. Yusuf, S., Kagdi, H., Maletic, J.I.: Assessing the comprehension of UML class diagrams via eye tracking. In: *Program Comprehension, ICPC'07. 15th IEEE International Conference on Program Comprehension (ICPC)*, pp. 113–122. IEEE (2007)
18. Hevner, A., Davis, C., Collins, R., Gill, T.: A neurodesign model for IS research. *Informing Sci. Int. J. Emerg. Transdiscipline* **17**, 103–132 (2014)
19. Léger, P.M., Sénécal, S., Courtemanche, F., de Guinea, A.O., Titah, R., Fredette, M., Labonte-LeMoine, É.: Precision is in the eye of the beholder: application of eye fixation-related potentials to information systems research. *J. Assoc. Inf. Syst.* **15**(10), 651 (2014)
20. Riedl, R., Léger, P.M.: *Fundamentals of NeuroIS*. Springer, Berlin (2016)
21. Fuller, R., Davis, C.: Requirements elicitation techniques as communication channels: a framework to widen the window of understanding. In: *Information Systems Analysis and Design: Foundations, Methods and Practices, Advances in Management Information Systems (AMIS) Monograph Series* (2009)

# A Neuro-Cognitive Explanation for the Prevalence of Folder Navigation and Web Browsing

Ofer Bergman and Yael Benn

**Abstract** We describe our mapping of the neural correlates associated with different ways by which users access digital information. Despite advances in search technology and its flexibility, users prefer to retrieve files using hierarchical folders navigation. This requires an explanation. In two studies, using a dual task and functional magnetic resonance imaging (fMRI), we show that folder navigation uses brain structures involved in physical navigation, hence requiring little verbal attention. In contrast, search recruits classic language structures (Broca's area). We further examine search versus browsing preferences on a popular supermarket website and show that users prefer browsing rather than searching for products. Qualitative analysis revealed that this preference was due to verbal-cognitive overload. Our next two studies will use the dual task paradigm and fMRI to examine the cognitive and neural correlates of search versus browsing in the web environment. We hypothesize that results will replicate our previous findings for files.

**Keywords** Web browsing · Search · Navigation · Dual task paradigm · fMRI

## 1 Introduction

Personal and public information differ in several ways: Personal information items (e.g., files and emails) are stored and retrieved by a user, who is typically familiar with their own organization scheme. In contrast, public information items (such as web pages) are generally not organized by the user who retrieves them, and hence

---

O. Bergman (✉)  
Bar-Ilan University, Ramat Gan, Israel  
e-mail: bergmao@mail.biu.ac.il

Y. Benn  
Manchester Metropolitan University, Manchester, UK  
e-mail: Y.Benn@mmu.ac.uk; Y.Benn@sheffield.ac.uk

Y. Benn  
The University of Sheffield, Sheffield, UK

are often unfamiliar [1]. However, the main retrieval options for both personal and public information items are somewhat similar: *Search* is a method by which users first generate a query specifying an attribute of the target item, and when the search engine returns a set of results, the user can select the relevant item; *Navigation/Browsing* refers to the method whereby users manually traverse the information organization structure until they reach the target information item. Folder navigation (navigation for short) and web browsing (browsing for short) are similar but not identical, as navigation is strictly hierarchal (following the folder hierarchy), while browsing allow users to reach any web page from any other web page. Still, both navigation and browsing require the user to remember the path leading to the information item. In contrast, search is more flexible, allowing users to apply any attribute that they happen to remember about the information item to be used in the search query. However, despite this flexibility and recent advances in search technology, navigation dominates Personal Information Management (PIM) retrieval [1–6], and browsing is prevalent in web retrieval [4, 7]. This phenomenon requires an explanation.

In here, we describe our journey to uncover why people often use navigation and browsing despite these methods’ apparent limitations compared with search. The first two studies we conducted related to PIM. Study 1 [8] used a dual task paradigm to demonstrate that navigation requires less verbal attention than search. Study 2 [9] used fMRI and provided an explanation for the results of the first study, revealing that virtual folder navigation recruited the same brain structures as real-world navigation, while searching involved Broca’s area, which is associated with linguistic processing. Study 3 [10] examined browsing and searching behavior on a large UK-supermarket website. Studies 4 and 5, which we are currently planning, will utilize a dual-task paradigm and fMRI to examine the cognitive and neural correlates of information retrieval on the web.

## 2 Why is Folder Navigation and Web Browsing so Prevalent?

Folder navigation forces users to remember the exact path to a specific information item, which can be difficult, especially if time has elapsed between an item’s storage and retrieval [11]. Similarly, web browsing depends on the users’ memory in cases where they are familiar with the web site, or on their ability to guess where their target page is located for unfamiliar sites [12]. Compared with navigation and browsing, search is more flexible, allowing users to retrieve an item using any attribute they happen to remember (e.g., a word it contains) [11]. These benefits, along with significant recent improvements in desktop search engines, should induce strong user preferences for search over navigation [11, 13–15]. However, research consistently shows that users prefer navigation over search [2–6] and that file search is used only as a last resort [1]. Search is generally more common on the

Web compared to within PIM. However, even on the Web, search is less frequent than we might expect [16]. Users focus on re-finding rather than seeking novel information, and make extensive use of their “back” button and other options such as history [17–19], suggesting that browsing is a prominent retrieval method [4, 7].

Why do people prefer navigation, despite its apparent limitations compared with search? In previous work [1] we proposed several possible explanations, including users’ familiarity with their own folder structure, which stays relatively stable over time. By contrast, the flexibility of search may compromise consistency, as users are able to retrieve the same file using different search terms, and receiving different results. Here, we explore the neurocognitive basis of search, navigation, and web browsing. Computer users typically retrieve information items while in the midst of engaging in another activity (e.g., writing an article), which they intend to continue pursuing after the retrieval. Hence, being able to keep the current task in mind would eliminate the time and cognitive cost of recalling where they were. As such, the preference for a retrieval option that demands less verbal attention is rational. Therefore, we hypothesized that: (a) navigation and browsing requires less verbal attention than search (studies 1, 3 and 4); (b) navigating and browsing in the virtual and physical environments rely on the same deep-brain structures, while search relies on linguistic brain structures (studies 2 and 5).

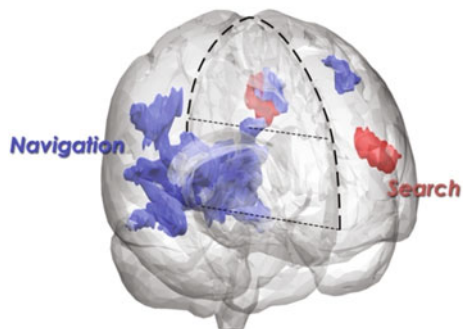
### **3 Study 1: Navigation Requires Less Verbal Attention**

In our first study [8], we tested the hypothesis that file navigation requires less verbal attention than search by applying a dual-task paradigm. Using a within-subject design, we read a list of words to each of our 62 participants. We then asked each participant to navigate or search to a target file (counterbalancing the order), and then to recall the words from the list. Comparing the number of words recalled in each condition revealed that participants remembered significantly more words when retrieving by navigation than by search. The improved performance at the secondary verbal-memory task when navigating indicates that verbal resources (such as the phonological loop) were available while navigating, suggesting that navigation requires less verbal attention than search. Our results also cast doubt on the assumption that search is more efficient and easier than navigation: search took nearly three times longer than navigation, was more vulnerable to mistakes and retrieval failures, and was subjectively evaluated as more difficult.

## 4 Study 2: Folder Navigation Uses the Same Brain Structures as Real-World Navigation

Study 1 left an open question: why does folder navigation require less verbal attention than search? In [9] we hypothesized that file navigation relies on the same brain structures that are used for real-world navigation, as it is not the folders' names, but rather the hierarchical structures that are used for retrieval. Navigation-related brain structures are located deep in the brain, largely in the posterior part [20], and are distinct from traditionally linguistic brain structures [21]. In contrast, search requires users to identify a precise and relevant search term, which is likely to involve language-related brain structures. We tested these hypotheses using fMRI. Seventeen participants were asked to search and navigate to files on their own computer. This was achieved by projecting their laptop display onto the mirror in the MRI head-coil. Interaction was achieved via MRI-compatible mouse, which rested on a flat surface on the participants' lap, and was connected to their laptop using MRI-compatible USB cable. While they searched and navigated, their brain activity was recorded. Two control tasks, one for search and one for navigation, were used to establish a base-line brain activity evoked by visual and motor actions. Results indicated that folder navigation recruited the posterior limbic (including the retrosplenial cortex) and parahippocampal regions, similar to those previously observed during real-world navigation in both animals [22–24] and humans [20, 25–27]. In contrast, search activated the left inferior frontal gyrus, commonly known as Broca's area, which is often observed during linguistic processing (see Fig. 1 for a 3D illustration of the brain activation in both conditions). Combined with evidence from study 1 [8], these results suggest that mechanisms that allow for the retrieval of an item from a specific location serve us in both real and virtual environments. These deep-rooted neurocognitive routines, of navigating through the same path used in storing the (information) item, are related to spatial memory, and have minimal reliance on linguistic processing, leaving the language system available for other tasks.

**Fig. 1** A 3D model illustrating bilateral posterior regions activated for folder navigation (blue), and left inferior frontal activation for search (red)



## 5 Study 3: Web Browsing is Prevalent

To test the hypothesis that web browsing is prevalent when the environment is familiar, and that some of the reason for this preference is due to the increased linguistic processing required for search, we carried out an eye-tracking study. Forty participants performed their weekly shopping on a large UK supermarket website [10]. Results indicate that the number of product-pages reached via browsing was significantly higher than those reached by searching ( $t(39) = -2.4$ ,  $p = 0.023$ ). Avoidance of linguistic overload was manifested in the eye-tracking data: the most common items people looked at were the food pictures, followed by the product title and the price. High-content verbal information (e.g., ingredients/nutrition lists) was rarely viewed, even by those reporting to be on a diet or having special dietary requirements. Ten of the participants were invited to view where their eyes looked during the shopping and reflect on it. High linguistic load was often reported (e.g., trying to remember shopping lists), as well as linguistic-related difficulties with search (e.g., difficulties in spelling or identifying appropriate search terms). Such difficulties often resulted in returning to the browsing option. In addition, as with the PIM system, search was often a result of a failed browsing attempt. These qualitative results indicate that search requires greater verbal processing than browsing; however, this should be empirically tested using quantitative research as proposed in studies 4 and 5. In addition, future research could examine cognitive load more directly using EEG [28].

## 6 Study 4: Does Browsing Require Less Verbal Attention?

To test our hypothesis that web browsing is prevalent because it requires less verbal attention, we will again use the dual-task paradigm. Given that research shows that revisiting web pages is the rule rather than the exception [17–19], we will select web pages collected from the participants' browser history. Using a technique called Elicited Personal Information Retrieval (EPIR), which we previously employed in several studies (e.g., [29]) including studies 1 and 2, participants will be asked to retrieve familiar web pages by screening an image of the web page on another computer. The image will not show the URL, and the verbal information on the page will be obscured, so as to avoid participants' use of the text in the search query, hence maximizing ecological validity [30]. Participants' screens will be recorded during the retrievals using dedicated software. As a secondary task, a variation of the verbal shadowing task [31] will be applied. During the retrieval process, participants will be asked to listen to a stream of nonsensical sentences, and when one of three specific words appear (e.g., "book," "phone," or "light"), to press a button positioned at the bottom corner of the screen. Accuracy and response-time will be recorded. We hypothesize that during browsing, accuracy on the secondary task will be higher and response-time shorter when compared with performance on the secondary task during search.

## 7 Study 5: The Neural Correlates of Search and Web Browsing

Study 5 will investigate the neural correlates that are involved in searching and browsing for information on the web. While we hypothesize that for the most part, results will replicate previous findings, we suggest that the need to re-examine a browsing path may place increased linguistic demands on browsing compared to navigating. To examine this hypothesis, we will recruit 20 participants for a block-design fMRI study. Prior to the study, participants will be asked to explore several websites of familiar structure (e.g., restaurant websites). During the experiment, all participants will be asked to both search and browse for information (e.g., location or price of an item) on these websites. The control conditions for this experiment are difficult to devise and have been the main focus in the development of this study. Currently, we are still considering the different options. It is predicted that browsing will result in less language-related activations and more hippocampus and posterior structures activations, while searching will result in increased language-related activations.

## References

1. Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N., Whittaker, S.: Improved search engines and navigation preference in personal information management. *ACM Trans. Inf. Syst.* **26**, 1–24 (2008)
2. Barreau, D.K., Nardi, B.A.: Finding and reminding: file organization from the desktop. *SIGCHI Bull.* **27**, 39–43 (1995)
3. Boardman, R., Sasse, M.A.: “Stuff goes into the computer and doesn’t come out”: a cross-tool study of personal information management. In: *SIGCHI Conference on Human Factors in Computing Systems*, pp. 583–590. ACM Press, 985766, April (2004)
4. Teevan, J., Alvarado, C., Ackerman, M.S., Karger, D.R.: The perfect search engine is not enough: a study of orienteering behavior in directed search. In: Dykstra-Erickson, E., Tscheligi, M. (eds.) *SIGCHI Conference on Human Factors in Computing Systems*, pp. 415–422. ACM Press, New York (2004)
5. Kirk, D., Sellen, A., Rother, C., Wood, K.: Understanding photowork. In: *SIGCHI Conference on Human Factors in Computing Systems*, pp. 761–770. ACM, New York (2006)
6. Capra, R.G., Pérez-Quiñones, M.A.: Using web search engines to find and re-find information. *Computer* **38**, 36–42 (2005)
7. Obendorf, H., Weinreich, H., Herder, E., Mayer, M.: Web page revisitation revisited: implications of a long-term click-stream study of browser usage. In: *Proceedings of the SIGCHI Conference on Human factors in computing systems*, pp. 597–606. ACM, New York (2007)
8. Bergman, O., Tene-Rubinstein, M., Shalom, J.: The use of attention resources in navigation versus search. *Pers. Ubiquit. Comput.* **17**, 583–590 (2013)
9. Benn, Y., Bergman, O., Glazer, L., Arent, P., Wilkinson, I.D., Varley, R., Whittaker, S.: Navigating through digital folders uses the same brain structures as real world navigation. *Sci. Rep.* **5**, 1–8 (2015)

10. Benn, Y., Webb, T.L., Chang, B.P., Reidy, J.: What information do consumers consider, and how do they look for it, when shopping for groceries online? *Appetite* **89**, 265–273 (2015)
11. Lansdale, M.W.: The psychology of personal information management. *Appl. Ergonomics* **19**, 55–66 (1988)
12. Shneiderman, B., Plaisant, C.: *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Publ. Co., Reading (2010)
13. Fertig, S., Freeman, E., Gelernter, D.: “Finding and reminding” reconsidered *SIGCHI Bulletin* **28**, 66–69 (1996)
14. Russell, D., Lawrence, S.: Search everything. In: Jones, W., Teevan, J. (eds.) *Personal information management*, pp. 153–166. University of Washington Press, Seattle and London (2007)
15. Cutrell, E., Dumais, S.T., Teevan, J.: Searching to eliminate personal information management. *Commun. ACM* **49**, 58–64 (2006)
16. Bergman, O., Whittaker, S.: *The Science of Managing Our Digital Stuff*. MIT Press (2016)
17. Cockburn, A., Greenberg, S.: Issues of page representation and organisation in web browser’s revisit tools. *Australas. J. Inf. Syst.* **7**, 120–127 (2000)
18. Bruce, H., Jones, W., Dumais, S.: Keeping and re-finding information on the web: what do people do and what do they need? *Proc. Am. Soc. Inf. Sci. Technol.* **41**, 129–137 (2004)
19. Aula, A., Jhaveri, N., Kaki, M.: Information search and re-access strategies of experienced web users. In: *Proceedings of the 14th International Conference on World Wide Web*, pp. 583–592. ACM, New York (2005)
20. Maguire, E., Frackowiak, R.S., Frith, C.D.: Recalling routes around London: activation of the right hippocampus in taxi drivers. *J. Neurosci.* **17**, 7103–7110 (1997)
21. Catani, M., Jones, D.K.: Perisylvian language networks of the human brain. *Ann. Neurol.* **57**, 8–16 (2005)
22. Furuya, Y., Matsumoto, J., Hori, E., Boas, C.V., Tran, A.H., Shimada, Y., Ono, T., Nishijo, H.: Place-related neuronal activity in the monkey parahippocampal gyrus and hippocampal formation during virtual navigation. *Hippocampus* **24**, 113–130 (2014)
23. Furtak, S.C., Wei, S.M., Agster, K.L., Burwell, R.D.: Functional neuroanatomy of the parahippocampal region in the rat: the perirhinal and postrhinal cortices. *Hippocampus* **17**, 709–722 (2007)
24. Bingman, V.P., Mench, J.A.: Homing behavior of hippocampus and parahippocampus lesioned pigeons following short-distance releases. *Behav. Brain Res.* **40**, 227–238 (1990)
25. Epstein, R., Kanwisher, N.: A cortical representation of the local visual environment. *Nature* **392**, 598–601 (1998)
26. Maguire, E.: The retrosplenial contribution to human navigation: a review of lesion and neuroimaging findings. *Scand. J. Psychol.* **42**, 225–238 (2001)
27. Aguirre, G.K., Detre, J.A., Alsup, D.C., D’Esposito, M.: The parahippocampus subserves topographical learning in man. *Cereb. Cortex* **6**, 823–829 (1996)
28. Müller-Putz, G.R., Riedl, R., Wriessnegger, S.C.: Electroencephalography (EEG) as a research tool in the information systems discipline: foundations, measurement, and applications. *Commun. Assoc. Inf. Syst.* **37**, 46 (2015)
29. Bergman, O., Whittaker, S., Falk, N.: Shared files—the retrieval perspective. *J. Am. Soc. Inform. Sci. Technol.* **65**, 1949–1963 (2014)
30. Ingwersen, P.: Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *J. Documentation* **52**, 3–50 (1996)
31. Hermer-Vazquez, L., Spelke, E.S., Katsnelson, A.S.: Sources of flexibility in human cognition: dual-task studies of space and language. *Cogn. Psychol.* **39**, 3–36 (1999)



# Physiological, Psychological, and Behavioral Measures in the Study of IS Phenomena: A Theoretical Analysis of Triangulation Strategies

Kevin Hill and Stefan Tams

**Abstract** Recent NeuroIS research has suggested that physiological measures could contribute to an improved explanation and prediction of IS phenomena. However, few studies have examined a combination of different kinds of measures, raising the question of how the propagated improvement in explaining and predicting IS phenomena can be achieved. Therefore, research is needed that sheds light on the interrelationship amongst physiological measures (i.e., NeuroIS), psychological measures (i.e., perceptual, self-report), and behavioral measures (i.e., directly observed behaviors). Drawing on the methodological triangulation approach, this research essay endorses the use of multiple measures in the study of IS phenomena, and it discusses two strategies that can be useful in this endeavor: convergent validation and holistic representation. The former aims to explain and predict variance in IS dependent variables with greater certainty, while the latter intends to increase the amount of variance explained. The essay concludes that—although both strategies have merit—holistic representation is where NeuroIS could play an especially important role.

**Keywords** NeuroIS • Self-report • Perceptual • Behavioral • Physiological data

---

K. Hill (✉)

Department of Human Resources Management,  
HEC Montréal, Montréal, QC, Canada  
e-mail: kevin.hill@hec.ca

S. Tams

Department of Information Technologies,  
HEC Montréal, Montréal, QC, Canada  
e-mail: stefan.tams@hec.ca

## 1 Introduction

It has long been postulated that one of the major contributions of NeuroIS to the general field of IS research would be through improvements in the explanation and prediction of IS phenomena [1, 2]. To date, however, only a few NeuroIS studies have employed multiple measures [3], which may be reflective of lingering confusion as to the specific opportunities provided by NeuroIS to advance the IS field. In order to encourage more NeuroIS research, the present essay examines the question of precisely *how* NeuroIS can contribute to an improved explanation and prediction of IS phenomena by highlighting the distinct benefits of methodological triangulation, involving the active comparison of results using physiological measures (i.e., NeuroIS measures) alongside psychological (i.e., self-reported) and behavioral (i.e., observed) measures.

To this end, our essay begins by providing the necessary background on methodological triangulation. Two strategies are specifically discussed: convergent validation and holistic representation. While each of these triangulation strategies can improve the explanation and prediction of IS phenomena, we emphasize in this essay that each makes a distinct contribution. Our aim is to sensitize NeuroIS researchers to the nature of this distinction as well as to the importance of aligning one's methodological strategy to the type of contribution that is needed to advance understanding in a given domain. We conclude with a discussion of the anticipated contributions of methodological triangulation to the IS field in light of the nature of NeuroIS phenomena and recent empirical results that are consistent with the holistic representation approach.

## 2 The Triangulation Approach

Triangulation involves using a combination of different methods in the study of the same phenomenon [4]. Researchers have identified two distinct triangulation strategies: convergent validation and holistic representation [5].

The convergent validation approach to triangulation was elaborated by Campbell and Fiske [6] as a means for distinguishing substantive variance in a theoretical construct from unwanted method variance (systematic variance associated with the use of a given method), the latter of which undermines construct validity and introduces bias in the estimates of theoretically proposed relationships. This approach is employed widely by those developing new measures of an existing theoretical construct and those seeking greater confidence in their estimates of relationships between distinct theoretical constructs. For instance, Ortiz de Guinea et al. [7] used a convergent validation approach (the multi-trait multi-method matrix) to assess the validity of three IS constructs (engagement, arousal and cognitive load). Their results demonstrated good correspondence between self-reported and neurophysiological measurement of these constructs, increasing

researchers' confidence that these different forms of measurement are reflective of the same underlying construct.

To the extent that NeuroIS measures are significantly positively correlated with traditional psychological and behavioral measures and to the extent that these measures converge in their predictions of outcomes, this corroborative evidence reinforces the researcher's certainty in terms of the assessment of the specific construct under investigation and in its capacity to predict theoretically relevant IS outcomes [5, 6].

While some increase in explanatory power may be achieved through the combined use of convergent measures, the primary contribution of each form of measurement in the convergent validation approach is to compensate for distinct sources of error variance associated with other forms. That is, to the extent that the same underlying dimension of the phenomenon is reflected in each measure, none is expected to explain significant increments in the variance of theoretically related outcomes over and above that accounted for by the others (see Fig. 1).

In sharp contrast to convergent validation, the holistic representation approach to triangulation is prevalent among mixed-methods researchers who hold that, beyond method-specific variance, different methods also tap into *distinct theoretically relevant dimensions* of a phenomenon and, therefore, that different forms of measurement will diverge in their predictions of outcomes [5, 8, 9]. By tapping unique dimensions, the combined use of multiple measures provides a more complete picture of the phenomenon under investigation and, through their combined effects, more powerful predictive relationships. In other words, to the extent that holistic representation applies, it would not be expected to find that physiological, psychological and behavioral measures correlate significantly and each measure would offer researchers the possibility of explaining unique theoretically relevant variance in outcomes (see Fig. 2).

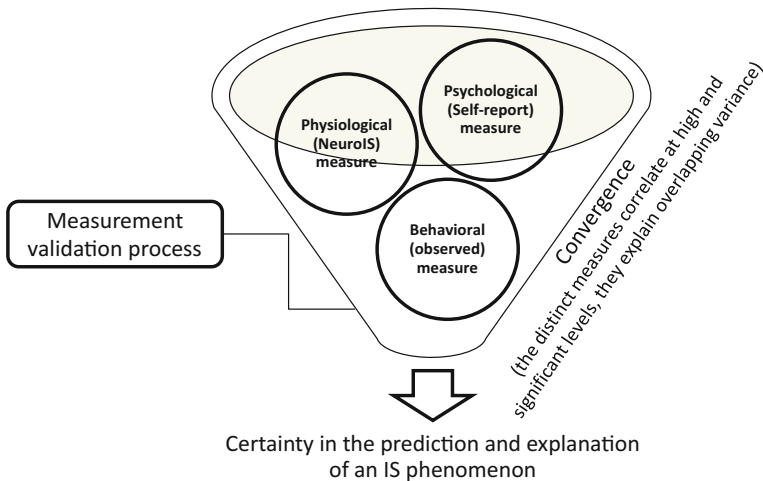


Fig. 1 Convergent validation approach to methodological triangulation

**Fig. 2** Holistic representation approach to methodological triangulation

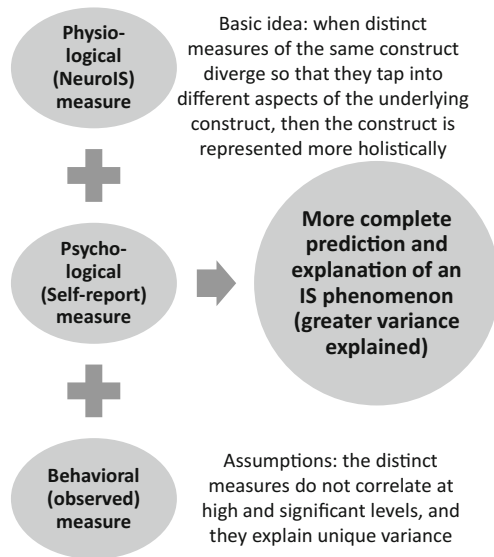


Table 1 summarizes the two distinct ways that mixed-methods research employing NeuroIS measures alongside psychological and behavioral ones can yield improved explanation and prediction of IS phenomena. Researchers contemplating the use of NeuroIS measures should begin by asking themselves what is the central limitation of the current theoretical explanation of the phenomenon in question that can be addressed through NeuroIS, which will guide their selection between convergent validation and holistic representation as the dominant methodological strategy for resolving this challenge. It goes without saying that this choice must be guided by existing theory and empirical evidence [3].

**Table 1** A guide to selecting and drawing conclusions from two triangulation strategies

Specific limitation of current theoretical explanation	Triangulation approach	Inter-method correlation	Prediction of relationships	Nature of study contribution
Concerns that prior results are biased by method-specific variance	<b>Convergent validation</b> <i>Objective</i> Corroboration of current explanation	Significant correlation between measures of construct across methods	Consistent (no incremental variance explained)	More certainty in existing explanation of phenomenon
Concerns that prior results provide an incomplete understanding of the phenomenon	<b>Holistic representation</b> <i>Objective</i> Extension of current explanation	Non-significant correlation between measures of construct across methods	Novel (significant incremental variance explained)	More complete explanation of phenomenon

Depending on the approach selected, Table 1 further specifies the empirical criteria that would enable researchers to draw conclusions as to how they have addressed the identified problem, as well as the precise nature of the theoretical contribution achieved by their investigation. In either case, the contribution is determined on the basis of two tests: (1) the test of the significance of the correlation between the construct measured by the distinct methods and (2) the test of the significance of the incremental variance explained in a theoretically related outcome by the introduction of the NeuroIS measure.

### 3 What Relationship to Expect

While each approach to methodological triangulation offers distinct opportunities to improve explanation and prediction, here we elaborate upon the theoretical reasons to expect that holistic representation holds the most promise as a methodological strategy for combining physiological, psychological and behavioral measures in the study of IS phenomena.

Convergent validation rests on the assumption that distinct methods will converge in their assessment of a theoretical construct. Yet, physiological measures represent functions and processes related to the human body that are largely unconscious, whereas self-report measures frequently are designed to represent beliefs of which people should be consciously aware [10]. Furthermore, behavioral measures capture how conscious and unconscious processes manifest themselves in actual behavior [11]. Yet, behavior is additionally constrained by factors external to the self [12]. Therefore, consistent with the holistic representation approach, each of these methods is likely to tap a distinct theoretical dimension of a phenomenon. Research that has deliberately tested for convergence between unconscious, conscious, and behavioral indicators of the same phenomenon finds that people hold conscious beliefs that are inconsistent with unconscious indicators or else are inconsistent with their behaviors, be these non-verbal behaviors or discrete choices [13–15]. Moreover, people are largely unaware of the extent to which their beliefs are inconsistent in these respects [16].

For these reasons, we suggest that NeuroIS research is likely to complement what can be understood by the examination of psychological or behavioral measures by *capturing the unconscious dimensions* of IS phenomena and *explaining unique variance* in related dependent variables. The validity of this claim has been verified empirically for technostress research, where it was found that physiological and psychological measures of technostress did not correlate and that the physiological measure explained variance in computerized task performance above and beyond the variance explained by the psychological measure [17]. We encourage future mixed-methods research in this and in any area where theory suggests that conscious and unconscious processes jointly operate to predict outcomes

(e.g. automatic and controlled processing) [18]. Combining NeuroIS measures with psychological and behavioral ones also holds promise for research on phenomena, such as cognitive absorption or flow, which cannot be ascertained through conscious assessment without interrupting the cognitive processes that are inherent to the construct [19].

## 4 Conclusion

In this essay, we have promoted more mixed-methods NeuroIS research and a concerted triangulation approach to identifying the contributions of such research. As a first step, we recommend that NeuroIS researchers consult prior research and theory to obtain an a priori understanding of the expected relationship between distinct measures and to derive hypotheses as to how they each operate (either in common or uniquely) to predict outcomes. This a priori understanding should then inform the study design as well as the criteria by which results will be evaluated toward rendering conclusions. If different measures correlate at high and significant levels, they can be used together to reinforce the researcher's certainty in the prediction and explanation of an IS phenomenon. If not, the physiological (NeuroIS) measure can be examined to see if it explains unique variance in a theoretically related outcome above and beyond that explained by other measures. To the extent that NeuroIS researchers follow these steps, they will be in a position to more clearly demonstrate *how* the combined use of NeuroIS measures with psychological and behavioral ones contributes to the IS field.

In sum, NeuroIS measures are neither better nor worse than other measures currently employed in IS research, they are neither more accurate nor less accurate than other measures. To the extent that they can be used to capture the same dimension of an underlying IS construct as that captured by other measures, they help researchers to account for method bias and develop more certainty in their predictions. To the extent that they capture a different dimension of an underlying IS construct, they hold the potential to explain significant incremental variance in theoretically relevant outcomes. Given NeuroIS measures focus on the unconscious, rather than the conscious or behavioral dimensions of human experience, we suggest that their contribution to IS research will primarily take the form of increased richness of theoretical explanations and more powerful predictions of IS phenomena (see Tams et al. [17] for a detailed analysis of these ideas in the context of technostress research).

**Acknowledgements** This research was supported by the Social Sciences and Humanities Research Council of Canada.

## References

1. Riedl, R., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Dimoka, A., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Müller-Putz, G., Pavlou, P.A., Straub, D.W., vom Brocke, J., Weber, B.: On the foundations of NeuroIS: reflections on the Gmunden Retreat 2009. *Commun. AIS*. **27**, 243–264
2. Dimoka, A., Pavlou, P.A., Davis, F.D.: NeuroIS: The potential of cognitive neuroscience for information systems research. *Inf. Syst. Res.* **22**(4), 687–702 (2011)
3. Riedl, R., Davis, F., Hevner, A.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**(10), i–xxxv (2014)
4. Denzin, N.K.: *The research act: a theoretical introduction to sociological methods*, 3rd edn. Prentice-Hall, United States of America (1989)
5. Jick, T.D.: Mixing qualitative and quantitative methods: triangulation in action. *Adm. Sci. Q.* **24**(4), 602–611 (1979)
6. Campbell, D.T., Fiske, D.W.: Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* **56**(2), 81–105 (1959)
7. de Guinea, A.O., Titah, R., Léger, P.: Measure for measure: a two study multi-trait multi-method investigation of construct validity in IS research. *Comput. Hum. Behav.* **29**(3), 833–844 (2013)
8. Flick, U.: *An introduction to qualitative research*, 5th edn. Sage, Los Angeles (2014)
9. Mathison, S.: Why Triangulate? *Educ. Researcher* **17**(2), 13–17 (1988)
10. Davidson, D.: Mental events. In Lycan, W.G., Prinz, J.J. (eds.), *Mind and cognition*, 3rd edn, pp. 55–65. Blackwell Publishing, Malden (2008)
11. Ho, S.Y., Bodoff, D.: The effects of web personalization on user attitude and behavior: an integration of the elaboration likelihood model and consumer search theory. *MIS Q.* **38**(2), 497–520 (2014)
12. Ross, L., Nisbett, R.E.: *The person and the situation: perspectives of social psychology*. Pinter & Martin Publishers, London (2011)
13. Sapienza, P., Toldra-Simats, A., Zingales, L.: Understanding trust. *Econ. J.* **123**(573), 1313–1332 (2013)
14. Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K.: Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* **74**(6), 1464–1480 (1998)
15. Dovidio, J.F., Kawakami, K., Gaertner, S.L.: Implicit and explicit prejudice and interracial interaction. *J. Pers. Soc. Psychol.* **82**(1), 62–68 (2002)
16. Nisbett, R.E., Wilson, T.D.: Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* **84**(3), 231–259 (1977)
17. Tams, S., Hill, K., Ortiz de Guinea, A., Thatcher, J., Grover, V.: NeuroIS-Alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *J. Assoc. Inf. Syst.* **15**(10), 723–753 (2014)
18. Huang, Y., Kuo, F., Luu, P., Tucker, D., Hsieh, P.: Hedonic evaluation can be automatically performed: An electroencephalography study of website impression across two cultures. *Comput. Hum. Behav.* **49**, 138–146 (2015)
19. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS—Information Systems and the Brain*. Springer-Verlag Berlin Heidelberg (2016)

# The Psychophysiology of Flow: A Systematic Review of Peripheral Nervous System Features

Michael T. Knierim, Raphael Rissler, Verena Dorner,  
Alexander Maedche and Christof Weinhardt

**Abstract** As information systems (IS) are increasingly able to induce highly engaging and interactive experiences, the phenomenon of flow is considered a promising vehicle to understand IS user behavior and to ultimately inform the design of flow-fostering IS. However, despite growing interest of researchers in the phenomenon, knowledge about how to continuously assess flow during IS usage is limited. Hereby, recent developments in NeuroIS and psychophysiology propose novel possibilities to overcome this limitation. This article presents the results of a systematic literature review (SLR) on peripheral nervous system indicators of flow. The findings revealed that currently four major approaches exist towards physiological measurement. Propositions for simple and unobtrusive measurement in IS research are derived in conclusion.

**Keywords** Flow theory · Psychophysiology · Systematic review · NeuroIS

---

M.T. Knierim (✉) · R. Rissler · V. Dorner · A. Maedche · C. Weinhardt  
Karlsruhe Institute of Technology (KIT), Institute of Information Systems and Marketing  
(IISM), Karlsruhe, Germany  
e-mail: michael.knierim@kit.edu

R. Rissler  
e-mail: raphael.rissler@partner.kit.edu

V. Dorner  
e-mail: verena.dorner@kit.edu

A. Maedche  
e-mail: alexander.maedche@kit.edu

C. Weinhardt  
e-mail: christof.weinhardt@kit.edu

R. Rissler  
SAP SE, Walldorf, Germany



## 1 Introduction

In today's digital economy, IS are a significant investment for companies and constitute an indispensable part of employees daily work [1]. Due to technological developments such as multi-media-rich user interfaces, IS are increasingly able to induce highly engaging, interactive, and holistic experiences [1]. One such experience called *flow*—defined as “the holistic sensation that people feel when they act with total involvement” [2, p. 36]—is considered to be of theoretical and practical significance for IS research, helping to explain pre-as well as post-adoptive user behavior [3–5]. As flow becomes more relevant in the business context, understanding how to design IS that induce or foster flow represents a valuable contribution from IS research.

However, despite increasing interest of IS scholars in flow [6], a central challenge is the limited knowledge about real-time measurement. Researchers typically rely on self-report scales which are administered post-task (e.g., [7, 8]). As flow occurs during task execution, post-task self-reported measures cannot assess parameters like the length or depth of flow during task execution, and are subject to reporting inaccuracies [9]. The recent rise of the NeuroIS field with the inclusion and development of psychophysiological measures therefore provides new possibilities for objective and continuous measurements of psychological constructs in the context of IS [10, 11]. Especially towards flow in IS use, the benefit of increasingly reliable measurement [12, 13], but also the design of psychophysiological adaptive IS [14] have been outlined. While previous research on flow-adaptive IS has mainly focused on structured tasks (e.g. gaming and learning [15, 16]) where difficulty levels can be adjusted on the system side, future systems could be extended to more open tasks in business contexts like those of knowledge workers (e.g., software engineers, designers, scientists) through integration of mechanisms that reduce flow-interruption or enhance self-regulation. Yet, the advancement of these lines of research is challenged by the lack of integration of possibilities to physiologically detect flow.

Against this backdrop, we conducted a SLR and examined 20 articles to address the following research question: *What is the state-of-the-art in psychophysiological flow measurement?* This paper makes two key contributions to IS research. First, we systematically review and provide an overview of existing studies utilizing psychophysiological measurements of flow based on the peripheral nervous system. Second, we integrate and synthesize knowledge and provide propositions on how to measure flow using physiological data.

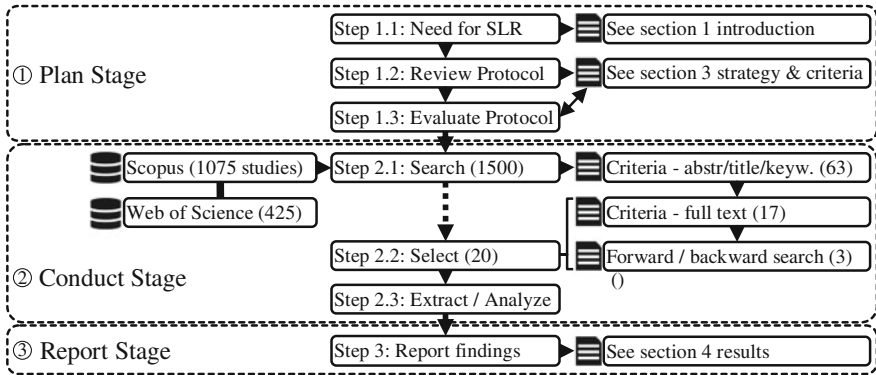
## 2 Theoretical Background

Mihaly Csikszentmihalyi developed a theory of flow in the 1970s [2], positing that flow can be characterized by nine distinct dimensions: (1) challenge–skill balance, (2) clear goals, (3) unambiguous feedback, (4) autotelic experience, (5) action-awareness merging, (6) sense of control, (7) loss of self-consciousness, (8) transformation of time, and (9) concentration on the task at hand. Deriving from these characteristics, rather recently several theoretic propositions have been made how flow is reflected in central (CNS) and peripheral (PNS) nervous system activity. In this study, we focus on the peripheral nervous system as related features are of heightened interest in IS research due to high user acceptance of such measurements. Moreover, there is uncovered potential to distinctly detect flow with PNS features [17, 18].

Due to increased concentration on a task that is appraised as challenging but not threatening and accompanied by positive affective valence, Peifer [9] describes flow to be reflected by optimized physiological activation (i.e., moderate peripheral arousal). Comparably, Keller and colleagues [17] postulate flow to be an experience similar to stress resulting from intense mental effort due to high involvement in an activity and high task difficulty. De Manzano and colleagues [19] describe flow physiology as being reflective of positive affect, increased arousal, and increased mental effort, caused by focused attention on a task. In this line of thought Ullén and colleagues [20] follow the concept of effortless attention, arguing that flow is simultaneously constituted by high attention, increased mental effort, but also a physiological coping mechanism. The latter refers to an increase in relaxing activity of the parasympathetic branch of the autonomous nervous system [18, 20]. Lastly, Léger and colleagues [21] propose that high concentration and attention in flow are reflected by a stable, less volatile state of physiological and affective activation. In summary, the physiology of flow has been described in terms of increased, stable peripheral physiological activation levels, positive affective valence, concurrent calming influences, and mental strain (e.g., stress and mental load). To align differences in these propositions, more research is needed to consolidate empirical findings into a common understanding.

## 3 Method

In order to address our research question, we conducted a SLR according to the guidelines of Kitchenham and Charters [22] as well as Webster and Watson [23]. Overall, we subdivide our systematic review into plan, conduct, and report stages (Fig. 1).



**Fig. 1** Stages of the systematic literature review (SLR)

**Search strategy.** We searched *Web of Science* and *Scopus* [24, 25] with the search string (*flow OR cognitive engagement OR cognitive absorption*) AND (*physiological signal\* OR psychophysiology OR neurophysiology*). The search string was developed in five steps. First, we conducted an exploratory search using Google Scholar with the search term “psychophysiology AND flow”. Second, we reviewed the first 20 search results and identified six highly cited studies [9, 19, 26–29]. Third, we reviewed the full text of these six papers and extracted the terms “neurophysiology” and “physiological signal(s)” as highly relevant to our research question. Fourth, we identified “cognitive engagement” and “cognitive absorption” as relevant flow derivations. Finally, we used Boolean operators to create the final search string. To ensure a holistic search, we have not limited our search to a specific time period.

**Study selection criteria.** All studies that met the following criteria were included: (1) contains an empirical component, (2) is a peer-reviewed journal article, article in press, in conference proceedings, or book chapter, (3) refers to the psychological phenomenon of flow, (4) focuses on the peripheral nervous system. The selection criteria were first applied to abstract, title, and keyword section (excluding 1437 studies). In a further attempt, the criteria were applied to the full text of remaining studies (excluding 46 studies). Finally, a forward and backward search based on the remaining 17 studies was conducted. Thereby, we identified another 3 relevant studies. Overall our SLR identified 20 relevant studies.

## 4 Results

Results are summarized (Table 1) and split into three areas: (1) experiment design parameters, (2) theoretical perspectives, and (3) findings on physiological features.



Table Headers: *Mdl*. Relax Modulation of Activation by Relaxing Influence; *Mod*. Activ. Moderate Physiological Activation; *Sta*. Activ. Stable Physiological Activation; *Pos*. Affect Concurrent Positive Affect; *No Distinct*. No Distinction between Flow/Strain Proposed; *HR* Heart Rate; *IBI* Interbeat Interval; *HRV* Heart Rate Variability; *LF* Low Frequency HRV; *HF* High Frequency HRV; *LF/HF* Low Frequency to High Frequency HRV Relation; *FW-HM* Full-Width-At-Half-Maximum of Pulse Pressure; *MinP* Minimum Pulse Wave Amplitude; *MaxP* Maximum Pulse Wave Amplitude; *PRtg*. MaxP—MinP; *EDA* Electrodermal Activity; *SCL* Skin Conductance Level (Tonic); *SD-SCL* Standard Deviation of SCL; *RR* Respiratory Rate; *RDT* Thoracic Respiratory Depth; *RDA* Abdominal Respiratory Depth; *RC-YC* Respiratory Cycle; *CoLe* Cortisol Level; *Co-Rea* Cortisol Reactivity; *ZM* Zygomaticus Major; *CS* Corrugator Supercilii; *OO* Orbicularis Oculi; Dependent Variables/Measures: *F9D/3D* FSS 9 Item Subset; *JGEQ 14/36* Flow Subscale of Short/Long Game Experience Questionnaire; *RT* Relaxation/Task Conditions; *FKS* Flow Short Scale; *FKS A/F* FKS Subscales Absorption/Fluency; *B/F/O* Boredom/Fit/Overload Difficulty Conditions; *CA* Cognitive Absorption Scale; *CSPE* Combined Indicators Challenge-Skill Balance and Positive Emotion; *FSS* Flow State Scale; *R/F/S* Relaxation/Facebook(Flow)/Stress; *F1* Flow Index; *B/F/I* Boredom/Flow/Immersion Conditions; *BS/AWS* Between/Within Subject Design

**Explanation of the table.** To illustrate findings and our depiction thereof, consider the first study in Table 1 by Harmat and colleagues [18]. They conducted an experiment using the digital game Tetris and evaluated flow with a subset of the Flow State Scale (F9D). Insignificant relationships between heart rate (HR) and the self-report scale (●) were found. Moreover, a positive linear relationship between thoracic respiratory depth and the self-report scale ((↗)) were found. Ulrich and colleagues [30] used an arithmetic task and found an inverted U-shaped relationship ((^)) between skin conductance levels (SCL) and difficulty-manipulated task conditions termed boredom, fit, and overload (B/F/O). Partial findings like the inverted U-shaped relationship between low frequency heart rate variability (LF-HRV) in the first half of the experiment by Peifer and colleagues [31] are denoted with an asterisk ((^\*)).

**Experimental design** (*sample sizes, flow induction tasks, and dependent variables/measures*). Sample sizes vary strongly across studies ranging from seven to 77 experiment participants. Our sample counts include reported, usable observations only. Second, the majority of studies in our SLR (14/20) used games. This is important because designing tasks that reliably induce flow states is still a major challenge in flow research [32] and game paradigms have been criticized as to not sufficiently induce straining experiences [31]. Depending on research goals (e.g., in case of separating flow from stress experiences), this spectrum might be important for flow research in IS. Utile alternatives include high involvement tasks (e.g., [33]). Third, dependent variables differ in two operationalization formats, that are self-reports (14/20) and experiment conditions (8/20), with some studies utilizing both (4/20). Conditions are most often differentiated along the dimension of task difficulty. In total, nine different self-report instruments were used in 15 studies to measure flow.

**Theoretical perspectives** (*modulation by relaxing influence, moderate activation, stable activation, positive affect, no distinction*). This area refers to the theoretical propositions of how flow can physiologically be differentiated from strain (e.g., stress). We refer to perspectives and order studies in terms of diagnosticity (i.e., how proposed physiology patterns are to isolate flow from other states) [11]. While increased peripheral physiological arousal is a common denominator in both flow and strain [9], four central distinguishing patterns are described towards flow: (1) modulation of arousal by relaxing influences (Mdl. Relax), (2) moderate instead of high levels of arousal (Mod. Activ.), (3) stable, less volatile arousal (Sta. Activ.), and (4) concurrent presence of arousal and positive affect (Pos. Affect). The abbreviations in parentheses refer to how these perspectives are denoted in Table 1. To our knowledge, the first is currently the only characterization that sufficiently distinguishes flow from strain by the phenomenon of non-reciprocal co-activation of sympathetic and parasympathetic branches of the autonomic nervous system [18]. Three types of studies were derived. The first (high diagnosticity—sufficiency condition fulfilled) propose distinct physiological signatures by investigating arousal modulation by relaxation (6/20). The second (moderate diagnosticity—necessity condition fulfilled) propose indicative physiological signatures (6/20).

The third (low diagnosticity) propose either indistinct physiological signatures (2/20), or do not include hypotheses towards flow physiology specifically (6/20).

**Physiological findings** (*cardiac, pulse, electrodermal, respiration, hormonal, facial muscle, and pupillary reactions*). In summary, cardiac features are used most often (12/20), especially in the class of higher diagnostic studies (6/6). This is probably due to the property of the cardiovascular system to reflect both sympathetic and parasympathetic activation [34]. Therefore, distinguishing flow from strain is enabled by comparison of arousal levels, arousal variability or the isolated activity of sympathetic and parasympathetic activation. EDA is the second-most used feature (10/20), albeit mainly in studies with lower diagnosticity (7/8). EMG measures are used mainly in valence-related studies across classes (7/20). Support has been found for all four outlined theoretical propositions, with (1) being mainly related to sympathetic and parasympathetic (HF-HRV, RDT) autonomic activity, (2) being most often related to moderate cortisol (CoLe), skin conductance (SCL) and heart rate variability (Total HRV and LF-HRV) levels, (3) being related to skin conductance and hormonal level reactivity, and (4) being most often related to increased facial muscle activity (ZM).

## 5 Discussion, Future Directions and Conclusion

This literature review identified four central approaches to the physiological measurement of flow. All include increased levels of arousal, yet vary in their explanation to how arousal states differ from straining experiences such as stress. Of these four, three fulfill only necessity conditions to distinguish flow. The proposition of a non-reciprocal co-activation of sympathetic and parasympathetic nervous system in flow [18, 20] also fulfills sufficiency conditions. This approach would also lend itself to rather simple and unobtrusive measurement, comprised of ECG and EDA instruments that can be used as reliable indicators of parasympathetic [35] and sympathetic [36] activity. We also classified studies as more diagnostic that combine propositions (1)–(4). Support has been found for all directions through different physiological features, which is why we propose that ideally multiple propositions be taken into further investigation. NeuroIS research can especially contribute to the state of knowledge by advancing the line of research on these diagnostically higher perspectives. An exemplary approach in this direction is reported by Bian and colleagues [37]. Furthermore, NeuroIS researchers should pay attention to task selection and be aware of limitations of dependent variables. The inclusion of multiple criteria, for example dedicated, encompassing self-reports like FSS [7] or FKS [8] in conjunction with multiple task conditions (e.g., strongly varied in difficulty and coupled with intrinsic involvement [32]) are advised. In following these propositions, NeuroIS research can benefit from finding means of increased objective validity in flow measurement and also advance constructivist efforts to facilitate flow through adaptive IS. As initially mentioned, this could extend current efforts on flow-adaptive IS in structured tasks (e.g., gaming and

learning [15, 16]), to more open tasks in business-contexts like those of knowledge workers (e.g., software engineers, designers, scientists) where the experience and protection of flow states has been shown to be highly beneficial to performance and satisfaction outcomes. One example of how such systems could support users' flow is through the reduction of IT-mediated interruptions [38], that have been found as a primary disruptive factor in work settings [39]. In this direction, physiological measurements could be used by a system to determine when and how IT-mediated interruptions should be delivered to users. Furthermore, physiological data could be used to inform IS users about their mental state in order to facilitate flow experience through self-regulation. This proposition is derived from findings on HRV-biofeedback, which show that individuals can use biofeedback tools to modulate parasympathetic nervous system activity [40]. Similarly, with a wide array of business activities increasingly being conducted in small groups, biofeedback system integrations might be used to support interpersonal state regulation. This could also incorporate research on the rather novel concept of group flow [41, 42]. In this direction, central questions would have to be considered like whether group flow, an experience supposedly distinct from aggregated individual flow [41], would have a specific physiological signature within the individual (e.g., a comparatively more joyful experience than solitary flow [41]), or physiological group signatures with distribution characteristics like homogenous activation levels [43] or temporal characteristics like simultaneous or sequential physiological changes [13, 43]. As this review has pointed out that some PNS markers might be well suited to isolate flow experiences in the individual, they pose similar utility to further investigate these questions regarding group flow. In summary, even though more research is required to substantiate what is known on the PNS psychophysiology of flow, the findings of this review can be utilized to inform NeuroIS research and the investigation of several novel possibilities to modulate or protect an individual's flow with the help of adaptive IS.

## References

1. Agarwal, R., Karahanna, E.: Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. *Manag. Inf. Syst. Q.* **24**, 665–694 (2000)
2. Csikszentmihalyi, M.: *Beyond Boredom and Anxiety*. Jossey-Bass, San Francisco, CA (1975)
3. Webster, J., Trevino, L.K., Ryan, L.: The dimensionality and correlates of flow in human-computer interactions. *Comput. Human Behav.* **9**, 411–426 (1993)
4. Ghani, J.A., Supnick, R., Rooney, P.: The experience of flow in computer-mediated and in face-to-face groups. In: *Proceedings of the 12th International Conference on Information Systems*, pp. 229–237 (1991)
5. Mahnke, R., Benlian, A., Hess, T.: Flow experience in information systems research: revisiting its conceptualization, conditions, and effects. In: *Proceedings of the 35th International Conference on Information Systems*, pp. 1–22 (2014)
6. Rissler, R., Nadj, M., Adam, M.T.P.: Flow in information systems research: review, integrative theoretical framework, and future directions. In: *International Conference on Wirtschaftsinformatik*, pp. 1051–1065 (2017)



7. Jackson, S.A., Eklund, R.C.: Assessing flow in physical activity: the flow state scale-2 and dispositional flow scale-2. *J. Sport Exerc. Psychol.* **24**, 133–150 (2002)
8. Engeser, S., Rheinberg, F.: Flow, performance and moderators of challenge–skill balance. *Motiv. Emot.* **32**, 158–172 (2008)
9. Peifer, C.: Psychophysiological correlates of flow-experience. In: Engeser, S. (ed.) *Advances in Flow Research*, pp. 139–164. Springer Science, New York (2012)
10. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS Information Systems and the Brain*. Springer, Berlin Heidelberg (2016)
11. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**, i–xxxv (2014)
12. Bastarache-Roberge, M.-C., Léger, P.-M., Courtemanche, F., Sénécal, S., Frédette, M.: Measuring flow using psychophysiological data in a multiplayer gaming context. In: Davis, F. D. (ed.) *Information Systems and Neuroscience, Lecture Notes in Information Systems and Organisation*, pp. 187–191 (2015)
13. Labonté-Lemoine, É., Léger, P.-M., Resseguier, B., Bastarache-Roberge, M.-C., Frédette, M., Sénécal, S., Courtemanche, F.: Are we in flow? neurophysiological correlates of flow states in a collaborative game. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 1980–1988 (2016)
14. Adam, M.T.P., Gimpel, H., Maedche, A., Riedl, R.: Designing stress-sensitive adaptive enterprise systems: theoretical foundations and design blueprint. In: *Gmunden Retreat on NeuroIS*, pp. 1–22 (2014)
15. Afegan, D., Peck, E.M., Solovey, E.T., Jenkins, A., Hincks, S.W., Brown, E.T., Chang, R., Jacob, R.J.K.: Dynamic difficulty using brain metrics of workload. In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems—CHI’14*, pp. 3797–3806 (2014)
16. Liu, C., Agrawal, P., Sarkar, N., Chen, S.: Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *Int. J. Hum. Comput. Interact.* **25**, 506–529 (2009)
17. Keller, J.: The flow experience revisited: the influence of skills-demands-compatibility on experiential and physiological indicators. In: *Flow Experience*, pp. 351–374. Springer International Publishing, Cham (2016)
18. Harmat, L., de Manzano, Ö., Theorell, T., Högman, L., Fischer, H., Ullén, F.: Physiological correlates of the flow experience during computer game playing. *Int. J. Psychophysiol.* **97**, 1–7 (2015)
19. De Manzano, O., Theorell, T., Harmat, L., Ullén, F.: The psychophysiology of flow during piano playing. *Emotion* **10**, 301–311 (2010)
20. Ullén, F., De Manzano, Ö., Theorell, T., Harmat, L.: The physiology of effortless attention: correlates of state flow and flow proneness. In: Bruya, B. (ed.) *Effortless Attention: A New Perspective in the Cognitive Science of Attention and Action*, pp. 205–217. MIT Press, Cambridge (2010)
21. Léger, P.M., Davis, F.D., Cronan, T.P., Perret, J.: Neurophysiological correlates of cognitive absorption in an enactive training context. *Comput. Human Behav.* **34**, 273–283 (2014)
22. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering. *Tech. Rep. EBSE*. (2007)
23. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: writing a literature review. *Manag. Inf. Syst. Q.* **26**, xiii–xxiii (2002)
24. Bandara, W., Furtmueller, E., Gorbacheva, E., Miskon, S., Beekhuizen, J.: Achieving rigor in literature reviews: insights from qualitative data analysis and tool-support. *Commun. Assoc. Inf. Syst.* **37**, 154–204 (2015)
25. Hamari, J., Koivisto, J., Sarsa, H.: Does gamification work?—a literature review of empirical studies on gamification. In: *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 3025–3034 (2014)

26. Mauri, M., Cipresso, P., Balgera, A., Villamira, M., Riva, G.: Why is facebook so successful? psychophysiological measures describe a core flow state while using facebook. *Cyberpsychology, Behav. Soc. Netw.* **14**, 723–731 (2011)
27. Tozman, T., Magdas, E.S., MacDougall, H.G., Vollmeyer, R.: Understanding the psychophysiology of flow: a driving simulator experiment to investigate the relationship between flow and heart rate variability. *Comput. Human Behav.* **52**, 408–418 (2015)
28. Kivikangas, J.M., Puttonen, S.: *Psychophysiology of flow experience: an explorative study* (2006)
29. Nacke, L.E., Lindley, C.A.: Flow and Immersion in first-person shooters: measuring the player's gameplay experience. In: *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, pp. 81–88. ACM Press, New York (2008)
30. Ulrich, M., Keller, J., Grön, G.: Neural signatures of experimentally induced flow experiences identified in a typical fMRI block design with BOLD imaging. *Soc. Cogn. Affect. Neurosci.* **11**, 496–507 (2016)
31. Peifer, C., Schulz, A., Schächinger, H., Baumann, N., Antoni, C.H.: The relation of flow-experience and physiological arousal under stress—Can u shape it? *J. Exp. Soc. Psychol.* **53**, 62–69 (2014)
32. Tozman, T., Peifer, C.: Experimental paradigms to investigate flow-experience and its psychophysiology: inspired from stress theory and research. In: Harmat, L., Orsted, F., Ullén, F., Wright, J., and Sadlo, G. (eds.) *Flow Experience*, pp. 329–350. Springer, Heidelberg (2016)
33. Tozman, T., Zhang, Y.Y., Vollmeyer, R.: Inverted U-shaped function between flow and cortisol release during chess play. *J. Happiness Stud.* **18**, 247–268 (2017)
34. Berntson, G.G., Quigley, K.S., Lozano, D.: Cardiovascular psychophysiology. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (eds.) *Handbook of Psychophysiology*, pp. 182–210. Cambridge University Press, Cambridge (2007)
35. Porges, S.W.: *Orienting in a defensive world: mammalian modifications of our evolutionary heritage. A polyvagal theory* (1995)
36. Boucsein, W.: *Electrodermal Activity*. Springer Science+ Business Media, New York, Dordrecht, Heidelberg, London (2012)
37. Bian, Y., Yang, C., Gao, F., Li, H., Zhou, S., Li, H., Sun, X., Meng, X.: A framework for physiological indicators of flow in VR games: construction and preliminary evaluation. *Pers. Ubiquitous Comput.* **20**, 821–832 (2016)
38. Addas, S., Pinsonneault, A.: The many faces of information technology interruptions: a taxonomy and preliminary investigation of their performance effects. *Inf. Syst. J.* **25**, 231–273 (2015)
39. Parent-Thirion, A., Vermeylen, G., van Houten, G., Wilkens, M., Wilczynska, A.: *Sixth European working conditions survey (EWCS)*. Office for Official Publications of the European Communities, Luxembourg, LU (2015)
40. Lehrer, P.M., Gevirtz, R.: Heart rate variability biofeedback: how and why does it work? *Front. Psychol.* **5**, 1–9 (2014)
41. Walker, C.J.: Experiencing flow: is doing it together better than doing it alone? *J. Posit. Psychol.* **5**, 3–11 (2010)
42. Keith, M., Anderson, G., Dean, D., Gaskin, J., Wiser, T., Gremmert, S., Payne, K.: The effects of video gaming on work group performance. In: *Proceedings of the 37th International Conference on Information Systems*, pp. 1–20 (2016)
43. Noy, L., Levit-Binun, N., Golland, Y.: Being in the zone: physiological markers of togetherness in joint improvisation. *Front. Hum. Neurosci.* **9**, 1–14 (2015)
44. Murch, W.S., Chu, S.W.M., Clark, L.: Measuring the slot machine zone with attentional dual tasks and respiratory sinus arrhythmia. *Psychol. Addict. Behav.* 1–11 (2017)
45. Peifer, C., Schächtinger, H., Engeser, S., Antoni, C.H.: Cortisol effects on flow-experience. *Psychopharmacology* **232**, 1165–1173 (2015)

46. Gaggioli, A., Cipresso, P., Serino, S., Riva, G.: Psychophysiological correlates of flow during daily activities. In: Wiederhold, B.K. and Riva, G. (eds.) *Annual Review of Cybertherapy and Telemedicine 2013*, pp. 65–69. IOS Press (2013)
47. Keller, J., Bless, H., Blomann, F., Kleinböhl, D.: Physiological aspects of flow experiences: skills-demand-compatibility effects on heart rate variability and salivary cortisol. *J. Exp. Soc. Psychol.* **47**, 849–852 (2011)
48. Nacke, L.E., Grimshaw, M.N., Lindley, C.A.: More than a feeling: measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interact. Comput.* **22**, 336–343 (2010)
49. Drachen, A., Pedersen, A.L.: Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In: *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*, pp. 49–54 (2009)
50. Drachen, A., Nacke, L.E., Yannakakis, G., Pedersen, A.L.: Psychophysiological correlations with gameplay experience dimensions. In: *Brain, Body and Bytes, Workshop—CHI 2010*, pp. 1–4 (2010)
51. Nacke, L.E., Lindley, C.: Affective ludology, Flow and immersion in a first-person shooter: measurement of player experience. *Loading* **3**, 1–21 (2010)
52. Klarkowski, M., Johnson, D., Wyeth, P., Phillips, C., Smith, S.: Psychophysiology of challenge in play: EDA and self-reported arousal. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 1930–1936. ACM Press, New York (2016)

# Predicting Properties of Cognitive Pupillometry in Human–Computer Interaction: A Preliminary Investigation

Pierre-Majorique Léger, Patrick Charland, Sylvain Sénécal  
and Stéphane Cyr

**Abstract** This paper aims to investigate the predictive property of pupil dilation in an IT-related task. Previous work in the field of cognitive pupillometry has established that pupil size is associated with cognitive load. We conducted a within-subject experiment with 22 children aged between 7 and 9. For the hard questions, visit duration, pupil size and its quadratic effect were significant predictors. We discuss the potential of using this unobtrusive approach for neuro-adaptive and auto-adaptive applications.

**Keywords** Eye-tracking · Pupillometry · Cognitive load · HCI—learning

## 1 Introduction

Over the last few years, there has been growing interest in the concept of cognitive workload in the field of information technology (IT) and human–computer interaction (HCI) [1–5]. Cognitive workload is defined as the information processing load placed on a human in a particular task [6]. The advent of neuroscientific methods in the field of Neuro-information-Systems (NeuroIS) has opened the possibility of measuring this construct using implicit measures to capture the unconscious and automatic nature of the cognitive workload [7].

---

P.-M. Léger (✉) · S. Sénécal  
HEC Montréal, QC, Canada  
e-mail: pml@hec.ca

S. Sénécal  
e-mail: ss@hec.ca

P. Charland · S. Cyr  
Université du Québec à Montréal, QC, Canada  
e-mail: charland.patrick@uqam.ca

S. Cyr  
e-mail: cyr.stephane@uqam.ca

Previous work in the field of cognitive pupillometry has established that pupil size is associated with cognitive load [8–10]. A wide range of task-evoked pupillary response experiments have demonstrated a relationship between pupil size and mental processing demands in various contexts including computerized tasks [11, 12]. Pupil size is a metric that can be unobtrusively acquired by standard eye-tracking technology during an interaction with a computerized interface, while preserving the ecological validity of the task.

This paper aims to investigate the predictive property of pupil dilation in an IT-related task. Specifically, building on cognitive load theory, we are using pupil size to predict task performance in a problem-solving context. To answer our research question, we conducted a laboratory experiment with 22 children who had to perform a mathematics task on a tablet. The results section elaborates on the prediction possibilities of the right answer within pupillometric data. We conclude with a discussion on the potential of using this unobtrusive approach for neuro-adaptive and auto-adaptive applications.

## 2 Prior Research

Cognitive load theory (CLT) [13–16], is based on the premise that working memory and information treatment capacity are limited, at a given time, for certain tasks [17]. A task which leads to a high cognitive load would thus overload working memory and affect the performance and/or efficiency of the individual.

Until recently, the concept of workload in information system research has been studied using psychometric scales. In fact, the principal measurement method was based on self-reported scales [18, 19], where the participant must evaluate his level of mental effort during a task. The best-known measures are those developed by [14], as well as the TLX scale developed by NASA [20]. The limits of the self-reported questionnaires are well known in experimental psychology [21, 22], including those associated with the primacy (the subject remembers the beginning of the task) and recency effects (the subject remembers the end of the task) as well as the social desirability bias (the subject seeks to please the experimenter [23]).

However, the recent application of methods developed in neuroscience to research in the field of education implies resorting to new instruments that are complementary to the traditionally used ones. Methods such as magnetic resonance imagery (MRI), electroencephalography (EEG), eye-tracking (study of eye movement) or pupillometry (study of the diameter of the pupil) enable direct data collection in real time which informs researchers about the learners' unconscious cognitive processes [1]. These methods make it possible to establish inferences likely to inform researchers about a subject's cognitive load during a learning activity, and in real time.

Research undertaken by [24] paves the way for the establishment of a predictive link between the level of difficulty of addition exercises and the cognitive load measured by EEG. Though these findings are promising, having to rely on EEG

implies a cumbersome technical set-up as well as an analysis which is often costly. In parallel, [17] observe that many researchers in the field of educational studies have recently been using pupillometry, a real-time measure of the diameter of the pupil using an eye-tracker, and [14] consider pupillometry to be a very precise technique to measure variations in levels of cognitive load. The non-intrusive nature of this technique makes it suitable for research on children learning mathematics, as it does not require the installation of sensors on their scalp.

To the best of our knowledge, no study focusing on the cognitive load of children measured by pupillometry in the context of mathematics learning (arithmetical operations) has yet been published. It should be noted, however, that this method has well-known limitations, such as the conditions which affect variations in pupil size (light and distance of objects observed). Many researchers in a variety of fields were able to account for these limitations with robust experimental designs allowing for the control of these conditions, and [11, 25] propose different methods to control these variables at the data collection stage.

### 3 Methods

We conducted a within-subject experiment with 22 children aged between 7 and 9 (60% male). We used a Tobii x60 oculometer to measure and record the participants' pupil dilation as well as their eye movement patterns. Each parent and child participant received a \$50 compensation and an educational book. The experiment was approved by the Institution's Ethical Review Board.

Participants had to answer 2nd grade arithmetic questions (additions) on a commercially available iPad application dedicated to the learning of mathematics. For each question, they could select their answer from among three (3) to four (4) multiple choice options. The questions were also classified according to their level of difficulty by three experts in mathematics. Each question's level of difficulty was assessed as easy or difficult based on several criteria. A question was deemed easy if it was associated with a lower educational level and if it was contextualized or illustrated. A question was deemed hard if it was associated with a higher educational level, if it involved addition and subtraction with double digits, and if it was neither contextualized nor illustrated.

Specifically, each participant was asked to answer a total of five (5) randomized blocks of exercises containing six (6) mathematical questions (see Fig. 1). Pupil size and gaze visit duration were recorded on an area of interest that corresponded to the right answer.

As this was a real-life application, the presentation of the experiment's questions was subject to some constraints. In our model, we thus control these possible confounding effects. Square and horizontal correspond to the orientation of the multiple choice questions. Answer position corresponds to whether the right answer was in the first, second, third, or fourth position.

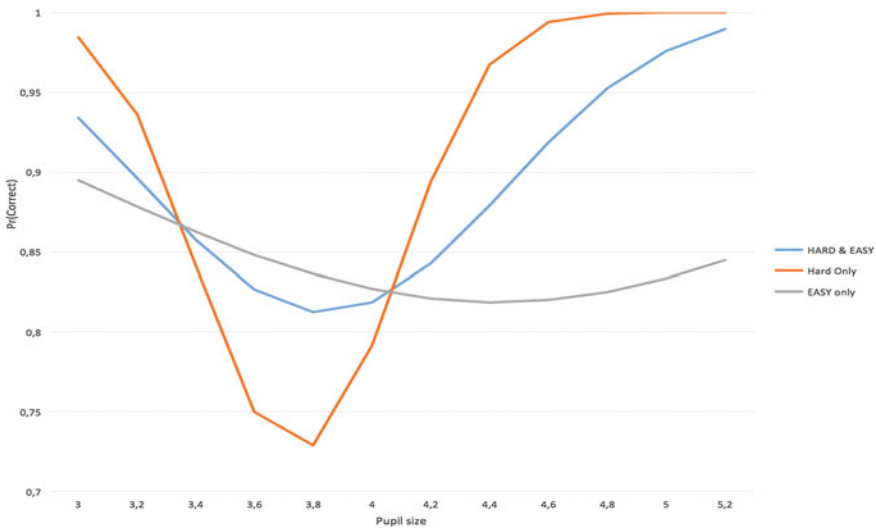


Fig. 1 Predicted probability of having correct answers

## 4 Results

A total of 255 observations were usable for analytical purposes; 191 of them were from easy questions, while 64 were from hard questions. A mixed-model logistic regression with a random intercept was used for the prediction model (SAS PROC GLIMMIX) [26]. Table 1 presents models that make it possible to predict the right answer with both the hard and easy questions, only with the hard questions and, finally, only with the easy questions.

When both types of questions are analyzed jointly, two variables predict the choice of the right answer: the time spent looking at the right answer (Visit duration,  $B = -1.88$ ,  $p = 0.001$ ) and the orientation of the answer choices (Horizontal,  $B = -1.03$ ,  $p = 0.015$ ). However, it has to be noted that there were no hard questions with horizontal presentation in the usable observations. For the hard questions, visit duration ( $B = -1.10$ ,  $p = 0.049$ ), pupil size ( $B = -38.04$ ,  $p = 0.099$ ) and its quadratic effect ( $B = 5.11$ ,  $p = 0.088$ ) were significant predictors. It should be noted that as expected, consistent with previous work [27, 28], there is a quadratic effect of pupil size on the chance of getting the right answer. Finally, for the easy questions, the orientation of the answer choices (Horizontal,  $B = -1.18$ ,  $p = 0.010$ ) and the visit duration on the right answer ( $B = -2.29$ ,  $p = 0.001$ ) were significant predictors. Hence, visit duration of gaze on the right answer had a negative effect in all models.

These results are presented graphically in Fig. 1 in order to provide a better illustration of the quadratic effect of the cognitive load on the participant's ability to choose the right answer to hard questions. We observed that participants who

**Table 1** Right answer prediction

Effect	Hard and Easy		Hard only		Easy only	
	Estimate	Pr >  t	Estimate	Pr >  t	Estimate	Pr >  t
Intercept	9.38	0.312	73.28	0.111	9.26	0.343
Square	-0.40	0.617	0.00		-0.56	0.490
Horizontal	-1.03	0.015	0.00		-1.18	0.010
Hard	73.10	0.272				
Answer position	-0.02	0.931	-0.39	0.379	0.08	0.732
Pupil size	-2.76	0.523	-38.04	0.099	-2.57	0.572
Visit duration	-1.88	<0.0001	-1.10	0.049	-2.29	<0.0001
Hard×Pupil	-39.79	0.247				
Pupi×Pupil	0.31	0.541	5.11	0.088	0.29	0.589
Hard×Pupil×Pupil	5.36	0.224				

invested sufficient cognitive resources in the task had more chances of getting the right answer to hard questions. The first part of the curve (low effort/high probability) should be understood with the concept of prior or integrated knowledge. For example, most of the pupils have already integrated the automatic response to a simple operation such as “2 + 2 = 4”. When they are presented with this kind of very easy stimuli, their response is unconscious and automatic and they don’t invest much mental effort to the task in order to come up with the right answer.

## 5 Discussion and Concluding Remarks

In this paper, we tested the predictive properties of pupil size with an ecologically valid tablet-based arithmetic task. We found that in the case of difficult arithmetic problems, pupil dilation at the moment of fixation on the right answer can contribute to predicting the extent to which the child is investing sufficient cognitive resources to successfully solve the problem.

New laptop computers integrating embedded low-end eye-tracking functions are now commercially available.<sup>1</sup> It is very likely that eye-tracking will become a mainstream input device on many computers over the next few years. It is important to explore this field further in order to gain a better understanding of the way in which eye-tracking characteristics, such as pupil diameter, can be used in neuro-adaptive and auto-adaptive devices. These results are important in order to eventually develop auto-adaptive learning environments where the difficulty level of tasks and assessments could be determined based on the user’s cognitive data, such as pupil diameter.

---

<sup>1</sup><http://gizmodo.com/msi-s-eye-tracking-laptop-is-the-future-but-not-the-pr-1758485727>.



Our next step will be to test this protocol with an adult population which we shall subject to a more complex IT task. For example, building upon previous work [29], we are currently preparing to auto-adapt an information dashboard for a monitoring task in an ERP simulation, based on a workload index [30, 31].

## References

1. Léger, P.M., Sénécal, S., Courtemanche, F., de Guinea, A.O., Titah, R., Fredette, M., Labonté-LeMoyné, É.: Precision is in the eye of the beholder: application of eye fixation-related potentials to information systems research. *J. Assoc. Info. Syst.* **15**(10), 651 (2014)
2. Loos, P., et al: NeuroIS: neuroscientific approaches in the investigation and development of information systems. *Bus. Info. Syst. Eng.* **2**(6):395–401 (2010)
3. Dumont, L., et al., Using transcranial direct current stimulation (tDCS) to assess the role of the dorsolateral prefrontal cortex in technology acceptance decisions: a pilot study. *Proc. Gmunden Retreat NeuroIS* (2014)
4. Randolph, A.B., Labonté-LeMoyné, É., Léger, P.M., Courtemanche, F., Sénécal, S., Fredette, M.: Proposal for the use of a passive BCI to develop a neurophysiological inference model of IS constructs. In: *Information Systems and Neuroscience*, pp. 175–180. Springer, Heidelberg (2015)
5. Mirhoseini, S.M.M., Léger, P.-M., Sénécal S.: The influence of task characteristics on multiple objective and subjective cognitive load measures. In: *Information Systems and Neuroscience* pp. 149–156. Springer, Heidelberg, (2017)
6. Sheridan, T.B., Stassen, H.G.: Definitions, models and measures of human workload. In: Moray, N., (eds.) *Mental Workload*, pp. 219–234. Plenum; New York (1979)
7. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS*. Springer, Berlin (2016)
8. Hess, E.H., Polt, J.M.: Pupil size in relation to mental activity during simple problem solving. *Science* **143**(3611), 1190–1192 (1964)
9. Kahneman, D., Beatty, J.: Pupil diameter and load on memory. *Science* **154**(3756), 1583–1585 (1966)
10. Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* **91**(2), 276 (1982)
11. Klingner, J., Tversky, B., Hanrahan, P.: Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology* **48**(3), 323–332 (2011)
12. Piquado, Tepring, Isaacowitz, Derek, Wingfield, Arthur: Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology* **47**(3), 560–569 (2010)
13. Paas, F.G., Van Merriënboer, J.J.: Instructional control of cognitive load in the training of complex cognitive tasks. *Edu. Psychol. Rev.* **6**(4), 351–371 (1994)
14. Paas, F., et al.: Cognitive load measurement as a means to advance cognitive load theory. *Edu. Psychol.* **38**(1), 63–71 (2003)
15. Paas, F., Van Gog, T., Sweller, J.: Cognitive load theory: new conceptualizations, specifications, and integrated research perspectives. *Edu. Psychol. Rev.* **22**(2), 115–121 (2010)
16. Clark, R.C., Nguyen, F., Sweller, J.: *Efficiency in Learning: Evidence-based Guidelines to Manage Cognitive Load*. Wiley (2011)
17. Kruger, J.L., Doherty, S. Measuring cognitive load in the presence of educational video: towards a multimodal methodology. *Australas. J. Edu. Technol.* **32**(6) (2016)
18. Sweller, J., Van Merriënboer, J.J., Paas, F.G.: Cognitive architecture and instructional design. *Edu. Psychol. Rev.* **10**(3), 251–296 (1998)

19. Van Mierlo, C.M., Jarodzka, H., Kirschner, F., Kirschner, P.A.: Cognitive load theory in e-learning. In: *Encyclopedia of Cyber Behavior*, pp. 1178–1211. IGI Global (2012)
20. Rubio, S., Díaz, E., Martín, J., Puente, J.M.: Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and workload profile methods. *Appl. Psychol.* **53**(1), 61–86 (2004)
21. Postman, L., Phillips, L.W.: Short-term temporal changes in free recall. *Q. J. Exp. Psychol.* **17**(2), 132–138 (1965)
22. Glanzer, M., Cunitz, A.R.: Two storage mechanisms in free recall. *J. Verbal Learn. Verbal Behav.* **5**(4), 351–360 (1966)
23. Beauvois, J.L. (2003). Judgment norms, social utility, and individualism. *Sociocognitive Approach Social Norms*, 123–147
24. Spüler, M., Walter, C., Rosenstiel, W., Gerjets, P., Moeller, K., Klein, E.: EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning. *ZDM* **48**(3), 267–278 (2016)
25. Klingner, J.: Fixation-aligned pupillary response averaging. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 275–282). ACM (2010, March)
26. Littell, R.C., et al.: *SAS for Mixed Models*, 2nd edn., Cary, NC. SAS Institute, 2006
27. Berka, C., et al.: EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviat. Space Environ. Med.* **78**(5), B231-B244 (2007)
28. Charland, P., Williot, A., Léger, P.-M., Mercier, J., Skelling, Y., Lapierre, H.-G. Problem solving in physics: a pupillometric perspective. In: *Proceedings of the XVIIIth International Congress of the World Association of Research in Education*, Eskisehir, Turkey, 1 June 2016
29. Léger, P.-M., et al.: Neurophysiological correlates of cognitive absorption in an enactive training context. *Comput. Human Behav.* **34**, 273–283 (2014)
30. Léger, Pierre-Majorique: Using a simulation game approach to teach enterprise resource planning concepts. *J. Info. Syst. Edu.* **17**(4), 441 (2006)
31. Léger, P.M., et al.: *ERPsims, ERPsims Lab (erpsim.hec.ca)*. HEC Montréal, QC (2007)

# Human Versus Machine: Contingency Factors of Anthropomorphism as a Trust-Inducing Design Strategy for Conversational Agents

Anna-Maria Seeger and Armin Heinzl

**Abstract** Conversational agents are increasingly popular in various domains of application. Due to their ability to interact with users in human language, anthropomorphizing these agents to positively influence users' trust perceptions seems justified. Indeed, conceptual and empirical arguments support the trust-inducing effect of anthropomorphic design. However, an opposing research stream that has widely been overlooked provides evidence that human-likeness reduces agents' trustworthiness. Based on a thorough analysis of psychological mechanisms related to the contradicting theoretical positions, we propose that the agent substitution type acts as a situational moderator variable on the positive relationship between anthropomorphic design and agents' trustworthiness. We argue that different agent types are related to distinct user expectations that influence the cognitive evaluation of anthropomorphic design. We further discuss how these differences translate into neurophysiological responses and propose an experimental set-up using a combination of behavioral, self-reported and eye-tracking data to empirically validate our proposed model.

**Keywords** Conversational agents · Trustworthiness · Anthropomorphism  
Eye-tracking

## 1 Introduction

Substantial advances in artificial intelligence make conversational technology increasingly relevant. Conversational agents are software systems that are able to process, understand and produce natural language interactions [1, 2]. These systems are also referred to as chatbots or intelligent virtual assistants [3]. Business analysts

---

A.-M. Seeger (✉) · A. Heinzl  
University of Mannheim, Mannheim, Germany  
e-mail: seeger@uni-mannheim.de

A. Heinzl  
e-mail: heinzl@uni-mannheim.de

are expecting conversational agents to revolutionize the way humans interact with information systems in various fields of application [4, 5]. Conversational agents promise to provide convenient, instant and accurate responses to a wide range of user inquiries on the basis of natural language. Users are expected to benefit from greater accessibility and more intuitive interactions. Providers are expected to benefit by reducing costs and improving quality of standardized and recurring tasks [6]. An agent that is able to satisfy users' expectations, thus, creates a win-win situation. Real world use cases include shopping bots that assist consumers in finding and purchasing desired products and services, health assistant bots that provide patients with personalized health information and guidance as well as enterprise software bots that enable professionals to interact with enterprise systems such as Customer-Relationship-Management (CRM). The successful design of conversational agents, however, is contingent upon an understanding of users' expectations and perceptions in order to assure that users are willing to rely on these agents. Therefore, users' trust is a prerequisite for successful adoption. Indeed, numerous information systems (IS) studies have addressed the role of trust for technology acceptance and use [7–9].

Two opposing theoretical positions exist that explain human-agent trust by adopting either a human-human or a human-machine trust perspective [17]. Table 1 provides an overview of the two perspectives. The Computers are Social Actors (CASA) paradigm [10, 19, 20] is a prominent conceptual basis for research interested in understanding how to make computer agents more trustworthy. Studies in this tradition adopt the human-human trust perspective. CASA research builds upon the media-equation hypothesis that proposes that humans place social expectations, norms and beliefs on computers [10, 11]. This stream of research produced experimental evidence indicating that *anthropomorphism*—the extent to which computational systems are perceived to have human characteristics—increases users' trust into computer agents [21–24]. Inspired by these findings, designers could conclude that making conversational agents more human-like is essential to create a sustainable trust relationship between agents and users. But is human-likeness of conversational agents really unconditionally beneficial for agents' trustworthiness? Another stream of literature adopts the human-machine trust perspective and argues that humans place more trust into computerized systems as opposed to humans [25, 26]. Researchers explain this phenomenon with the *automation bias*—humans' propensity to trust computerized decision support in order to reduce own cognitive efforts [27]. Cues of automation are used as a heuristic in decision making because humans perceive computer systems as more objective and rational if compared to other humans [18]. Experimental evidence supports the trust-inducing effect of highly computerized systems [16, 25, 28, 29].

These two opposing positions bring about an interesting research puzzle. While the human-human trust perspective suggests that anthropomorphic design is beneficial for agents' trustworthiness, the human-machine trust perspective suggests to minimize anthropomorphic design to make agents' more trustworthy. Some researchers have investigated the difference between human-human and human-machine trust [17, 28, 29]. However, the issue regarding the trust-inducing effect of

**Table 1** Overview of perspectives on human–agent trust

Perspective on human–agent trust	Human–human trust	Human–machine trust
Description	The same psychological constructs and mechanisms can be applied to explain interpersonal and human–agent trust	Interpersonal trust conceptualizations provide only limited explanation. Distinct psychological constructs and mechanisms need to be considered to explain human–agent trust
Theoretical foundation	Computers are social actors: media-equation hypothesis	Automation bias: authority hypothesis
Main references	Nass et al. [10]; Reeves and Nass [11]; Bickmore and Cassell [13]; Cassell and Bickmore [12]	Dijkstras et al. [15]; Madhavan and Wiegmann [17]; Dijkstras [16]; Mosier and Skitka [18] Muir [14]
Anthropomorphism and initial trust	Anthropomorphized agents are related to higher initial trust	Anthropomorphized agents are related to lower initial trust
Main explanation	Anthropomorphism makes novel systems more familiar and controllable	Computer systems are believed to be more capable, rational and objective than humans

anthropomorphism remains unresolved. Understanding what situational factors influence the validity of the two positions is important to increase our conceptual understanding of human–agent trust and to inform designers of conversational technologies. We address this gap with the following research question:

What factors determine whether anthropomorphic design increases users’ initial trust into a conversational agent?

As conversational agents use human language to interact with users, anthropomorphism appears to be a natural characteristic of this technology. Nevertheless, we posit that also in the light of software agents that are able to use human language, it is not unconditionally beneficial to assign them human characteristics and behavior. Thereby, this research enhances the literature on trust into technology by considering the distinct nature of conversational technologies. In order to address the formulated research question, we are investigating the psychological mechanisms that relate anthropomorphic design to users’ trusting behavior towards a conversational agent. By doing so, we seek to identify the situational factors that explain whether or not anthropomorphic design has a positive effect on agents’ trustworthiness. We are examining this relationship by considering extant research in the context of trust into technology [7, 30, 31], the CASA paradigm [10] and the automation bias literature [25]. To test our proposed research model, we plan to conduct a NeuroIS experiment that allows us to broaden our understanding of the cognitive processes related to the evaluation of anthropomorphic design and trustworthiness of conversational agents.

## 2 Theoretical Development: Trust and Anthropomorphism

*Trust* on an individual level is defined as “a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another” [32]. The initial evaluation of characteristics of another actor determines the perceived trustworthiness and, ultimately, influences the decision to trust [33]. Three conceptually distinct trust dimensions have been acknowledged by extant research on technology use [30, 34]. First, competence reflects perceptions of a trustee’s ability to produce the desired outcome. Second, benevolence reflects the extent to which the trustee is perceived to be motivated to put the interest of the trustor first. Integrity, finally, reflects the extent to which the trustee is perceived to adhere to generally accepted principles and to be honest. Keeping the objective of the present research in mind, we posit to further distinguish between qualification- and goodwill-based trustworthiness in order to account for potential variations related to anthropomorphic design. The dimensions of benevolence and integrity are classified as goodwill-based trustworthiness as these consider a trustee’s intentions and motives to fulfill the raised expectations. Intentions and thoughts are a central differentiating aspect between humans and computers. Therefore, we believe it is important to contrast these from the competence dimension which is purely qualification-based. The distinction between the volitional and non-volitional dimensions of initial trust is in accordance with the reconceptualization of trust proposed by Barki et al. [35].

*Anthropomorphism* refers to the human tendency to attribute humanlike characteristics such as intentions, emotions or motivations to non-human agents [36]. According to psychological theory, the tendency to anthropomorphize is not universal but is triggered when humans feel the urge to increase their perceived control of an otherwise unpredictable agent [37]. When the behavior of an agent is unpredictable, anthropomorphizing this agent helps to increase the perceived level of familiarity and control with regard to that agent [21]. Correspondingly, anthropomorphism is positively related to perceived predictability. Predictability is a construct closely related to trust as it reflects the extent to which one is certain about the motives and intentions of a trustee [34]. Yet, predictability in contrast to trust is a neutral construct and can have positive or negative implications. Perceived predictability can foster the positive effect of perceived trustworthiness on trusting behavior (H4, H5) [34]. The study of anthropomorphism in the context of human-computer interaction is closely related to the CASA paradigm [10]. Studies adopting the CASA perspective find evidence that anthropomorphic design, for example via the use of social cues (names, appearance) or human behavior (politeness, gestures), increases perceptions of computer agents’ trustworthiness [21–23, 38]. IS studies interested in users’ trust towards recommendation agents [39–42] and trust in e-commerce [9, 43] confirm the positive effect of anthropomorphic design. In this context, it is important to also consider research on the role of perceived agency on social expectations and behavior (i.e. trusting behavior) in

human–computer interactions [44, 45]. Appel et al. [45], for example, found support that knowledge about the agency (human vs. computer) of the interaction partner is related to feelings of social presence (human agency: high social presence). However, they also found evidence indicating that the displayed human characteristics (i.e. anthropomorphic design) are more important for social behavior in human–computer interactions than the knowledge about the agency of the interaction partner. In examining the role of anthropomorphism on trust towards conversational agents, it is thus important to make the agency condition explicit to minimize potential confounding effects. In sum, research on agency and anthropomorphic design supports the perspective of the CASA paradigm on the positive relationship between human-likeness and trust.

While, thus, one stream of literature argues that anthropomorphism—by increasing perceptions of control and familiarity—is positively related to trustworthiness (H1, H2) and predictability (H3), automation bias literature proposes the opposite [25, 27]. In accordance with that perspective, humans tend to trust computational systems more than other humans because humans are expected to be imperfect while the opposite is true for automation [28]. Therefore, humans use cues of automation as a heuristic to assess the perceived competence of an agent [27]. As humans naturally seek to minimize cognitive effort, heuristics provide a convenient way to perform such assessments [46]. Accordingly, anthropomorphic design is negatively related to trustworthiness and predictability as cues of humanness indicate lower qualification and also cause more cognitive evaluation efforts.

We, however, propose that both perspectives are valid in the context of conversational agents and that the agent substitution type acts as a moderator on the effect of anthropomorphic design on perceived trustworthiness and perceived predictability. We propose to differentiate between the agent as human-substitute and system-substitute. The former refers to instantiations where a conversational agent is implemented in order to substitute a human expert (e.g. sales person, teacher). The latter refers to conversational agents that are implemented to provide a more user-friendly interface to computer systems (e.g. enterprise software, databases). We expect that, in accordance with the CASA paradigm, agents as human-substitutes in contrast to system-substitutes benefit from increased anthropomorphism in terms of trust. We theorize that different expectations are triggered by the substitution type that translate into cognitive processes related to assessing an agents' trustworthiness.

More precisely, we expect that due to humans' desire to decrease uncertainty anthropomorphic design will be positively related to trustworthiness and predictability for human-substitute agents through increased feelings of control and familiarity (H6a, H7a, H8a). Anthropomorphizing unknown and novel interaction partners increases the perceived level of control and similarity because humans can use existing social knowledge in assessing the non-human other [36]. Exclusively human characteristics including intention, emotion and consciousness are assigned to an anthropomorphized non-human agent [21, 22]. This is beneficial for agents of a human-substitute type because in their role they need to meet not only

qualification-related but also goodwill-related expectations. On the other hand, we expect that due to humans’ qualification-focused expectations and their desire to decrease mental effort the positive effect of anthropomorphic design will be negatively moderated by a system-substitute agent type (H6b, H7b, H8b). The rationale behind this is that the conveyance of human characteristics causes cognitive evaluation effort that does not add value to the predictability and qualification assessment of a system-substitute agent who is primarily expected to efficiently perform non-human tasks.

We further expect that the differences in cognitive processing related to the trust assessment can be revealed by the use of neurophysiological measures. Riedl et al. [47] conducted a brain imaging study and found mentalizing effort in interactions with computer agents. Moreover, they found less effort if compared to interaction with humans. Because increased mentalizing implies increased cognitive effort, in accordance with the Riedl et al. [47] study, we expect to find more cognitive effort caused by anthropomorphic design if compared to non-anthropomorphic design. The underlying rationale for this theorizing is also supported by evidence showing that whenever people perceive human attributes in other agents even in objects, they tend to activate mentalizing (e.g. [48]). Because eye movement measures allow to infer cognitive states of attention and mental processing [49], we plan to use eye-tracking to investigate the trust evaluation effort. Based on our theoretical discussion, we propose the following research model (Fig. 1).

### 3 Proposed Experimental Design

To empirically validate our research model, we plan to conduct a controlled experiment that allows us to capture behavioral, self-reported and neurophysiological measures. We propose a 2 × 2 within-subjects factorial design to examine

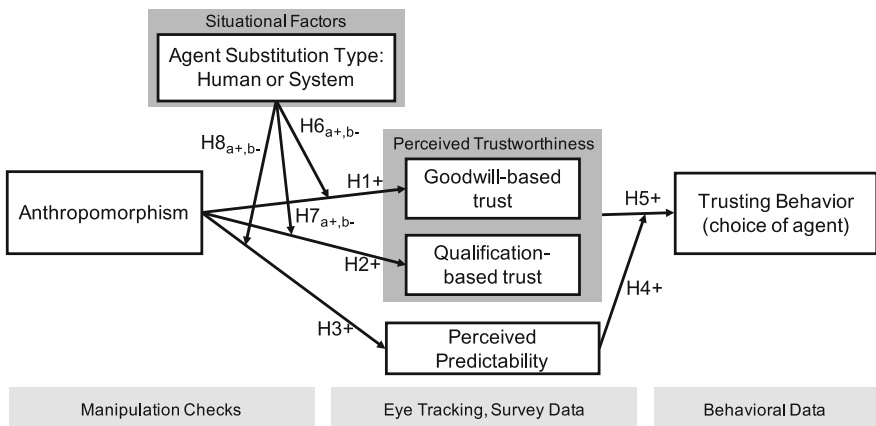


Fig. 1 Proposed research model



the effects of anthropomorphism and agent type on users' perceptions of agents' trustworthiness and trusting behavior. Two levels of anthropomorphic design (high vs. low) and two types of agents (human-substitute vs. system-substitute). To ensure that agency of the conversational agent does not confound our findings, we are informing the participants that all conversational agents are representations of computer algorithms (non-human agency; see [45]). Participants will be provided with a task scenario. They are assigned a role as an associate in a marketing team of a company. They are told that their manager wants to capitalize on the latest progress in chatbot technology and asked them to evaluate and decide which enterprise chatbot should be implemented to efficiently perform transactions in the CRM system (system-substitute) and which customer service chatbot should be implemented as a first touching point for customers (human-substitute). In order to make the manipulation of the agent type more explicit, participants will initially be informed how the respective task is currently performed. For each chatbot type the participants are provided with a highly and a slightly anthropomorphized agent (high vs. low). To provide a cover story the participants will be asked to evaluate and decide on the implementation of two other tools for the team. The order of the decision tasks will be randomized.

**Measurements.** Established self-rating scales for trustworthiness and predictability will be adapted from prior literature [9, 30]. During the study, we will use eye-tracking to capture participants' eye movement, fixation and pupil dilation. According to the eye-mind hypothesis [50], eye movement data is closely related to cognitive processing of cues in view of a person. The use of eye-tracking to understand cognitive processes in human-computer interactions is in accordance with prior IS studies (e.g. [43, 51]). In relation to our hypothesis that anthropomorphism results in more cognitive effort required to evaluate agents' trustworthiness, we expect this to translate into longer fixations. Combining this with self-reports on perceived trustworthiness, we expect that more intensive processing (fixation data) aligns with higher perceived trustworthiness in the human-substitute condition and with lower perceived trustworthiness in the system-substitute condition. In addition, we attempt to include data on pupil dilation to measure the uncertainty related to the assessment of anthropomorphic design in the two agent type conditions. According to research in neuroscience, fluctuations in pupil diameter are triggered by states of arousal in cognitive demanding situations such as decision-making under uncertainty [52, 53]. A series of experiments has successfully related perceptions of uncertainty and unexpected outcomes to increased pupil diameter [54–56]. Based on these findings we are confident that the diagnosticity—the precision of a physiological measure to capture the target construct [57]—of pupil dilation as a measure for uncertainty is established. In line with this body of research Xu and Riedl (2011), for example, propose to include pupil dilation to measure perceptions of uncertainty in e-commerce decision-making tasks [58]. Similarly, we expect that the uncertainty triggered by the inadequate use of anthropomorphism (human- vs. system-substitute type) is reflected in fluctuations of pupil diameter.

In addition to the self-rating scales and neurophysiological measures, the choice decision made between the offered chatbots represents the behavioral trust measure. Finally, the following control variables will be included due to their established importance in the context of trust and anthropomorphism in human–computer interactions: gender [59], trust propensity [30], computer self-efficacy [60], dispositional anthropomorphism [61] and need for cognition [15].

## 4 Discussion and Expected Contributions

We identified an existing contradiction regarding the use of anthropomorphic design to stimulate users' trust into computer agents. Because conversational agents are characterized by their ability to interact in human language, it appears intuitive to conclude that such systems benefit from anthropomorphic design. By building upon theoretical knowledge on anthropomorphism, cognitive heuristics and trust we challenge this intuition. More precisely, we are proposing that the agent substitution type changes user expectations and perceptions regarding anthropomorphism. Our experimental approach seeks to assess these differences through a combination of self-rating, eye-tracking and behavioral data. By adopting a NeuroIS perspective, we seek to confirm that a misuse of anthropomorphisms results in cognitive responses that damage an agents' trustworthiness. We expect this research to enhance existing understanding of cognitive processes triggered by anthropomorphic system design and their effect on trust perceptions. In this context, future research will also need to consider the role of the uncanny valley effect—the phenomenon that as non-human objects appear more human-like (anthropomorphic design) they increase perceptions of familiarity and trust until a certain threshold is reached that triggers sudden perceptions of disturbance and rejection due to the objects' non-human imperfections [62]. Finally, this project also provides new areas for IS research on user trust. Future studies can investigate how trust-violation and -repair dynamics differ between human- and system-substitution type and how this relates to cognitive and emotional responses.

## References

1. Schuetzler, R.M., Grimes, M., Giboney, J.S., Buckman, J.: Facilitating natural conversational agent interactions: lessons from a deception experiment. In: Proceedings of the 35th International Conference on Information Systems, Auckland, pp. 1–16 (2014)
2. Nunamaker Jr., J.F., Derrick, D.C., Elkins, A.C., Burgoon, J.K., Patton, M.W.: Embodied conversational agent-based Kiosk for automated interviewing. *J. Manag. Inf. Syst.* **28**, 17–48 (2011)
3. Kassner, L., Hirmer, P., Wieland, M., Steimle, F., Königsberger, J., Mitschang, B.: The social factory: connecting people, machines and data in manufacturing for context-aware exception

- escalation. In: Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS), pp. 1673–1682 (2017)
4. Gartner: (2016) <https://www.gartner.com/doc/3471559?ref=SiteSearch&sthkw=chatbot&fhl=search&srcId=1-3478922254>
  5. Forbes: (2016) <http://www.forbes.com/sites/parmyolson/2016/02/23/chat-bots-facebook-telegram-wechat/#4c95b5762633>
  6. Braun, A.: *Chatbots in der Kundenkommunikation*. Springer, Berlin (2013)
  7. Wang, W., Benbasat, I.: Attributions of trust in decision support technologies: a study of recommendation agents for e-commerce. *J. Manag. Inf. Syst.* **24**, 249–273 (2008)
  8. Pavlou, P.A., Gefen, D.: Building effective online marketplaces with institution-based trust. *Inf. Syst. Res.* **15**, 37–59 (2004)
  9. Gefen, D., Straub, D.W.: Consumer trust in B2C e-commerce and the importance of social presence: experiments in e-products and e-services. *Omega* **32**, 407–424 (2004)
  10. Nass, C., Steuer, J., Tauber, E.R.: Computers are social actors. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp. 72–78, New York (1994)
  11. Reeves, B., Nass, C.: The media equation: how people treat computers, television? New media like real people? *Comput. Math Appl.* **5**, 33 (1997)
  12. Cassell, J., Bickmore, T.: External manifestations of trustworthiness in the interface. *Commun. ACM* **43**, 50–56 (2000)
  13. Bickmore, T., Cassell, J.: Relational agents. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems—CHI '01, New York (2001)
  14. Muir, B.M.: Trust between humans and machines, and the design of decision aids. *Int. J. Man Mach. Stud.* **27**, 527–539 (1987)
  15. Dijkstra, J.J., Liebrand, W.B.G., Timminga, E.: Persuasiveness of expert systems. *Behav. Inf. Technol.* **17**, 155–163 (1998)
  16. Dijkstra, J.J.: User agreement with incorrect expert system advice. *Behav. Inf. Technol.* **18**, 399–411 (1999)
  17. Madhavan, P., Wiegmann, D.A.: Similarities and differences between human–human and human–automation trust: an integrative review. *Theor Issues Ergon. Sci.* **8**, 277–301 (2007)
  18. Mosier, K.L., Skitka, L.J.: Human decision makers and automated decision aids: made for each other. *Autom. Hum. Perform. Theory Appl.* pp. 201–220 (1996)
  19. Nass, C., Moon, Y.: Machines and mindlessness: social responses to computers. *J. Soc. Issues* **56**, 81–103 (2000)
  20. Riedl, R., Mohr, P., Kenning, P., Davis, F., Heekeren, H.: Trusting humans and avatars: behavioral and neural evidence. In: Proceedings of the Thirty Second International Conference on Information Systems, Shanghai (2011)
  21. Epley, N., Waytz, A., Akalis, S., Cacioppo, J.T.: When we need a human: motivational determinants of anthropomorphism. *Soc. Cogn.* **26**, 143–155 (2008)
  22. Waytz, A., Heafner, J., Epley, N.: The mind in the machine: anthropomorphism increases trust in an autonomous vehicle. *J. Exp. Soc. Psychol.* **52**, 113–117 (2014)
  23. Gong, L.: How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Comput. Hum. Behav.* **24**, 1494–1509 (2008)
  24. de Visser, E.J., Monfort, S.S., McKendrick, R., Smith, M.A.B., McKnight, P.E., Krueger, F., Parasuraman, R.: Almost human: anthropomorphism increases trust resilience in cognitive agents. *J. Exp. Psychol. Appl.* **22**, 331–349 (2016)
  25. Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P.: The role of trust in automation reliance. *Int. J. Hum. Comput. Stud.* **58**, 697–718 (2003)
  26. Skitka, L.J., Mosier, K.L., Burdick, M.: Does automation bias decision-making? *Int. J. Hum. Comput. Stud.* **51**, 991–1006 (1999)
  27. Mosier, K.L., Skitka, L.J., Heers, S., Burdick, M.: Automation bias: decision making and performance in high-tech Cockpits. *Int. J. Aviat. Psychol.* **8**, 47–63 (1998)
  28. de Visser, E.J., Krueger, F., McKnight, P.: The world is not enough: trust in cognitive agents. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting (2012)

29. Merritt, S.M., Heimbaugh, H., LaChapell, J., Lee, D.: I trust it, but I don't know why. *Hum. Factors J. Hum. Factors Ergon. Soc.* **55**, 520–534 (2013)
30. McKnight, D.H., Choudhury, V., Kacmar, C.: Developing and validating trust measures for e-commerce: an integrative typology. *Inf. Syst. Res.* **13**, 334–359 (2002)
31. Komiak, S.Y.X., Wang, W., Benbasat, I.: Trust building in virtual salespersons versus in human salespersons: similarities and differences. *e-Service J.* **3**, 49–64 (2004)
32. Rousseau, D.M., Sitkin, S.B., Burt, R.S.: Not so different after all: a cross-discipline view of trust. *Acad. Manag. Rev.* **23**, 393–404 (1998)
33. Tomlinson, E.C., Mayer, R.C.: The role of causal attribution dimensions in trust repair. *Acad. Manag. Rev.* **34**, 85–104 (2009)
34. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. *Acad. Manag. Rev.* **20**, 709–734 (1995)
35. Barki, H., Robert, J., Dulipovici, A.: Reconceptualizing trust: a non-linear Boolean model. *Inf. Manag.* **52**, 483–495 (2015)
36. Epley, N., Waytz, A., Cacioppo, J.T.: On seeing human: a three-factor theory of anthropomorphism. *Psychol. Rev.* **114**, 864–886 (2007)
37. Waytz, A., Morewedge, C.K., Epley, N., Monteleone, G., Gao, J.-H., Cacioppo, J.T.: Making sense by making sentient: effectance motivation increases anthropomorphism. *J. Pers. Soc. Psychol.* **99**, 410–435 (2010)
38. Riedl, R., Javor, A., Gefen, D., Felten, A., Reuter, M.: Oxytocin, trust, and trustworthiness: the moderating role of music (2017)
39. Qiu, L., Benbasat, I.: Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. *J. Manag. Inf. Syst.* **25**, 145–182 (2009)
40. Al-Natour, S., Benbasat, I., Cenfetelli, R.T.: Creating rapport and intimate interactions with online virtual advisors. *SIGHCI 2007 Proceedings* (2007)
41. Al-Natour, S., Benbasat, I., Centefelli, R.: Trustworthy virtual advisors and enjoyable interactions: designing for expressiveness and transparency. In: *Proceedings of the European Conference on Information Systems (ECIS)*. Regensburg, Germany (2010)
42. Benbasat, I., Dimoka, A., Pavlou, P.A., Qiu, L.: Incorporating social presence in the design of the anthropomorphic interface of recommendation agents: insights from an fMRI study. In: *Proceedings of the International Conference on Information Systems (ICIS)* (2010)
43. Cyr, D., Head, M., Larios, H., Pan, B.: Exploring human images in website design: a multi-method approach. *MIS Q.* **33**, 539–566 (2009)
44. Pütten, von der, A.M., Krämer, N.C., Gratch, J., Kang, S.-H.: “It doesn't matter what you are!” Explaining social effects of agents and avatars. *Comput. Hum. Behav.* **26**, 1641–1650 (2010)
45. Appel, J., Pütten, von der, A., Krämer, N.C., Gratch, J.: Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human–computer interaction. *Adv. Hum. Comput. Interact.* pp. 1–10 (2012)
46. Fiske, S.T., Taylor, S.E.: *Social Cognition: From Brains to Culture*. Sage, London (2013)
47. Riedl, R., Mohr, P.N.C., Kenning, P.H., Davis, F.D., Heekeren, H.R.: Trusting humans and avatars: a brain imaging study based on evolution theory. *J. Manag. Inf. Syst.* **30**, 83–114 (2014)
48. Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., Kircher, T.: Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE* **3**, 1–11 (2008)
49. Poole, A., Ball, L.J.: Eye tracking in HCI and usability research. *Encycl. Hum. Comput. Interact.* **1**, 211–219 (2006)
50. Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. *Cogn. Psychol.* **8**, 441–480 (1976)

51. Léger, P.M., Sénécal, S., Courtemanche, F.: Precision is in the eye of the beholder: application of eye fixation-related potentials to information systems research. *J. Assoc. Inf. Syst.* **15**, 651–678 (2014)
52. Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M., Cohen, J.D.: Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cogn. Affect. Behav. Neurosci.* **10**, 252–269 (2010)
53. Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D.: Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron* **84**, 355–362 (2014)
54. Satterthwaite, T.D., Green, L., Myerson, J., Parker, J.: Dissociable but inter-related systems of cognitive control and reward during decision making: evidence from pupillometry and event-related fMRI. *Neuroimage*. pp. 1017–1031 (2007)
55. Preuschoff, K., Hart, B.T.: Pupil dilation signals surprise: evidence for noradrenaline’s role in decision making. *Front. Neurosci.* **5**, article 115 (2011)
56. Lavín, C., Martín, R.S., Jubai, E.R.: Pupil dilation signals uncertainty and surprise in a learning gambling task. *Front. Behav. Neurosci.* **7**, article 218 (2014)
57. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**, i–xxxv (2014)
58. Xu, Q., Riedl, R.: Understanding online payment method choice: an eye-tracking study. In: *Proceedings of the 32th International Conference on Information Systems*, pp. 1–12, Shanghai (2011)
59. Riedl, R., Hubert, M., Kenning, P.: Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *MIS Q.* **34**, 397–428 (2010)
60. Compeau, D.R., Higgins, C.A.: Computer self-efficacy: development of a measure and initial test. *MIS Q.* pp. 189–211 (1995)
61. Waytz, A., Cacioppo, J., Epley, N.: Who sees human? *Perspect. Psychol. Sci.* **5**, 219–232 (2010)
62. MacDorman, K.F., Green, R.D., Ho, C.C., Koch, C.T.: Too real for comfort? Uncanny responses to computer generated faces. *Comput. Hum. Behav.* **25**, 695–710 (2009)

# Affective Processing Guides Behavior and Emotions Communicate Feelings: Towards a Guideline for the NeuroIS Community

Peter Walla

**Abstract** Like most researchers from other disciplines the NeuroIS community too faces the problem of interchangeable terminology regarding emotion-related aspects of their work. This article aims at solving this issue by clearly distinguishing between *emotion*, *feeling* and *affective* processing and by offering clear definitions. Numerous prior attempts to agree on only an emotion definition alone have failed, even in the emotion research community itself. A further still widely neglected problem is that language as a cognitive cortical function has no access to subcortical affective processing, which forms the basis for both feelings and emotions. Thus, any survey question about anything emotional cannot be answered properly. This is why it is particularly important to complement self-report data with objective measures whenever emotion-related processes are of interest. While highlighting that cognitive processing (e.g. language) is separate from affective processing, the present paper proposes a brain function model as a basis to understand that subcortical affective processing (i.e. neural activity) guides human behavior, while feelings are consciously felt bodily responses that can arise from suprathreshold affective processing and that are communicated to others via emotions (behavioral output). To provide an exemplary consequence, according to this model fear is not an emotion, but a feeling. The respective emotion is a scared face plus other behavioral responses that show an observer that one feels fear as a result of affective processing. A growing body of literature within and outside the NeuroIS community began to reveal that cognitive, explicit responses (self-report) to emotion stimuli often deviate from implicit affective neural activity that can only be accessed via objective technology. This paper has the potential to facilitate future NeuroIS

---

P. Walla (✉)

CanBeLab, Department of Psychology, Webster Vienna Private University, Praterstrasse 23,  
1020 Vienna, Austria

e-mail: peter.walla@webster.ac.at

P. Walla

School of Psychology, Newcastle University, Newcastle, Callaghan, Australia

P. Walla

Faculty of Psychology, University of Vienna, Vienna, Austria

© Springer International Publishing AG 2018

F.D. Davis et al. (eds.), *Information Systems and Neuroscience*,

Lecture Notes in Information Systems and Organisation 25,

[https://doi.org/10.1007/978-3-319-67431-5\\_16](https://doi.org/10.1007/978-3-319-67431-5_16)

research as well as to provide an innovative understanding of emotion for the entire science community.

**Keywords** Emotion · Feeling · Affective processing · Conscious  
Non-conscious · Behavior · Emotion model · Subjective · Objective  
Implicit versus explicit

## 1 The Problem

### 1.1 Introduction

If you ever felt angry about a person you deeply love you know what love/hate is. How can one have two emotions at the same time? A quick answer is that love and hate are no emotions, they are feelings. A more elaborate answer is that given the current confusion in emotion research it is difficult to find a clear answer and only the use of a more sophisticated and accurate vocabulary and a clear understanding of human brain function can help.

Driven by the problematic and interchangeable use of the terms “affective”, “emotion” and “feeling” this article makes an effort to suggest a very concrete understanding of those words’ meanings with the purpose of proposing a distinct emotion model or better a brain function model including affective processing that is the basis for emotions. It is true, but unacceptable that most scholarly work in the field of emotion research mentions the problem of missing proper definitions without offering a solution. Within Information Systems (IS) the NeuroIS community, established since 2007, is strongly focusing on emotional aspects related to information technology (IT) and IS (e.g. [1–6]) and suffers from an absent agreement on how to define emotion. Most often feeling and emotion are used interchangeably. The herewith proposed model and terminology is meant to help the NeuroIS community to more efficiently disseminate its research outcome and to better communicate their results at conferences. Ideally, it leads to a consistent view and use of those terms. In the best case this effort also leads to a novel understanding of everything around emotion in principle. The title “affective processing guides behavior and emotions communicate feelings” already brings it to the point, but to fully understand this short and sharp statement one must go into some further detail, which starts with a good understanding of the overall function of the entire brain in the first place. Further below, a respective concept is explained and elaborated on including its neurobiological roots. Whether or not the science community will accept this solution depends on various factors and might be a matter of time and solid evaluation, but it is definitely about time to make some progress. Continuous interchangeable use of terms describing emotion-related phenomena is hindering further developments and should thus become history.

## 1.2 *Emotion in IS and NeuroIS*

Since 2009 the link between information systems (IS) and the neurosciences (NeuroIS) is discussed in the frame of the *Gmunden Retreat* (Austria) that became a yearly event with a rapidly increasing number of participants [7].

After all, it is the brain that produces behavior, perceives and appreciates design, accepts or rejects technology, thinks, makes decisions and communicates and it consists of neurons. Obviously, it makes sense to take neurosciences into account. The NeuroIS community acknowledges that and thus forms an important and very promising group of scholars as part of the large IS community.

During the early days of NeuroIS Dimoka et al. [8] concluded that there is great potential for drawing upon cognitive neuroscience theories and using brain imaging tools in IS research to enhance its own theories. In their highly valuable commentary [8] they were asking “how can the cognitive neuroscience literature inform IS research?” and “how can IS researchers use brain imaging tools to complement their existing sources of data?” While the answers to those questions will indeed provide useful further insight into IT- and IS-related research it would also be beneficial to invite neuroscientists more frequently to co-author respective written output in order to help implementing neuroscience theories and interpreting collected data. The NeuroIS community already made enormous progress within IS by recognizing that objective and unbiased measures of cognitive and affective processes are important to complement traditional data sources and by actually recording and analyzing those objective measures for their research [9].

Importantly, it has been emphasized that pure behavior research is potentially biased due to its reliance on self-report (i.e. explicit responses) [10]. Within the context of technostress the relationship between physiological (objective) and self-reported (subjective) data was investigated [11] and the authors argue that both kinds of data tap into different aspects of technostress and that only the combination of both can provide the most complete understanding of technostress impact. This means an enormous step forward. Respective empirical results show that a physiological measure explains performance on the computer-based task over and above explicit responses, which certainly reflects that the brain knows more than it admits to consciousness.

Neurophysiological methods such as electroencephalography (EEG) [12, 13], skin conductance (SC) and facial electromyography (fEMG) were used in the frame of NeuroIS investigations [14] and only recently also startle reflex modulation (SRM) has been introduced (see below).

The use of objective technologies is an important first step; however, one also wants to understand why objective measures are often better than subjective measures and the answer to that question is, because affective processing content is not directly accessible to language. Some scholars already noticed a limited self-monitoring capacity particularly related to emotionally driven decisions [15]. From a neurobiological perspective it is clear that self-report cannot properly reflect raw affective responses due to language being a cortical function, while affective



processing happens deeply subcortical. In contrast, self-report can of course easily reflect cognitive responses, which are mainly cortical themselves. The fact that words cannot easily reflect what's going on deep inside the brain has been shown in several studies about discrepancies occurring between explicit responses to affective stimulation and objective measures [16–36].

Besides those discrepancies also the way Gregor et al. [37] wrote about emotion in IS research underlines the problem of respective interchangeable terminology use. In their work, the authors speak of three interacting emotion systems: language, physiology and behavior. Remarkably, they used a multiple measurement approach (i.e. paper-based self-report measures, qualitative comments as well as EEG measures) and highlighted the multiple aspect nature of emotion. This paper already points in a very innovative direction regarding the understanding of emotion, but instead of using distinct vocabulary the authors labeled all three emotion systems by borrowing names related to other functions (language, physiology and behavior). Below, the solution begins with first explaining the brain's function from a neurobiological perspective and then by defining affective processing, feeling and emotion.

## 2 The Solution

### 2.1 *The Brain's Function*

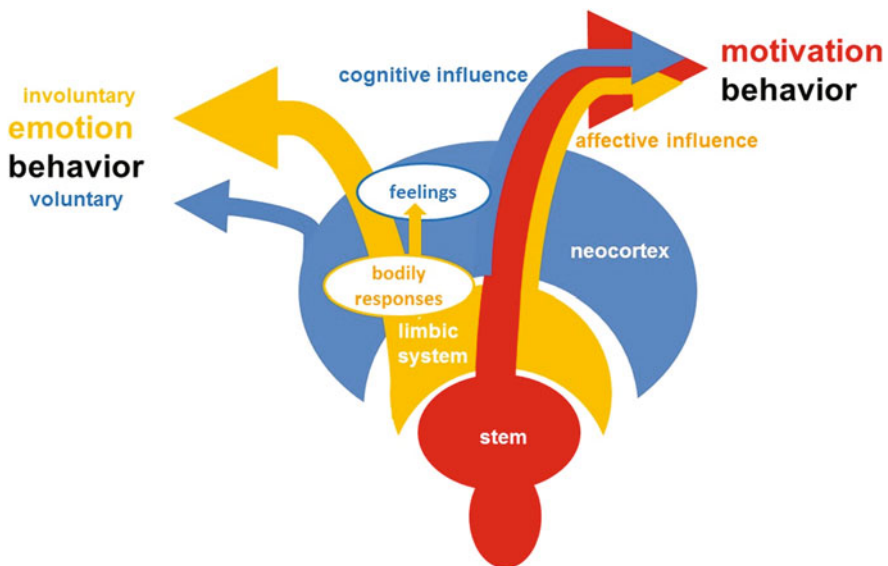
The heart pumps blood to deliver chemical substances as well as entire cells to distinct body parts. The lung extracts oxygen from the air to provide energy for the whole organism and the brain processes information to produce adapted behavior (besides maintaining homeostasis). Every organ has its function that it operates via a specific mechanism. The information the brain processes is the result of sensory input. There are no sounds, pictures, odors nor any other actual physical or chemical environmental stimuli in our brains, there are only neural signals triggered by sensory neurons and sent toward the central nervous system, which consists of the brain and the spinal cord. Seeing, hearing, smelling, tasting and touching as well as balance, body acceleration, temperature, pain and general body awareness signals inform the brain continuously about ongoing changes in the external and internal world [38].

After the translation of external and internal stimuli into the brain's language (i.e. graded potentials and action potentials) actual information processing begins. Importantly, two different information aspects are central, one cognitive and the other affective. Cognitive information focuses on semantic features that lead to an understanding of *what* something is, while affective information is evaluative leading to a decision on *how* something is [29].

It makes sense to believe that affective processing evolved before cognitive processing as a first mechanism to adapt behavior on the basis of evaluative

decisions rather than semantic rationale understanding. This idea is supported by the fact that affective information is processed by older brain structures whereas cognitive information is processed by much younger cortical networks. Primitive non-human animals still make their decisions solely based on affective processing and so is our own early childhood behavior primarily guided by affective processing. However, the brain evolved over time and at some point cognitive processing established as a consequence of natural selection [39]. But crucially, one must understand that even in us human animals any behavior is initially triggered deep inside the brain by old structures that can still be found in primitive vertebrates such as reptiles and that both affective and cognitive information processing adjusts it on the way to its execution, again with affective processing being the basis. The right part of Fig. 1 reflects such motivated behavior.

To this point, you may have noticed that the term “emotion” has not been mentioned yet even though the function of the brain has been fully explained. This is so, because emotion is here understood as behavior and not information processing. “Emotion” (from Latin “emovere” = to carry away, to remove or in other words to move out or express) is here understood as behavioral output of affective processing and because it is not processing itself, it does not directly contribute to behavior adaptation. See further details in the next paragraph and the left side of Fig. 1 that shows paths related to emotional behavior.



**Fig. 1** Schematic model demonstrating that any behavior is initially triggered deep inside the brain by old neural structures belonging to the brain stem. Crucially, before actual execution it is adapted through affective and cognitive information processing that takes influence and thus modify the way we behave (right side: motivation behavior). On the left, note the distinction between involuntary and voluntary emotion behavior (emotions are behavioral output!)

Affective processing can happen without even generating an emotion, which has serious implications, because not all disordered affective processing might show up as observable or measurable emotions (i.e. behavioral patterns). It is also of great interest to the industry and of course the IS and NeuroIS community, because the emotions a marketing expert plans to elicit or an IS scholar tries to measure might not always match up with underlying affective processing, which might negatively influence the interpretation and discussion of scientific results or lead to flop products that nobody wants. These are just some of many more examples that highlight this undoubtedly radical, but enormously helpful approach to “emotion”.

## ***2.2 The Proposed Emotion Model***

As a matter of fact, some scholars understand emotions as neural activities, others see them as felt affective phenomena and yet others as facial expressions. Independent from an exact definition of emotion it is problematic to assume that observing someone’s facial expression elicits respective responses in the observers’ brain. Given that in most cases the facial expressions were fake and that faces are often not necessarily reflective of deep inner affective states this becomes even more problematic.

The herewith proposed model states that affective processing (i.e. neural activity) represents actual information processing, while an emotion is not at all information processing, it is produced behavior as the word “emotion” in emotion suggests. Critically, emotions are not necessarily directly reflective of affective processing, which means that one should be more interested in affective processing and not emotions. Affective processing guides our behavior, not emotions!

In the mammalian brain, any motivation-based behavior (i.e. muscle contraction causing movement and locomotion) is triggered in the brain stem and on the way to its motoric execution it is first affective information processing (i.e. affective decision making) that can adapt it to environmental changes by evaluating stimuli (approach or withdraw). This basic processing stage equals a judgment of external environmental as well as internal own body stimuli regarding their pleasant/unpleasant aspects. It is automatic and independent from cognitive processes. It evolved long before cognition and consciousness and thus also before language came into existence.

Since humans are mammals too it must be accepted that this automatic evaluative process also forms the basis for any human behavior. In humans though, cognitive processes can influence and overrule affective decision making, which is usually referred to as emotion control, but should from now on be understood as affection control. Nevertheless, getting back to affective processing one can say that if it crosses certain thresholds (supra-threshold neural activity) it leads to bodily responses (hypothalamus, visceral, etc.) that can be perceived and thus lead to feelings. All organisms capable of consciousness, which is a prerequisite of perception, can have feelings (basically all mammals). So, feelings are conscious phenomena, but they are not cognitive, they are perceived bodily responses. To

consciously experience (to feel) a bodily response is like to consciously experience (to see) visual information.

Like in the previous paragraph, we haven't heard anything yet about THE most often used term, *emotion*. Again, what is an emotion, the probably most inflationary word that is in everybody's mouth? According to the current model, emotions are possible behavioral results of affective processing, they are separate to feelings. However, there is a possible link between emotions and feelings in that emotions as behavioral responses can communicate feelings. As stated in the abstract the feeling of fear can be communicated by a respective facial expression. Perhaps, emotions evolved to let conspecifics know how one feels. It is here suggested to call those emotions that are genuine results of affective processing *involuntary* emotions (maybe they resemble so called micro-expressions). As mentioned above, cognition can interfere and organisms that are capable of cognitive information processing can intentionally modify emotions (e.g. facial expressions) and use them for strategic nonverbal communication purposes. This is an evolutionary advantage and such emotions are here referred to as *voluntary* emotions (e.g. fake facial expressions, exaggerated expressions, etc.). Table 1 provides an overview of all important definitions related to the newly proposed emotion model.

To link this model to existing emotion theories it can be said that it contains aspects of the well-known James-Lange-emotion theory, which also links bodily responses to feelings. In his 1884 paper [40], "What is an emotion?" James wrote that "the emotional brain-processes not only resemble the ordinary sensorial brain-processes, but in very truth are nothing but such processes variously combined". The crucial change in terminology now is that those brain processes are here called "affective processing", while emotions are defined as their behavioral consequences, while feelings, like James suggested, are felt bodily responses.

Darwin [39], when writing about affections, mentioned changes in the functioning of glands and muscles, which basically are the only effectors that get activated as a consequence of prior information processing in the brain. The current idea to put strong emphasis on behavioral output when talking about emotion resembles that view. Fear is a feeling that arises when respective neural activity elicits respective physiological bodily responses and the scared face is the emotion.

James also says that "the immense number of parts modified in each emotion is what makes it so difficult for us to reproduce in cold blood the total and integral expression of any one of them. We may catch the trick with the voluntary muscles, but fail with the skin, glands, heart, and other viscera." In terms of the current model this means that voluntary emotions can never fully copy involuntary emotions.

**Table 1** Summary of definitions

Affective processing	Neural activity coding for valence ("how" aspects)
Feeling	Felt bodily response arising from suprathreshold affective processing
Emotion	Behavioral output of affective processing communicating feelings
Cognitive processing	Neural activity coding for semantic information ("what" aspects)

### 3 Conclusions

This article covers two major topics of interest. First, due to explicit language functions being cortical mechanisms self-report data cannot adequately reflect affective brain responses that happen deeply subcortical. This inevitably leads to misleading results whenever survey-based data alone are analyzed. Second, the interchangeable terminology related to emotion can be solved by accepting that emotions are possible behavioral responses to affective processing (e.g. facial expressions) and feelings are felt bodily responses that arise as a consequence of strong (suprathreshold) affective processing.

Affective processing equals neural activity representing the most basic decision-making quality that guides human behavior. An emotion has nothing to do with information processing, it's behavior. Indeed, it has to be emphasized that this model is quite reductionist, but short and accurate explanations are better than long, confusing and also no explanations.

### References

1. Kallinen, K., Ravaja, N.: Effects of the rate of computer-mediated speech on emotion-related subjective and physiological responses. *Behav. Inf. Technol.* **24**(5), 365–373 (2005)
2. Zorn, T.: The emotionality of information and communication technology implementation. *J. Commun. Manag.* **7**(2), 160–171 (2002)
3. Alonso-Martín, F., Malfaz, M., Sequeira, J., Gorostiza, J., Salichs, M.: A multimodal emotion detection system during human-robot interaction. *Sensors* **13**(11), 15549–15581 (2013)
4. Shaikh, M., Prendinger, H., Ishizuka, M.: Emotion Sensitive News Agent (ESNA): A system for user centric emotion sensing from the news. *Web Intell. Agent Syst.* **8**(4), 377–396 (2010)
5. Zhang, Q., Lee, M.: Emotion development system by interacting with human EEG and natural scene understanding. *Cogn. Syst. Res.* **14**(1), 37–49 (2012)
6. Callejas, Z., López-Cózar, R.: Influence of contextual information in emotion annotation for spoken dialogue systems. *Speech Commun.* **50**(5), 416–433 (2008)
7. Riedl, R., Banker, R., Benbasat, I., Davis, F., Dennis, A., Dimoka, A., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Müller-Putz, G., Pavlou, P., Straub, D., Brocke, J., Weber, B.: On the foundations of NeuroIS: reflections on the Gmunden retreat 2009. *Commun. Assoc. Inf. Syst.* **27**, 243–264 (2010)
8. Dimoka, A., Banker, R., Benbasat, I., Davis, F., Dennis, A., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Pavlou, P., Müller-Putz, G., Riedl, R., Brocke, J., Weber, B.: On the use of neurophysiological tools in IS research: developing a research agenda for NeuroIS. *MIS Q.* **36**(3), 679–702 (2012)
9. Riedl, R., Davis, F., Hevner, A.: Towards a NeuroIS research methodology: intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**(10), i–xxxv (2014)
10. Brocke, J., Liang, T.: Guidelines for neuroscience studies in information systems research. *J. Manag. Inf. Syst.* **30**(4), 211–234 (2014)
11. Tams, S., Hill, K., Guinea, A., Thatcher, J., Grover, V.: NeuroIS—alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *J. Assoc. Inf. Syst.* **15**(10), 723–753 (2014)
12. Léger, P.-M.: Riedl, R., and vom Brocke, J.: Emotions and ERP information sourcing: the moderating role of expertise. *Ind. Manag. Data Syst.* **114**(3), 456–471 (2014)

13. Müller-Putz, G., Riedl, R., Wriessnegger, S.: Electroencephalography (EEG) as a research tool in the information systems discipline: foundations, measurement, and applications. *Commun. Assoc. Inf. Syst.* **37**, 911–948 (2015)
14. Minas, R., Potter, R., Dennis, A., Bartelt, V., Bae, S.: Putting on the thinking cap: using NeuroIS to understand information processing biases in virtual teams. *J. Manag. Inf. Syst.* **30** (4), 49–82 (2014)
15. Astor, P., Adam, M., Jerčić, P., Schaaff, K., Weinhardt, C.: Integrating biosignals into information systems: a NeuroIS tool for improving emotion regulation. *J. Manag. Inf. Syst.* **30** (3), 247–278 (2013)
16. Mühlberger, A., Wieser, M.J., Pauli, P.: Darkness-enhanced startle responses in ecologically valid environments: a virtual tunnel driving experiment. *Biol. Psychol.* **77**(1), 47–52 (2008)
17. Walla, P., Richter, M., Färber, S., Leodolter, U., Bauer, H.: Food evoked changes in humans: startle response modulation and event-related potentials (ERPs). *J. Psychophysiol.* **24**(1), 25–32 (2010). doi:[10.1027/0269-8803/a000003](https://doi.org/10.1027/0269-8803/a000003)
18. Geiser, M., Walla, P.: Objective measures of emotion during virtual walks through urban neighbourhoods. *Appl. Sci.* **1**(1), 1–11 (2011). doi:[10.3390/app1010001](https://doi.org/10.3390/app1010001)
19. Grahl, A., Greiner, U., Walla, P.: Bottle shape elicits gender-specific emotion: a startle reflex modulation study. *Psychology* **7**, 548–554 (2012). doi:[10.4236/psych.2012.37081](https://doi.org/10.4236/psych.2012.37081)
20. Nesbitt, K., Blackmore, K., Hookham, G., Kay-Lambkin, F., Walla, P.: Using the startle eye-blink to measure affect in players. In: *Serious Games Analytics*, pp. 401–434. Springer International Publishing (2015). [http://dx.doi.org/10.1007/978-3-319-05834-4\\_18](http://dx.doi.org/10.1007/978-3-319-05834-4_18)
21. Mavratzakis, A., Molloy, E., Walla, P.: Modulation of the startle reflex during brief and sustained exposure to emotional pictures. *Psychology* **4**, 389–395 (2013). doi:[10.4236/psych.2013.44056](https://doi.org/10.4236/psych.2013.44056)
22. Walla, P., Rosser, L., Scharfenberger, J., Duregger, C., Bosshard, S.: Emotion ownership: different effects on explicit ratings and implicit responses. *Psychology* **3A**, 213–216 (2013). doi:[10.4236/psych.2013.43A032](https://doi.org/10.4236/psych.2013.43A032)
23. Walla, P., Koller, M., Brenner, G., Bosshard, S.: Evaluative conditioning of brand attitude—comparing explicit and implicit measures. In: Conference paper accepted for the 2016 European Marketing Academy Conference in Oslo (2016)
24. Koller, M., Walla, P.: Towards alternative ways to measure attitudes related to consumption: Introducing startle reflex modulation. *J. Agric. Food Ind. Org.* **13**(1), 83–88 (2015)
25. Walla, P., Koller, M., Brenner, G., Bosshard, S.: Evaluative conditioning of established brands: implicit measures reveal other effects than explicit measures. *J. Neurosci. Psychol. Econ.* (2017) <http://dx.doi.org/10.1037/npe0000067> (March 2 online first publication)
26. Walla, P., Schweiger, M.: Samsung versus Apple: smartphones and their conscious and non-conscious affective impact. In: Full Conference Paper in Information Systems and Neuroscience. Volume 16 of the series Lecture Notes in Information Systems and Organisation, pp. 73–82 (2017)
27. Kunaharan, S., Walla, P.: Clinical neuroscience—towards a better understanding of non-conscious versus conscious processes involved in impulsive aggressive behaviours and pornography viewership. *Psychology* **5**, 1963–1966 (2014). doi:[10.4236/psych.2014.518199](https://doi.org/10.4236/psych.2014.518199)
28. Koller, M., Walla, P.: Established liked versus disliked brands: brain activity, implicit associations and explicit responses. *Cogent Psychol.* **3**, 1176691 (2016)
29. Lyons, G.S., Walla, P., Arthur-Kelly, M.: Toward improved ways of knowing children with profound multiple disabilities (PMD): introducing startle reflex modulation. *Dev. Neurorehabil.* **16**(5), 340–344 (2013). doi:[10.3109/17518423.2012.737039](https://doi.org/10.3109/17518423.2012.737039)
30. Walla, P., Koller, M.: Emotion is not what you think it is: startle reflex modulation (SRM) as a measure of affective processing in NeuroIs. In: *NeuroIs Conference Proceedings*, Springer (2015) [http://dx.doi.org/10.1007/978-3-319-18702-0\\_24](http://dx.doi.org/10.1007/978-3-319-18702-0_24)
31. Walla, P., Koller, M., Meier, J.: Consumer neuroscience to inform consumers—physiological methods to identify attitude formation related to over-consumption and environmental damage. *Front. Hum. Neurosci.* 20 May 2014 (2014). <http://dx.doi.org/10.3389/fnhum.2014.00304>

32. Koller, M., Walla, P.: Measuring affective information processing in information systems and consumer research—introducing startle reflex modulation. In: ICIS Proceedings, Breakthrough Ideas, Full Paper in Conference Proceedings, Orlando 2012. <http://aisel.aisnet.org/icis2012/proceedings/BreakthroughIdeas/1/>
33. Walla, P., Brenner, G., Koller, M.: Objective measures of emotion related to brand attitude: a new way to quantify emotion-related aspects relevant to marketing. *PLoS ONE* **6**(11), e26782 (2011). doi:[10.1371/journal.pone.0026782](https://doi.org/10.1371/journal.pone.0026782)
34. Mavratzakis, A., Herbert, C., Walla, P.: Emotional facial expressions evoke faster orienting responses, but weaker emotional responses at neural and behavioural levels compared to scenes: a simultaneous EEG and facial EMG study. *Neuroimage* **124**, 931–946 (2016)
35. Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., Allan, K.: Dissociation of the neural correlates of implicit and explicit memory. *Nature* **392**(6676), 595–598 (1998). doi:[10.1038/33396](https://doi.org/10.1038/33396)
36. Walla, P., Endl, W., Lindinger, G., Deecke, L., Lang, W.: Implicit memory within a word recognition task: an event-related potential study in human subjects. *Neurosci. Lett.* **269**(3), 129–132 (1999) [http://dx.doi.org/10.1016/S0304-3940\(99\)00430-9](http://dx.doi.org/10.1016/S0304-3940(99)00430-9)
37. Gregor, S., Lin, A., Gedeon, T., Riaz, A., Zhu, D.: Neuroscience and a nomological network for the understanding and assessment of emotions in information systems research. *J. Manag. Inf. Syst.* **30**(4), 13–48 (2014)
38. Walla, P., Panksepp, J.: Neuroimaging helps to clarify brain affective processing without necessarily clarifying emotions. In: Fountas, K.N. (eds.) *Novel Frontiers of Advanced Neuroimaging*. ISBN: 978-953-51-0923-5, InTech (2013) <http://dx.doi.org/10.5772/51761>
39. Darwin, C.: *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, 1st edn. John Murray, London (1859)
40. James, W.: What is an emotion? *Mind* **9**(34), 188–205 (1884). Retrieved from <http://www.jstor.org/stable/2246769>

# Beyond Traditional Neuroimaging: Can Mobile fNIRS Add to NeuroIS?

Caspar Krampe, Nadine Gier and Peter Kenning

**Abstract** NeuroIS research has shown that the application of neuroimaging methods (e.g. fMRI) could add to our understanding of human–human and human–computer interaction. However, taking the specific constraints of some neuroimaging methods into account, there is an ongoing discussion regarding the application and implementation of existing and innovative neuroimaging methods. Against this background, this work introduces an innovative neuroimaging method, namely mobile functional Near-Infrared Spectroscopy (fNIRS) to NeuroIS. By indicating that mobile fNIRS appears to be a valid neuroimaging tool, our work aims to encourage researchers to utilise mobile fNIRS in the field of NeuroIS.

**Keywords** fNIRS · NeuroIS · Mobile neuroimaging · Decision neuroscience

## 1 Neuroimaging Tools in NeuroIS—Potentialities and Obstacles

In our digital world, the usage of electronic hardware, operating software and their potential failures keep challenging our understanding of human processing. Facing this challenge, researchers in the nascent field of NeuroIS have started to utilise neuroimaging tools to expand theoretical concepts and provide insights into neural mechanisms underlying human–computer interaction. Most of the research was done using functional magnet resonance imaging (fMRI). As a consequence, fMRI is a widely used neuroimaging method, exploring or “mapping” the user’s brain.

---

C. Krampe (✉) · N. Gier · P. Kenning

Faculty of Business Administration and Economics, Heinrich-Heine-Universität,  
Düsseldorf, Germany

e-mail: caspar.krampe@hhu.de

N. Gier

e-mail: nadine.gier@hhu.de

P. Kenning

e-mail: peter.kenning@hhu.de



For example, NeuroIS researchers investigated human neural functioning in online environments by simulating a purchase scenario within fMRI [1, 2]. Consequently, in line with recent research [3], results indicated that brain responses might provide better predictions for purchase intentions than self-reported measurements [2], signifying the added value of neuroimaging methods in NeuroIS. However, the question remains whether ecological validity is given from insights that were obtained within an artificially created and stationary experimental environment [4]. As a consequence, there might be a discrepancy between the neural insights gathered in a more or less artificial setting and the user's real-world behavior. Therefore, there is a need to develop and discuss new and complementary methods that allow NeuroIS to image brain activity in naturalistic settings.

A neurophysiological method that is capable of measuring brain activity outside a laboratory is the electroencephalography (EEG). The advantageous usability of EEG was also recognised by researchers in NeuroIS, who utilised EEG to assess brain activity during human–computer interaction [5–8]. For example, the engagement with and usability of websites as well as computer games can be quantified by the use of electrophysiological signals [5, 6]. Moreover, whilst measuring brain activity utilising EEG, the distraction of work processes and potential reasons for poor decision-making have been investigated. The obtained insights have several implications for developing ERP-systems [7], giving reason to believe that human–computer interaction can be captured by means of neurophysiological tools. However, EEG measurements have the disadvantage of being very sensitive to movement artefacts. Meaning that if participants are freely moving within naturalistic settings, the application of EEG might be questionable.

Against this background, the use of mobile functional Near-Infrared Spectroscopy (fNIRS) (<http://nirx.net/nirsport/>) might be a fruitful avenue for NeuroIS. Like stationary fNIRS, mobile fNIRS is a relatively inexpensive and comfortably applicable alternative compared to other frequently and mostly stationary used neuroimaging techniques in NeuroIS research. However, due to the fact that mobile fNIRS is still a relatively novel technique, its validity has to be proven. To address this, the following section describes the methodological background and physiological parameters of fNIRS. Additionally, recent research is presented, indicating that mobile fNIRS is capable of successfully replicating a well-known brain mechanism from the field of decision neuroscience, namely the “winner-takes-it-all” effect [9, 10], showing that mobile fNIRS might also be a valid neuroimaging method in NeuroIS.

## **2 Functional Near-Infrared Spectroscopy—Functionality and Application**

Very briefly, functional Near-Infrared Spectroscopy is a neuroimaging tool to measure neural cortical activity by utilising the light absorption characteristics of de/oxygenated haemoglobin. In order to understand the functionality of fNIRS, the following section describes its methodological background and physiological parameters.

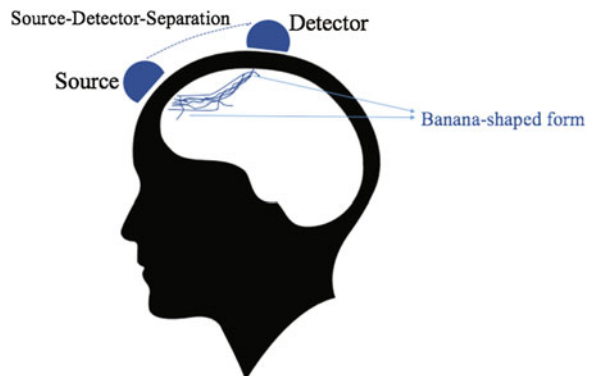
## 2.1 Methodological Background of fNIRS Measurement

Up until the present day from forty years ago, Jöbsis [11] was the first to explain how the optical properties of cerebral oxygenated and deoxygenated haemoglobin can be used to assess brain activity [12–14]. By irradiating near-infrared light into participants’ heads, scattered residuals of light can be captured, allowing the indirect quantification of neural activity to be measured. Until today the commercially available fNIRS is based on these technical and physiological principles. Using specific wavelengths of light (760 and 850 nm) that are mainly absorbed by oxygenated and deoxygenated hemoglobin, neural activity can indirectly be examined. More precisely, fNIRS projects near-infrared light through the scalp and records optical density fluctuations resulting from metabolic changes within the brain (Fig. 1). Like the BOLD-signal in fMRI, fNIRS uses cerebral blood flow as a proxy for neural activity, resulting in a high correlation between the two quantities [15, 16]. The spatial resolution and penetration depth of fNIRS is dependent on the distances between light sources and detectors, but generally fNIRS is capable of imaging depths of up to two centimeters [17]. This allows the measurement of cortical brain regions located near the scalp surface, making it particularly suitable for measuring brain regions of the prefrontal cortex, which plays a crucial role in the interpretation of information and decision-making processes [5].

In line with the “golden standard” in neuroimaging—fMRI—researchers have to follow a three-step approach to make use of fNIRS.

The first step is the acquisition of data. Participants are fitted with a cap or, by utilising mobile fNIRS, with a headband, comprising light sources and detectors that cover parts or the whole cortex. Most commonly in mobile fNIRS a headband with an 8-source/8-detector NIRS layout is used. Subsequently, researchers have to check for signal quality. Hereby, they should be aware of the fact that mobile fNIRS detectors gather optical light signals. Therefore, it is important to avoid external light interferences that could possibly distort the relevant signal. The protection against external light interferences can be ensured by using a cap covering the mobile fNIRS tool.

**Fig. 1** Schematic representation of fNIRS methodology



The second step is the data analysis. In order to perform statistical tests, raw data has to be pre-processed. In doing so, light signals are separated in order to distinguish oxygenated from deoxygenated haemoglobin. Furthermore, artefacts for example heart rate or head movements are eliminated and task-specific events are defined.

Third and finally, the task-specific events are contrasted and statistically examined. In order to identify the underlying associated brain regions, the statistically calculated values are depicted on a standard brain to visually locate the activation and interpret the results.

Taking into account that mobile fNIRS is a portable device, researchers have to be aware of specific characteristics such as the mentioned light interferences. In addition, the portable application of mobile fNIRS leads inevitably to participants' movement. In order to reduce artefacts within the data signal, expansive or galvanic movements should be avoided.

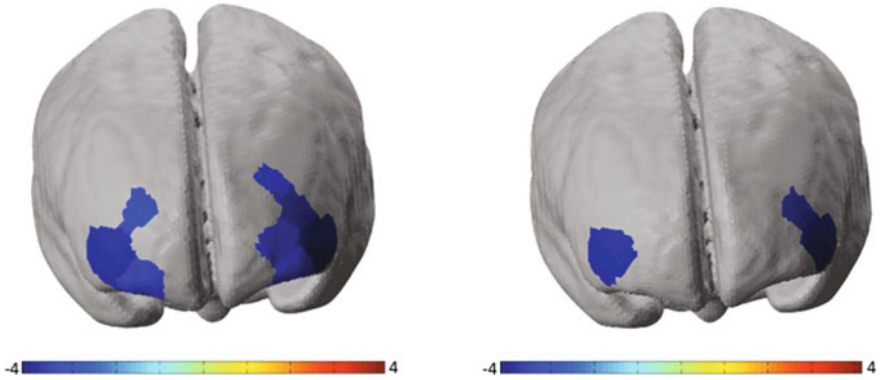
## ***2.2 Evidence for the Applicability of Mobile fNIRS Measurement***

In order to answer the question whether mobile fNIRS is indeed appropriate to assess brain activity relevant for NeuroIS, we conducted a study investigating a well-known and replicated neuroscientific effect in the field of the decision neuroscience, namely the “winner-take-it-all” effect [9, 10]. It is characterized by a decreased neural activity in the dorsolateral prefrontal cortex (dlPFC)—located laterally on the brain surface—and an increased neural activity in the ventromedial prefrontal cortex (vmPFC)—located medially within the brain—when consumers are exposed to a binary decision-making set integrating their favourite brand. With regard to this robust effect and based on the technical capabilities of mobile fNIRS, we suggest that mobile fNIRS is only able to partially detect this typical activation pattern [17]. More precisely, we hypothesised:

1. Mobile fNIRS is able to capture decreased neural activity in the dlPFC.
2. Mobile fNIRS is not capable of indicating increased neural activity of deeper-lying brain regions associated with the vmPFC.

To test these hypotheses, 23 participants were equipped with the mobile fNIRS headband. This headband consisted of an 8-source/8-detector NIRS layout, covering most of the prefrontal cortex, in particular bilateral dorsolateral prefrontal cortex, bilateral premotor cortex and bilateral orbitofrontal cortex.

Whilst measuring their neural prefrontal cortex activity, participants were asked to mentally decide between two different brands of the same product type, following the instructions used in the original study [9, 10]. In total, 100 decisions had to be taken, of which half of them integrated a predefined target brand in randomised order. Based on the participants' subjective ranking of the brands, two



**Fig. 2** Significant activation pattern contrasting the two groups TB versus non-TB during target-brand decisions at  $p < 0.01$  (left) and  $p < 0.001$  (right). Color scale specifies  $t$ -values in blue and red indicating negative and positive values respectively

groups were classified. Participants who rated the predefined target brand as their favourite brand (TB) were separated from participants who assigned another brand as their favourite (non-TB).

Following the three-step approach afore-mentioned, the raw data of each participant was truncated in order to delete negligible time intervals before and after the experimental task. Next, artefacts and irrelevant frequencies (e.g. heart rate) were removed by applying a band-pass filter. Furthermore, haemodynamic states were computed. In the last step, as in the original task, the neural activity of the two groups (TB vs. non-TB) were contrasted on target brand decision-making events.

In line with our hypotheses, a significant decreased activity of the dIPFC was identified for target brand decisions in the TB-group in comparison to the non-TB-group. As expected, no increased activity was found for brain regions associated with the vmPFC (Fig. 2).

To conclude, our results partially replicate the “winner-take-it-all” effect as suggested, indicating the validity of mobile fNIRS. Nevertheless, based on our research results it is evident that mobile fNIRS is not capable of measuring sub-jacent brain regions, such as the vmPFC, due to its technical capabilities. Therefore, NeuroIS researchers have to wisely decide a priori whether this neuroimaging method is suitable to explore their scientific entity.

### 3 Application of Mobile fNIRS to NeuroIS

Assuming mobile fNIRS as a valid proven method, it might potentially improve research in the field of NeuroIS. In order to demonstrate its application in NeuroIS research, two future research applications are described in the following section.

**Application 1: Machine Usability.** Undoubtedly, digitalisation has changed our daily life, for example at work [18]. In fact, human work life is often determined by innovative machines, that incorporate software intended to make them “smart”. Regarding productivity, users’ perceived usability and acceptance of such innovative machines is crucial. Consequently, it is essential to test the usability and practical application of innovative machines within their usual working environment (e.g. in regard to graphic-user-interfaces and human–computer interaction). However, the investigation of machine usability and human–machine interaction by means of fMRI is sometimes incompatible due to the magnetic nature of many machines. Therefore, in fMRI studies only simplified and less naturalistic versions of a scientific entity are feasible to be measured. As a consequence, relevant interaction and processing steps might be eliminated, increasing the discrepancy of the measured and actual human behaviour and its associated neural reactions. Here, mobile fNIRS has the advantage to assess relevant constructs within a naturalistic setting, including the natural usage of a machine and its associated effects evoked by the environment. Therefore, mobile fNIRS could be a promising, complementary and ecologically valid neuroimaging tool in field studies on machine usability.

**Application 2: Enterprise Resource Planning (ERP).** ERP-systems help organisations to deal with management processes that take place in modern business [19]. Generally speaking, an ERP-system is a software tool that manages all the company’s data to provide information to those who need it at the time they need it [20]. In addition to classical ERP-systems, in the future, in vivo-signals simultaneously gathered by using mobile fNIRS might indicate the cognitive load of employees and, therefore, allows an ERP-system to integrate neural data in order to optimize companies’ workflow. By assessing the load-dependent activation of the dlPFC [16]—a brain region measurable by means of mobile fNIRS—workload can be quantified for each employee individually. The ERP-system detects their cognitive load and automatically assists employees to reduce or adjust their workload. Moreover, the total amount of work (e.g. in a call center) might be distributed considering the mental capacity of each employee, enhancing the efficacy by competently shifting tasks and workload from one employee to another.

Against this background and based on these two examples, NeuroIS researchers and IS practice might consider mobile fNIRS as a novel neuroimaging tool. Compared to other frequently used neuroimaging tools in NeuroIS, mobile fNIRS provides some advantages that could encourage IS researchers to apply mobile fNIRS in future.

**Acknowledgments** The authors would like to thank the editorial board and the reviewers who provided insightful comments on earlier drafts of this chapter. Moreover, we would like to thank all of our colleagues who provided insights and expertise that greatly assisted the research work. In particular, we would like to thank Tim Eberhardt and Enrique Strelow for the data collection.

## References

1. Riedl, R., Hubert, M., Kenning, P.: Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *MIS Q.* **34**(2), 397–428 (2010)
2. Dimoka, A.: What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Q.* 373–396 (2010)
3. Venkatraman, V., et al.: Predicting advertising success beyond traditional measures: new insights from neurophysiological methods and market response modeling. *J. Mark. Res.* **52**(4), 436–452 (2015)
4. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., et al.: On the use of neurophysiological tools in information systems research: developing a research agenda for NeuroIS. *MIS Q.* **36**(3), 679–702 (2012)
5. Li, M., et al.: Enhancing user-game engagement through software gaming elements. *J. Manag. Inf. Syst.* **30**(4), 115–150 (2014)
6. Gregor, S., et al.: Neuroscience and a nomological network for the understanding and assessment of emotions in information systems research. *J. Manag. Inf. Syst.* **30**(4), 13–48 (2014)
7. Minas, R.K., et al.: Putting on the thinking cap: using NeuroIS to understand information processing biases in virtual teams. *J. Manag. Inf. Syst.* **30**(4), 49–82 (2014)
8. Léger, P.M., et al.: Precision is in the eye of the beholder: application of eye fixation-related potentials to information systems research. *J. Assoc. Inf. Syst.* **15**(10), 651 (2014)
9. Deppe, M., et al.: Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J. Neuroimaging* **15**(2), 171–182 (2005)
10. Koenigs, M., Tranel, D.: Prefrontal cortex damage abolishes brand-cued changes in cola preference. *Soc. Cogn. Affect. Neurosci.* **3**, 1 (2008)
11. Jöbsis, F.F.: Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* **198**(4323), 1264–1267 (1977)
12. Villringer, A., et al.: Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neurosci. Lett.* **154**(1), 101–104 (1993)
13. Funane, T., et al.: Quantitative evaluation of deep and shallow tissue layers' contribution to fNIRS signal using multi-distance optodes and independent component analysis. *Neuroimage* **85**, 150–165 (2014)
14. Kopton, I.M., Kenning, P.: Near-infrared spectroscopy (NIRS) as a new tool for neuroeconomic research. *Front. Hum. Neurosci.* **8**(549), 10.3389 (2014)
15. Masataka, N., Perlovsky, L., Hiraki, K.: Near-infrared spectroscopy (NIRS) in functional research of prefrontal cortex. *Front. Hum. Neurosci.* **9** (2015)
16. Fishburn, F.A., et al.: Sensitivity of fNIRS to cognitive state and load. *Front. Hum. Neurosci.* **8** (2014)
17. McCormick, P.W., et al.: Intracerebral penetration of infrared light: technical note. *J. Neurosurg.* **76**(2), 315–318 (1992)
18. CeBIT: “d!conomy—no limits”. <http://www.cebit.de/de/ausstellung/specials/topthema-dconomy/> (2017). Accessed 19 April 2017 via <http://www.cebit.de/de/ausstellung/specials/topthema-dconomy/>
19. Hartmann, J., De Angeli, A., Sutcliffe, A.: Framing the user experience: information biases on website quality judgement. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM (2008)
20. Arik Ragowsky, T.M.S.: Enterprise resource planning. *J. Manag. Inf. Syst.* **19**(1), 11–15 (2002)

# Decision Inertia and Arousal: Using NeuroIS to Analyze Bio-Physiological Correlates of Decision Inertia in a Dual-Choice Paradigm

Dominik Jung and Verena Dorner

**Abstract** Decision inertia is a cognitive process describing the reluctance to incorporate new information in choices, manifesting in the tendency to repeat previous choices regardless of the consequences. In this work, we discuss recent research in decision inertia, and show that inter-individual differences in arousal may play an important role for understanding decision inertia. We derive a NeuroIS framework for the operationalization of decision inertia, and discuss our conceptualization with a view towards a general theory of decision inertia.

**Keywords** Decision inertia · Arousal · Multiple processes · Dual-choice paradigm

## 1 Introduction

Numerous studies have established that decision-makers can show a considerable unwillingness to reach a decision, and tend to repeat previous decisions regardless of undesirable consequences. For instance, about 25% of 43 million US-Americans failed to reevaluate their medical situation and to register in the free Medicare program as it was released [1], which expanded free coverage for prescription drugs to Medicare beneficiaries. Instead they repeated their previous decision to hold their old agreements regardless of the objectively better situation. Other studies report the tendency to repeat previous decisions in strategic managerial processes [2], as well as in the context of customer journey and IS continuance models [3, 4], or the updating behavior with regard to smartphone software [5].

For a long time, this behavior has been linked to decision avoidance [6], status quo bias [7], or treated as a random process or noise [8, 9]. Contrary to this

---

D. Jung (✉) · V. Dorner  
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany  
e-mail: d.jung@kit.de

V. Dorner  
e-mail: verena.dorner@kit.de

viewpoint, a number of studies have proposed motivational factors as explanations for the occurrence of decision inertia [10, 11] but evidence is mixed. While some studies find that commitment [10], or “preference for consistency” as motivational factors are correlated with decision inertia [11], other studies reject this relation [12]. With the latest developments in psychology and NeuroIS, interest in the cognitive and bio-physiological foundations of the drivers of inertia in decision-making has grown [9, 11]. This allows to target two theoretical gaps (i) to provide a deeper insight in the phenomenon which is not understood so far, and (ii) to provide insights in the development of counter measures and design recommendations to reduce decision inertia (e.g., interactive NeuroIS components in interactive systems, which could detect if individuals are likely to act out decision inertia before they actually do it).

In a first step, this research, however, focuses on cognitive processes and disregarded the potential influence of affective processes even though evidence suggests that the latter often play an important role in decision-making (e.g. [13–15]). We suggest that decision inertia is at least partially driven by arousal, influencing individual’s tendency to rely on intuition-based decision processes [16]. For that purpose, we discuss recent research of decision inertia with respect to bio-physiological effects following the first three phases of the NeuroIS framework from [17]. Hence, we gave ourselves the following research objectives:

1. Build a theoretical research model concerning decision inertia and arousal to investigate this relationship in an experimental setting.
2. Provide an experimental framework for the operationalization and investigation of decision inertia experiments in NeuroIS. This allows to measure decision inertia in a laboratory environment, as well in the usage of interactive systems.

## 2 Decision Inertia and Arousal

Decision inertia is generally considered a cognitive process and defined in one of two ways: (i) the inability to make a decision or to change from the current position [18–20], or (ii) as a general individual tendency of choice repetition conflicting with deliberation [10, 11, 21, 22]. We will focus on studies based on the latter conceptualization, because the inability to make a decision has been linked to the tendency of choice repetition in various studies [6, 23].

Decision-making is generally considered to be a combination of multiple cognitive processes, which are in general modeled as 2-systems, or deliberative-intuitive processes [16, 24, 25]. These models combine the assumption of decision-makers maximizing their expected utility with findings that report systematic deviations (biases) from economic rationality. In this context, there is evidence that arousal is a possible driver of these two systems interacting. It has been argued that emotions and arousal explicitly manifest in the activation of the autonomic nervous system, which is responsible for emotional responding, effort



management, attention and further related functions [26]. This activation counteracts deliberative processes, which require a high amount of cognitive capacity [27].

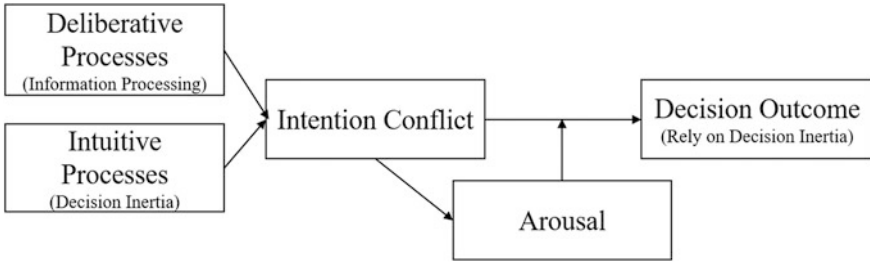
Intuitive processes, on the other hand, are fast and require substantially less cognitive capacities, combined with a low threshold for processing information [24, 28]. Findings from research on decision-making under risk suggest that high levels of arousal and strong emotions adversely influence subjective evaluations, hence decision-making [16, 29]. How the processes of deliberation and intuition interact precisely has become a subject of wide debate [30, 31].

Regarding decision inertia, we suggest that arousal could be a possible driver of intuitive processing, and reinforce the tendency to rely on decision inertia. Some evidence pointing towards this viewpoint is offered by recent studies on decision inertia in situations linked to high levels of arousal: Alison and colleagues report decision inertia occurring in decisions with equally perceived aversive outcomes especially when both decisions have life-threatening consequences [19, 20]. Decision inertia may thus be driven by the emotions and arousal linked to the negative consequences of a decision, or redundant deliberations about these negative consequences [20]. More evidence for this argumentation comes from Charness and Levin [21], who were able to reduce the effect of decision inertia in a belief-updating task by only rewarding the second decision of two subsequent decisions. This indicates that the affective response to the first decision may be a relevant driver of decision inertia.

However, hardly any research has focused on this relationship. Further evidence for these considerations comes from neurology: Yu et al. investigated the neural basis of repetition behavior in an fMRI study [32]. The decision not to switch away from the default was associated with an increased activity in the ventral striatum, which is associated with reward processing. These results suggest that decision inertia is linked to brain activities responsible for anticipating risk [32]. Hence, inter-individual differences in risk aversion, and the linked bio-physiological responses, could be a possible explanation for occurrences of decision inertia [32].

Following the definition of decision inertia, as cognitive process potentially conflicting with optimal behavior, which manifests in choice repetition, we know so far that (i) decision inertia manifests only in situations conflicting with optimal behavior [11]. Assuming that either the tendency to engage in decision inertia or engagement in decision inertia manifests in detectable bio-physiological responses [26, 32], we ought to be able to either predict or detect (ex-post) the occurrence of decision inertia at the time of a decision being made. In either case, we would expect arousal to act as a moderator in the relationship between intention building and the decision outcome, where higher levels of arousal are associated with a higher likelihood of engaging in decision inertia (see Fig. 1).

In summary, reviewing recent research indicates that decision inertia is an automatic process conflicting with optimal behavior [11], possibly driven by arousal and emotions, which may play a key role in the understanding of this phenomenon. To further investigate decision inertia, we derive a framework for the operationalization of decision inertia, which represents a generalization of the experimental procedure that underlies previous studies (e.g. [9–12, 21]).



**Fig. 1** Research model: Dual process perspective on decision inertia, based on [11, 16, 24, 25, 29]

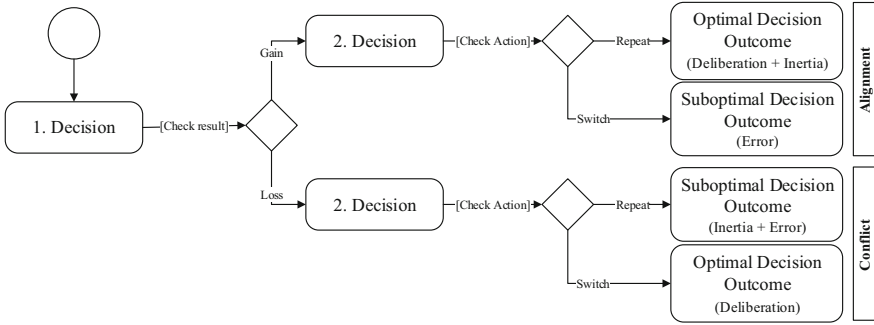
### 3 Experimental Paradigm

In general, studies on decision inertia follow a dual-choice paradigm [11]. The decision maker is confronted with two subsequent decisions, where the second decision depends in some way on the first decision. Specifically, the setup is such that the decision maker cannot possibly know the optimal decision for the first decision but, for the second decision, can calculate the best choice by taking into account the consequences of the first decision.

Hence, for the second decision, there exists a rational and a non-rational decision. Decision inertia occurs when the decision maker does not account (rationally) for the outcome of the first choice but mindlessly repeats their first choice. Pitz [10] operationalizes this dual-choice paradigm by setting the task of choosing twice from a bingo basket. Charness and Levin, and other set up an urn game where participants choose from two sets of urns [11, 21].

Other experimental designs include choosing from two directions of motions [9, 33], lottery tickets [23], or whether to repeat an unethical behavior [12]. The most popular design is the urn game, or belief-updating task, by Charness and Levin [21]. Based on the outcome of the first urn draw, the decision maker can compute the probabilities of how balls are distributed between the urns in the second draw. If they update their belief correctly, they have a higher likelihood of drawing one of the payoff-maximizing balls from the urns in the second draw. Figure 2 shows the operationalization of decision inertia in this dual-choice paradigm. In the first case, the repetition of a choice is linked to rational behavior, in the second case it is linked with decision inertia. The effect of decision inertia can be measured by comparing the error rates in the following two types of situations:

- **Situation 1 (Alignment):** The result of the first decision indicates that the optimal decision is to repeat the previous decision. Hence, decision inertia and deliberative processes are in line. Switching is an error.
- **Situation 2 (Conflict):** The result of the first decision indicates that the optimal decision is not to repeat the previous decision. Hence, if the participant does not switch, decision inertia is present.



**Fig. 2** Dual-Choice Paradigm: A framework for the operationalization of decision inertia in subsequent decision tasks derived from existing studies (e.g. [9–12, 21])

Compared to situation 1, the individual’s tendency to rely on decision inertia can be computed by:

$$\text{errorrate}_{\text{conflict}} - \text{errorrate}_{\text{alignment}} = \rho_{\text{decisioninertia}} \tag{1}$$

Turning to the role of arousal in decision inertia, we suggest integrating biosensors in the experiment due to their permitting undisturbing and objective measurements of arousal, as opposed to questionnaire-based self-reports [34]. Following our research model, we have to consider that situations where decision inertia and deliberative processes are linked to higher arousal, compared to situations where these processes are aligned. Furthermore, if we just consider the conflict-situation (situation 2), we assume that if an individual acts out decision inertia (suboptimal decision outcome) the arousal will be higher compared to the optimal behavior. In particular, we suggested that decision inertia is partly driven by arousal, influencing the intention building, and pushing the decision-outcome to rely on decision inertia. This is in line with Alós-Ferrer et al. which proposed to measure the response time as an additional indicator for the conflicting processes and decision inertia [11].

Hence, we propose to measure this relationship of decision inertia and arousal with the following bio-physiological correlates. The most popular measurements of an activation of the autonomic nervous systems are electrodermal and cardiovascular. The first one is measured by deviations of the individual’s skin conductance level or short-duration skin conductance responses [26]. The second characteristic includes measurements concerning heart rate and blood pressure. Table 1 summarizes behavioral and bio-psychological measurements for arousal and decision inertia in the proposed experimental paradigm.

**Table 1** Behavioral and bio-physiological measurements for arousal and decision inertia

Level	Description	Measure
Behavioral	Repetition of the previous decision, in contrast to optimal behavior	Decision outcome
Behavioral	Increased response time, compared to alignment of decision inertia and deliberation	Response time
Bio-psychological	Decision inertia and arousal is associated with an increased cognitive effort, resulting in increased heart rate variability	Arousal (ECG)
Bio-psychological	Additionally, electrodermal responding is quantified in differences in skin conductance level	Arousal (EDA)
Bio-psychological	Blood pressure is a cardiovascular measure of the autonomic nervous system activation. Inaction inertia could be linked to an increased activity	Arousal (BVP)

## 4 Conclusion

As we have seen, there is practical and experimental evidence that decision inertia or individual's tendency to repeat previous decisions regardless of the consequences underlies systematic processes. While recent research linked decision inertia to motivational factors with mixed findings, we argued that arousal is a relevant antecedent of decision inertial behavior. Reviewing recent decision inertia research, we derived a framework for the operationalization of decision inertia based on previous experimental task and behavioral and bio-psychological correlations. In this context we discussed the influence of arousal in decision-making, and on decision inertia in our framework. Regarding recent literature, we suggest that considering especially bio-psychological aspects of decision inertia may contribute significantly to improve our understanding of this multi-determined phenomenon. Furthermore, our framework allows to operationalize decision inertia in the lab, as well in the usage of interactive systems.

## References

1. Heiss, F., McFadden, D., Winter, J.: Who failed to enroll in Medicare Part D, and why? Early results. *Health Aff.* **25**, w344–w354 (2006)
2. Fredrickson, J.W., Iaquinto, A.L.: Inertia and creeping rationality in strategic decision processes. *Acad. Manag. J.* **32**, 516–542 (1989)
3. Limayem, M., Hirt, S.G., Cheung, C.M.K.: How habit limits the predictive power of intention: The case of information systems continuance. *MIS Q.* 705–737 (2007)
4. Polites, G.L., Karahanna, E.: Shackled to the status quo: The inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. *MIS Q.* **36**, 21–42 (2012)
5. Park, S.C.: Deciding to Update Mobile Applications: Reasons and Consequences of Inertia. *Asia Pac. J. Inf. Syst.* **26**, 268–289 (2016)

6. Anderson, C.J.: The psychology of doing nothing: forms of decision avoidance result from reason and emotion. *Psychol. Bull.* **129**, 139 (2003)
7. Samuelson, W., Zeckhauser, R.: Status quo bias in decision making. *J. Risk Uncertainty* **1**, 7–59 (1988)
8. Gold, J.I., Law, C.-T., Connolly, P., Bennur, S.: The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J. Neurophysiol.* **100**, 2653–2668 (2008)
9. Akaishi, R., Umeda, K., Nagase, A., Sakai, K.: Autonomous mechanism of internal choice estimate underlies decision inertia. *Neuron* **81**, 195–206 (2014)
10. Pitz, G.F.: An inertia effect (resistance to change) in the revision of opinion. *Can. J. Psychol. Rev. can. de psychol.* **23**, 24 (1969)
11. Alós-Ferrer, C., Hügelschäfer, S., Li, J.: Inertia and decision making. *Front. Psychol.* **7** (2016)
12. Zhang, S., Cornwell, J.F.M., Higgins, E.T.: Repeating the past: Prevention focus motivates repetition, even for unethical decisions. *Psychol. Sci.* **25**, 179–187 (2014)
13. Adam, M., Gamer, M., Krämer, J., Weinhardt, C.: Measuring emotions in electronic markets (2011)
14. Adam, M.T.P., Krämer, J., Weinhardt, C.: Excitement up! Price down! Measuring emotions in Dutch auctions. *Int. J. Electron. Commer.* **17**, 7–40 (2012)
15. Astor, P.J., Adam, M.T.P., Jähmig, C., Seifert, S.: The joy of winning and the frustration of losing: A psychophysiological analysis of emotions in first-price sealed-bid auctions. *J. Neurosci. Psychol. Econ.* **6**, 14 (2013)
16. Dhar, R., Gorlin, M.: A dual-system framework to understand preference construction processes in choice. *J. Consum. Psychol.* **23**, 528–542 (2013)
17. Vom Brocke, J., Liang, T.-P.: Guidelines for neuroscience studies in information systems research. *J. Manage. Inf. Syst.* **30**, 211–234 (2014)
18. Eyre, M., Alison, L., Crego, J., McLean, C.: Decision inertia: The impact of organisations on critical incident decision making. In: *Policing Critical Incidents: Leadership and Critical Incident Management*, pp. 201–239 (2008)
19. Alison, L., Power, N., Heuvel, C., Humann, M., Palasinski, M., Crego, J.: Decision inertia: Deciding between least worst outcomes in emergency responses to disasters. *J. Occup. Organ. Psychol.* **88**, 295–321 (2015)
20. Power, N., Alison, L.: Redundant deliberation about negative consequences: Decision inertia in emergency responders. *Psychol. Public Policy Law* (2016)
21. Charness, G., Levin, D.: When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. *Am. Econ. Rev.* **95**, 1300–1309 (2005)
22. Erev, I., Haruvy, E.: Learning and the economics of small decisions. *Handb. Exp. Econ.* **2** (2013)
23. Sautua, S.I.: Does uncertainty cause inertia in decision making? An experimental study of the role of regret aversion and indecisiveness. *J. Econ. Behav. Organ.* 1–14 (2017)
24. Strack, F., Werth, L., Deutsch, R.: Reflective and impulsive determinants of consumer behavior. *J. Consum. Psychol.* **16**, 205–216 (2006)
25. Alós-Ferrer, C., Strack, F.: From dual processes to multiple selves: Implications for economic behavior. *J. Econ. Psychol.* **41**, 1–11 (2014)
26. Mauss, I.B., Robinson, M.D.: Measures of emotion: A review. *Cogn. Emot.* **23**, 209–237 (2009)
27. Pribram, K.H., McGuinness, D.: Arousal, activation, and effort in the control of attention. *Psychol. Rev.* **82**, 116 (1975)
28. Sanbonmatsu, D.M., Kardes, F.R.: The effects of physiological arousal on information processing and persuasion. *J. Consum. Res.* **15**, 379–385 (1988)

29. Loewenstein, G., O'Donoghue, T.: Animal spirits: Affective and deliberative processes in economic behavior (2004)
30. Schwarz, N.: Emotion, cognition, and decision making. *Cogn. Emot.* **14**, 433–440 (2000)
31. Loewenstein, G., Lerner, J.S.: The role of affect in decision making. *Handb. Affect. Sci.* **619**, 3 (2003)
32. Yu, R., Mobbs, D., Seymour, B., Calder, A.J.: Insula and striatum mediate the default bias. *J. Neurosci.* **30**, 14702–14707 (2010)
33. Gold, J.I., Shadlen, M.N.: The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007)
34. Riedl, R., Davis, F.D., Hevner, A.R.: Towards a NeuroIS research methodology: Intensifying the discussion on methods, tools, and measurement. *J. Assoc. Inf. Syst.* **15**(1) (2014)

# IAT Measurement Method to Evaluate Emotional Aspects of Brand Perception— A Pilot Study

Harald Kindermann and Melanie Schreiner

**Abstract** The emotional perception of brands, explicit as well as implicit, is of interest to any brand manager. An implicit association test (IAT) could have the potential to detect unconscious attitudes and therefore evaluates intangible brand values. In a pilot study, we conducted an IAT online survey to test this implicit method to measure the emotional perception of established brand concepts. Analysis of emotional valence showed that the results compared to explicit brand evaluation with a simple question are roughly the same.

**Keywords** Implicit association test (IAT) · Measurement method  
Emotional response · Brand perception · Implicit brand attitude  
Brand management

## 1 Introduction

Emotions play an essential role in marketing and are interpreted as mental states influenced by the subjective interpretation or evaluation of a relevant effect on an individual's well-being status [1]. This kind of judgment happens consciously as well as unconsciously and automatically. In general, humans tend to avoid negative emotions and are attracted to positive emotions [1, 2]. This effect is commonly applied in marketing and used for brand building.

In a branding context, the pairing of emotional stimulus and brand evokes an unconscious learning process whereby implicit evaluations form brand attitude [3, 4]. This effect, known as evaluative conditioning, is used to actively influence emotional brand perception. It is beyond dispute that brands need to be associated with positive emotions rather than negative ones. Therefore, the emotional

---

H. Kindermann (✉) · M. Schreiner  
University of Applied Sciences Upper Austria, Digital Business, Wels, Austria  
e-mail: harald.kindermann@fh-steyr.at

M. Schreiner  
e-mail: melanie.schreiner@fh-steyr.at

perception of the brand, implicit and explicit, is of growing interest to brand management.

To capture all of these conditioning effects valid measurement methods are needed. In general, emotional responses can be divided into categories, for example awe or happiness, or dimensions such as valence and arousal [1]. For this study, we aim for a holistic measurement system to capture both category values and dimension in one go. Basically, measurement methods can be distinguished between explicit and subjective methods like self-report or implicit and objective methods like physiological methods. On the one hand explicit methods are valuable to capture current emotional status but on the other hand, they are vulnerable to group aspects like bias and individual aspects like awareness, willingness or ability to evaluate individual emotions [5, 6]. Implicit methods could balance at least some of these shortcomings mentioned. Additionally, implicit methods are able to capture implicit attitude, which plays a vital role in the purchase decision process [4, 7].

To capture the implicit evaluative conditioning effect on brands we rely on the implicit association test (IAT), which could have the potential to detect unconscious attitudes and to reveal intangible brand values [8]. When the IAT is based on association tasks of two already stored target concepts (positive or negative brand attitudes), differences in response times should occur [9]. In other words, if a brand (concept I) is strongly associated with positive emotion (concept II) the response time will be shorter than when associated with negative emotion. Using this method, the implicit emotional attitude toward a brand [8] is captured by valence. In further analyses, we will enhance the value of IAT application by evaluation of the specific response times of the different emotion categories that could lead to an emotional brand profile.

In this pilot study, we applied a broad method mix to evaluate the emotional perception of brands. First, an online survey was conducted to gather data by IAT method and self-response. Second, a laboratory experiment with a limited number of participants was conducted to capture additional physiological data. These results will be presented at the NeuroIS and will be published in an additional paper.

## **2 Method and Material**

### ***2.1 Participants***

One-hundred-sixty-five participants volunteered for the IAT online survey. Five participants were excluded due to incomplete data. The remaining data for 160 participants consisted of 107 females (66.9%) and 53 males (33.1%). The mean age was 29.33 (SD = 9.52).



## 2.2 *Stimuli*

The emotional concept to be tested included 16 emotion categories with either positive or negative valence. An online survey was conducted in the German language. Based on existing literature [10, 11] an emotion category system was developed. It included eight positive categories defined as recognition (Anerkennung), joy (Freude), happiness (Fröhlichkeit), luck (Glück), pride (Stolz), satisfaction (Zufriedenheit), well-being (Wohlbefinden), affection (Zuneigung) as well as eight negative categories defined as anger (Ärger), disgust (Ekel), frustration (Frustration), grief (Kummer), concern (Sorge), defiance (Trotz), fury (Wut), anger (Zorn). Neutral emotions were excluded. All emotion categories were presented as words during the IAT procedure.

For the brand concept, we included national and international popular brands. Each surveyed industry contained 2 brands to provide data for comparison. Commercial brands for coffee (Nespresso and Tchibo), outdoor clothing (Mammut and Jack Wolfskin) and beverages (Red Bull and Happy Day) were tested. Branding stimulus material included logos as well as images of branding campaigns and products. The brand logo was visible in all pictures.

In order to obtain results that are as comparable as possible positive stimuli were included in the procedure of the IAT, meaning that every brand concept was tested with these reference stimuli. All of the depicted images presented nature scenes, for example beaches, landscapes or forests which are associated positively [12, 13].

## 2.3 *Procedure and Design*

We conducted an IAT test similar to Greenwald et al. [9] and Lane et al. [14]. The implementation was performed with an online survey (soscisurvey with special IAT add-on). Each participant was randomly stimulated by one brand and one referenced nature picture. The add-ons software algorithm showed emotion categories randomly and selectively.

The IAT procedure consisted of practice and testing blocks (Table 1). In phase I, including blocks 1–4, the participants had to associate the nature stimulus with the positive emotion whereas the brand stimulus was associated with the negative emotion. In phase II (blocks 5–7) the association was inverted, that is brand stimulus + positive emotion. In the practice blocks, the participants trained the association whereas in the testing blocks the response time was measured and analyzed [9, 14].

After the IAT procedure, participants had to answer an online self-report questionnaire to capture possible moderator variables. Therefore, brand awareness was measured by 16 items in product categories on a bipolar scale (“cheap alternative”/“more expensive brand article”) and their product interest categorized by

industries on a three-item scale (“not interested”/“neutral”/“interested”). Additionally, brand attitude toward various brands was ascertained by a 7-point Likert scale (1 = “very unpleasant” to 7 = “very pleasant”).

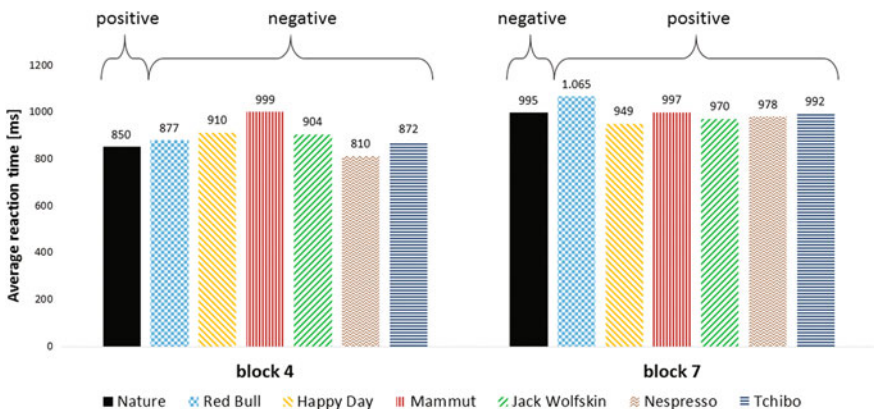
In a laboratory experiment, we applied physiological measurements to evaluate the emotional response in a more holistic way. Further, we used a remote eye-tracking system (SMI iView) to capture pupil dilation and blinks. Galvanic skin response (GSR) was ascertained by a MindMedia biofeedback system. The complete data will reveal the emotional dimensions of valence and arousal [4, 15] and will be presented at the NeuroIS in June 2017.

### 3 Results and Discussion

For IAT analysis, only the results of block 4 and block 7 were considered. Figure 1 shows the average reaction time for these two blocks, separated by brand and including the reference stimuli (nature), which was differentiated by emotional categories—positive or negative—that were assigned to the brands.

**Table 1** IAT procedure and blocks of the online survey

Block	Function	Left key assignment	Right key assignment
1	Practice	Pos. emotion categories	Neg. emotion categories
2	Practice	Nature stimulus	Brand stimulus
3	Practice	Pos. emotion + nature stimulus	Neg. emotion + brand stimulus
4	Test	Pos. emotion + nature stimulus	Neg. emotion + brand stimulus
5	Practice	Neg. emotion categories	Pos. emotion categories
6	Practice	Neg. emotion + nature stimulus	Pos. emotion + brand stimulus
7	Test	Neg. emotion + nature stimulus	Pos. emotion + brand stimulus



**Fig. 1** Average reaction time [ms]

As can be seen in Fig. 1, the reaction time in block 7 is generally higher than in block 4. This increase in reaction time could be presumably explained by the fact that the participants changed assignments. As mentioned, in block 4 the subjects learned to allocate negative emotional categories to a specific brand by pressing a key as fast as possible. In block 7 the subjects had to perform exactly the same task, but the category then allocated to the brand changed into positive. To control this confounding learning effect of block 4, it is necessary to evaluate the absolute reaction times, as depicted in Fig. 1, in comparison to the reference value alone (= nature-related stimuli: 995 ms—850 ms = 145 ms—see Table 2). With this reference value as a base, it is possible to calculate all the brand-specific difference values (e.g. for Red Bull: 188 ms—145 ms = 43 ms).

Following the cognitive fluency approach [16], it could be argued that a comparatively lower reaction time is in line with a lower cognitive load. Thus, Happy Day and Mammut, the brands with the shortest reaction times (see Table 2) could be associated best with all the positive emotional categories. On the other hand, Red Bull and Nespresso were the two brands with the worst results (see Table 2).

Even if the aim of this study is to reveal implicit coherences, we compared these IAT results with the spontaneous evaluation of these brands measured by means of a self-response 7-point Likert scale. The results are depicted in Fig. 2.

It is apparent that both the results of the implicit IAT and the explicit brand evaluation with a simple question are roughly the same. On the one hand, these insights are particularly striking because it was revealed that implicit mechanisms are in line with a self-classification. On the other hand, the question arises, as to which advantage can be derived from an implicit measure. The IAT method is distinctly more complex to apply and the preceding data preparation and analysis are time consuming.

In our opinion, the correlation between implicit and explicit measure supports the use of a questionnaire with simple scales. This aspect is important because business practice needs valid easy-to-use tools and this study supports the application of such scales. Although our results show some evidence for this implication, previous research has indicated the discrepancies between implicit and explicit measures [17–19]. Therefore, additional research is highly recommended.

**Table 2** Difference values for emotional valence

	Nature = reference	Red Bull	Happy Day	Mammut	Jack Wolfskin	Nespresso	Tchibo
Block 4 [ms]	850	877	910	999	904	810	872
Block 7 [ms]	995	1065	949	997	970	978	992
Difference values (block 7–block 4) [ms]	<b>145</b>	188	39	–2	66	167	120
<b>Difference to reference value [ms]</b>		<b>43</b>	<b>–106</b>	<b>–147</b>	<b>–79</b>	<b>22</b>	<b>–25</b>

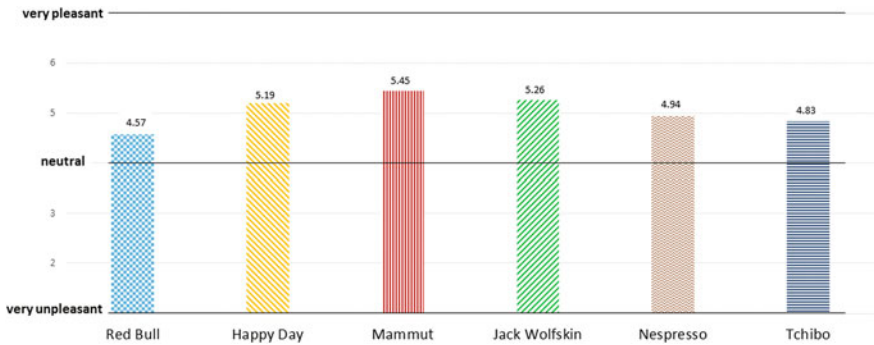


Fig. 2 Results of the spontaneous evaluation of the brands

## 4 Limitations and Future Research

One of the greatest challenges when applying the IAT measurement tool is the enormous effort of data preparation. This aspect makes such research highly time consuming, leading to low efficiency. This aspect must be considered in the light of the novelty of the results. So far, it is too early to come to a final evaluation of IAT measurement to capture emotional aspects of brands. We will do all analyses of GSR, pupil dilation and the different emotional categories next. These upcoming results may contribute to an adequate understanding of IAT measurement.

## References

1. Bagozzi, R.P., Gopinath, M., Nyer, P.U.: The role of emotions in marketing. *J. Acad. Mark. Sci.* **27**, 184–206 (1999)
2. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: Emotion, attention, and the startle reflex. *Psychol. Rev.* **97**, 377–395 (1990)
3. Gibson, B.: Can evaluative conditioning change attitudes toward mature brands? New evidence from the implicit association test. *J. Consum. Res.* **35**, 178–188 (2008)
4. Walla, P., Koller, M., Brenner, G., Bosshard, S.: Evaluative conditioning of established brands. Implicit measures reveal other effects than explicit measures. *J. Neurosci. Psychol. Econ.* (2017)
5. Gregg, A.P., Klymowsky, J.: The implicit association test in market research. Potentials and pitfalls. *Psychol. Mark.* **30**, 588–601 (2013)
6. Mauss, I.B., Robinson, M.D.: Measures of emotion: a review. *Cogn. Emot.* **23**, 209–237 (2009)
7. Spears, N., Singh, S.N.: Measuring attitude toward the brand and purchase intentions. *J. Curr Issues Res. Advert.* **26**, 53–66 (2004)
8. Gregg, A.P., Klymowsky, J., Owens, D., Perryman, A.: Let their fingers do the talking? Using the implicit association test in market research. *Int. J. Mark. Res.* **55**, 487 (2013)
9. Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K.: Measuring individual differences in implicit cognition. The implicit association test. *J. Pers. Soc. Psychol.* **74**, 1464–1480 (1998)

10. Ekman, P., Friesen, W.V., Ellsworth, P.: *Emotion in the Human Face. Guidelines for Research and an Integration of Findings*. Pergamon Press, Elmsford, NY (1972)
11. Holodynski, M.: Die Entwicklung von Emotion und Ausdruck. Vom biologischen zum kulturellen Erbe. ZIF-Mitteilungen 11–27 (2004)
12. Abraham, A., Sommerhalder, K., Abel, T.: Landscape and well-being. A scoping study on the health-promoting impact of outdoor environments. *Int. J. Pub. Health* **55**, 59–69 (2010)
13. Bradley, M.M., Lang, P.: The international affective picture system (IAPS) in the study of emotion and attention. *Handbook of Emotion Elicitation and Assessment*, p. 29 (2007)
14. Lane, K.A., Banaji, M.R., Nosek, B.A., Greenwald, A.G.: Understanding and using the implicit association test: IV. In: Wittenbrink, B., Schwarz, N. (eds.) *Implicit Measures of Attitudes*, pp. 59–102. Guilford Press, New York (2007)
15. Bradley, M.M., Miccoli, L., Escrig, M.A., Lang, P.J.: The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **45**, 602–607 (2008)
16. McFarland, S.G., Crouch, Z.: A cognitive skill confound on the implicit association test. *Soc. Cognit.* **20**, 483–510 (2002)
17. Geiser, M., Walla, P.: Objective measures of emotion during virtual walks through urban environments. *Appl. Sci.* **1**, 1–11 (2011)
18. Grahl, A., Greiner, U., Walla, P.: Bottle shape elicits gender-specific emotion. A startle reflex modulation study. *Psychology* **03**, 548–554 (2012)
19. Walla, P., Rosser, L., Scharfenberger, J., Duregger, C., Bosshard, S.: Emotion ownership. Different effects on explicit ratings and implicit responses. *Psychology* **04**, 213–216 (2013)

# Inferring Web Page Relevance Using Pupillometry and Single Channel EEG

Jacek Gwizdka

**Abstract** We continue investigating neuro-physiological correlates of information relevance decisions and report on research-in-progress, in which we study health-related information search tasks conducted on open web. Data was collected using an eye-tracker and a single-channel EEG device. Our findings show significant differences in pupil dilation on visits and revisits to relevant and irrelevant pages. Significant differences in EEG-measured power of alpha frequency band and in EEG-detected attention levels were also found in a few conditions. The results confirm feasibility of using pupil dilation and suggest plausibility of using low-cost EEG devices to infer relevance.

**Keywords** Information search · Relevance · Eye-tracking · Pupillometry  
EEG

## 1 Introduction and Related Work

Relevance remains a central construct for information search and retrieval (IS&R) systems. Saracevic, one of the key scholars in the area, reminds us that “relevance is timeless” [1] (p. 152) and so, the concern with better understanding factors affecting relevance decisions and the associated cognitive processes continues to be important. Recent years have seen increased research efforts that bring neuro-physiological methods to investigating cognitive aspects of relevance (e.g., [2–8]). We aim to contribute to these efforts. We have previously reported earlier studies, including an fMRI study that examined differences in brain activations between reading relevant and irrelevant texts [6] and an eye-tracking study, in which we showed measures correlated with processing of relevant and irrelevant texts and words [7, 9].

We present an exploratory analysis of pupil dilation and selected EEG data with a focus on finding differences between relevant and not relevant web pages.

---

J. Gwizdka (✉)

School of Information, University of Texas at Austin, Austin, TX, USA  
e-mail: neurois2017@gwizdka.com

**Related Work.** Key works that demonstrated usefulness of features extracted from eye-tracking data (EYE) as relevance indicators include Ajanki et al. [10], who used EYE features as implicit relevance feedback; Buscher et al. [11], who established relationship between several EYE features and text passage relevance; Simola et al. [12] who used EYE features to improve classification of processing states on three simulated information search tasks (word search, question-answer, and subjective interest); Gwizdka [7], who showed that reading irrelevant documents imposed lower mental load than relevant ones and achieved binary classification accuracy of 75%; and Gwizdka et al. [13], who used more sophisticated feature selection and classification and achieved binary relevance classification of paragraph-long texts using EYE features alone of up to 95%.

Pupil dilation is controlled by the Autonomic Nervous System [14]. Under constant illumination it has been associated with a number of cognitive functions, including mental workload [15], interest [16], surprise [17], and making decisions [18]. Generally, the sources of variation in pupil size are related to attention [19, 20]. Past work has shown relevance effects on pupil dilation. In [21] Oliveira et al. reported pupil dilatation for higher relevance stimuli for images. Gwizdka et al. investigated relevance of short text documents [7] and Wikipedia pages [22] and showed significant pupil dilation on relevant documents and particularly so in the one-two second period preceding relevance decision. They also showed significant differences in pupil dilation on fixations on relevant words [9].

Inexpensive EEG devices (e.g., Emotiv EPOC, NeuroSky) have been used successfully in research [23–27]. Such inexpensive devices have been occasionally compared with medical grade EEG devices. For example, Bobrov [24] found classification performance of data collected by EPOC was comparable to BrainProducts ActiCap in a task with recognition of two image types (face or house) and a relaxation state (3-class accuracy in EPOC vs. ActiCap: overall 48 vs. 54%; best 62 vs. 68%). Single channel EEG, Myndplay Brainband XL, was used in a study of sustained attention [28]. The authors showed correlations between low and high alpha power and participant behavior. However, the attention metric provided by the device was found not to be correlated.

In the area of inferring information relevance, we are aware of only one project [29], which showed plausibility of employing an inexpensive EEG device (Emotiv EPOC) in relevance classification with binary classification accuracy of up to 74% [13].

We extend previous results to more realistic search scenarios, in which tasks are conducted on open web. We also collect data from a single-channel EEG device, which has not been previously used to study user interaction with information search systems. While the presented analysis builds on previous results, it is exploratory in nature. Our research questions are as follows: *RQ1*. Does pupil dilation differ between relevant and irrelevant pages and initial visits and re-visits to these pages? *RQ2*. Do attention-related measures derived from single-channel EEG differ between these conditions too?

## 2 Method

We conducted a lab experiment in Information eXperience (IX) lab at University of Texas at Austin ( $N = 30$ ). Due to technical issues data was available for 26 participants (16 females; mean age 24.5). Participants were pre-screened for their native level of English, very good, uncorrected eye-sight, and non-expert topic familiarity. Each participant first performed a training task, which was followed by three search tasks (two assigned tasks and one self-generated) on health-related topics in fully counterbalanced order (i.e.  $3!$  yielding six rotations). The assigned tasks followed a simulated work-task approach [30] and were created to be complex (Table 1). We previously reported from this study data analysis results that were focused on participants' working memory span in relation to search effort [31]. The data analysis presented here has never been published.

Participants performed search tasks on publicly available websites and were asked to bookmark relevant web pages. Search results were retrieved from Google in real-time in the background by a dedicated proxy server. Architecture of our system is described in [32]. The search results were displayed on a custom search engine result page. We limited the number of results to seven per page to increase font size and ensure that eye fixations can be tracked accurately on individual search result snippets. Each user session typically lasted from 1.5 to 2 h. At the completion of a session, each participant received \$25.

**Apparatus.** Eye-tracking data was collected using remote eye-tracker Tobii TX-300.<sup>1</sup> EEG data was collected by single-channel EEG device—Myndplay Brainband XL<sup>2</sup> (Fig. 1) based on NeuroSky<sup>3</sup> chip and dry sensor technology. The one channel is located approximately at Fpz (prefrontal cortex) in 10/20 system. The collected brain signals are expected to reflect some aspects of executive control (e.g., decision making). The EEG hardware uses a sampling rate of up to 512 Hz. The hardware also calculates power of brainwave signal in selected frequency bands and measures “intensity” of being in three mental states (“attention”, “meditation” and “zone”). We used signal power of low and high alpha frequency band and output from NeuroSky's proprietary Attention Meter algorithm. The algorithm is described as indicating intensity of mental “attention.” Past research found increased activity in the alpha band to be correlated with lapses of attention [33]. Therefore, this frequency was considered as a potential marker for decreased attention.

**Dependent variables** included: (1) pupil dilation (pd); (2) EEG-derived low- $\alpha$  (la) and high- $\alpha$  (ha) frequency band power and attention metric (a)—both are calculated by the NeuroSky hardware once per second from 512 Hz raw EEG data. Eye-tracking data was cleaned by removing bad quality fixations (as marked by Tobii). To eliminate individual variability in pupil sizes and in magnitude of brain

---

<sup>1</sup><http://www.tobii.com/product-listing/tobii-pro-tx300/>.

<sup>2</sup><https://myndplay.com/>.

<sup>3</sup><http://neurosky.com/>.



**Table 1** Search tasks

---

**Task 1–Vitamin A:** *Your teenage cousin has asked your advice in regard to taking vitamin A for health improvement purposes. You have heard conflicting reports about the effects of vitamin A, and you want to explore this topic in order to help your cousin. Specifically, you want to know:*

---

(1) *What is the recommended dosage of vitamin A for underweight teenagers?*

---

(2) *What are the health benefits of taking vitamin A? Please find at least 3 benefits and 3 disadvantages of vitamin A*

---

(3) *What are the consequences of vitamin A deficiency or excess? Please find 3 consequences of vitamin A deficiency and 3 consequences of its excess*

---

(4) *Please find at least 3 food items that are considered as good sources of vitamin A*

---

**Task 2–Hypotension:** *Your friend has hypotension. You are curious about this issue and want to investigate more. Specifically, you want to know:*

---

(1) *What are the causes of hypotension?*

---

(2) *What are the consequences of hypotension?*

---

(3) *What are the differences between hypotension and hypertension in terms of symptoms? Please find at least 3 differences in symptoms between them*

---

(4) *What are some medical treatments for hypotension? Which solution would you recommend to your friend if he/she also has a heart condition? Why?*

---

**Examples of user self-generated tasks**

---

(a) *Chrohn’s disease—I know someone who was recently diagnosed and am curious about the disease*

---

(b) *Causes, symptoms, and treatments of Lyme Disease*

---

(c) *What is the recommended daily protein intake? Does it differ among sexes and ages? What are the consequences of excess or lack of proper amount?*

---

(d) *Health benefits of muscles. I am working on body building and losing fat but do not know the specific benefits of muscle besides the fact that it lowers my overall body fat percentage*

---

**Fig. 1** Myndplay Brainband XL.

Source <https://store.myndplay.com/products.php>



waves, we calculated a baseline for each participant ( $i$ ) by taking an average measurement over all tasks ( $x_{baseline}^i$ ) and calculating relative change in each measurement ( $rx_t^i$ ) from measurement at a time  $t$  ( $x_t$ ) as shown in Eq. (1).

$$rx_t^i = (x_t - x_{baseline}^i) / x_{baseline}^i \tag{1}$$

where  $x$  can be: pupil dilation, low-, high-alpha band brainwave signal power or attention meter values.

In this analysis, we only use data recorded during visits to content web pages, that is pages that were opened from search engine result pages (SERPs), or by following links from other content pages. Mean values of relative measures were calculated on each visit to web pages for *three types of two-second-long epochs*: (1) epochs at the beginning of page visit, (2) epochs at the end of page visit; and (3) epochs before relevance judgment (e.g., before bookmarking a page). To avoid possible influence of motor actions, the first or last 200 ms of after or before an event (page open/close) were not included. The choice of two-second-long epochs was informed by past research (e.g., [2, 7, 22]). In particular, Gwizdka et al. [13] found the best classification performance of EEG features for two-second-long epochs.

**Controlled variable** was constructed as a nine-level factor (Factor9). First, we labelled sequential data with five labels created from time segments when (1) irrelevant web page was visited first, (2) irrelevant web page was revisited; the same

**Table 2** Five-level factor for page relevance and visits types with counts of page visits

Level	Page relevance	Visit types	Count of visits	Count of visit types
1	Irrelevant page	First visit	398	258
2		Revisit		140
3	Relevant page*	First visit*	294	50
4		Revisit*		244
5	Rel. page visit with relevance judgment			295

\*Counts after removing visits when relevance judgments were made, which are contained in level 5

**Table 3** Relative pupil dilation on different page types, visits and epochs

Factor9	Factor9 level names	N	Mean	SD
pd1	Irrelevant page—start of first visit	248	-0.0381	0.054
pd2	Irrelevant page—start of re-visit	114	0.0085	0.088
pd3	Irrelevant page—end of first visit	249	-0.0366	0.092
pd4	Irrelevant page—end of re-visit	114	0.0217	0.092
pd5	Relevant page—start of first visit	48	-0.0369	0.045
pd6	Relevant page—start of re-visit	203	0.0043	0.094
pd7	Relevant page—end of first visit	48	-0.0253	0.065
pd8	Relevant page—end of re-visit	204	0.0088	0.086
pd9	Relevant page—relevance judgment	288	0.0018	0.082

was applied to time segments of visits to relevant pages. A separate label was applied to visits of relevant web pages when relevance judgment was made. This labelling yielded a five-level factor (Table 2). The first four levels were then combined with the type of epoch (start and end), yielding eight combinations. The third epoch type (before relevance judgment) was applicable only to the fifth level; it yielded the ninth level (Table 3).

### 3 Data Analysis and Results

Due to not-normal distribution of variables, we analyzed them using non-parametric tests (Kruskal–Wallis (K–W) and Mann–Whitney  $U$  (M–W)). The relative pupil dilation was significantly different between Factor9 levels (K–W  $\chi^2(8) = 157.8$ ,  $p < .001$ ). However, none of the EEG derived measures were.

Next, we run pair-wise comparisons of all variables and compared all combinations of individual Factor9 levels (except the same levels) (Table 4). Partially confirming previous results [22], pupil dilation (pd) was significantly larger on visits with relevance judgements in comparison with first visits to relevant pages (3.7% larger for start of visits (pd9, pd5), and 2.7% larger for end of visits (pd9, pd7), but not in comparison with re-visits to pages. Each comparison pair of re-visit to first-visit was significantly different, with pupils more dilated on revisits (pd2, pd1-by 4.7%, pd4, pd3–5.8%, pd6, pd5–4.1%, pd8, pd7–3.4%). The only two significant pair-wise comparison for EEG data were for relevance judgements in comparison with starts of first-visits and re-visits to not relevant pages (a9, a1-relative attention increased by 10.7% on relevance judgements, and for a9, a2-by 3.8%; la9, la2-power of lower alpha decreased by 27% on relevance judgements).

**Table 4** Significant pair-wise comparisons (Mann–Whitney  $U$  tests)

Factor9	1	2	3	4	5	6	7	8	9
1	–	–	–	–	–	–	–	–	–
2	pd2, pd1***	–	–	–	–	–	–	–	–
3			–	–	–	–	–	–	–
4			pd4, pd3***	–	–	–	–	–	–
5					–	–	–	–	–
6					pd6, pd5***	–	–	–	–
7							–	–	–
8							pd8, pd7*** <sup>2</sup>	–	–
9	a9, a1**	a9, a2 <sup>b</sup> , la9, la2***			pd9, pd5***		pd9, pd7*** <sup>2</sup>		–

<sup>b</sup>0.05 <  $p$  < 0.1, \*\*  $p$  < 0.01, \*\*\*<sup>2</sup>  $p$  = 0.002, \*\*\*  $p$  < 0.001

Dependent variables for which significant effect was found

pd( $n, m$ )—pupil dilation; a( $n, m$ )—attention metric; la( $n, m$ )—low alpha

where  $n, m$  are Factor9 levels for which pair-wise comparison a significant difference was found

## 4 Discussion and Conclusions

The presented analysis is preliminary and exploratory in nature. Our results demonstrate, to some extent, the expected differences in pupil dilation between first- and re-visits to relevant and irrelevant pages (RQ1). Compared with [7, 22], we found fewer such differences. One reason is that our stimuli (open web) and user interaction were more realistic and thus “noisier”. The EEG-derived measures were not found significantly different, except for a couple of conditions (RQ2). First, we used a device with single channel only. Another plausible reason may be our choice of two-second-long epochs. While such epochs worked well for features calculated from raw EEG data [13, 27], they seem to be less appropriate when used with data calculated at second-long-intervals using proprietary algorithms. When significant, however, the values of EEG measures differed in the expected direction.

In future work, we plan to create more detailed segments of user activities on web pages, thus separating reading/viewing from looking at task description, bookmarking and taking notes. We also plan to use raw EEG data instead of relying on proprietary algorithms and to conduct more detailed analysis of pupillary responses.

**Acknowledgements** This research was supported, in part, by IMLS Career award to Jacek Gwizdka # RE-04-11-0062-11.

## References

1. Sonnenwald, D.H.: Theory Development in the Information Sciences. University of Texas Press, Austin, TX (2016)
2. Allegretti, M., Moshfeghi, Y., Hadjigeorgieva, M., Pollick, F.E., Jose, J.M., Pasi, G.: When relevance judgement is happening? An EEG-based study. In: Proceedings of SIGIR’2015, pp. 719–722. ACM, New York (2015)
3. Moshfeghi, Y., Pinto, L.R., Pollick, F.E., Jose, J.M.: Understanding relevance: an fMRI study. In: Serdyukov, P., Braslavski, P., Kuznetsov, S.O., Kamps, J., Rüger, S., Agichtein, E., Segalovich, I., Yilmaz, E. (eds.) Advances in Information Retrieval (ECIR’2013), pp. 14–25. Springer, Berlin (2013)
4. Barral, O., Kosunen, I., Ruotsalo, T., Spapé, M.M., Eugster, M.J.A., Ravaja, N., Kaski, S., Jacucci, G.: Extracting relevance and affect information from physiological text annotation. *User Model User-Adapt. Interact.* **26**, 493–520 (2016)
5. Eugster, M.J.A., Ruotsalo, T., Spapé, M.M., Barral, O., Ravaja, N., Jacucci, G., Kaski, S.: Natural brain-information interfaces: recommending information by relevance inferred from human brain signals. *Sci. Rep.* **6**, 38580 (2016)
6. Gwizdka, J.: Looking for information relevance in the brain. In: Gmunden Retreat on NeuroIS 2013, p. 14. Gmunden, Austria (2013)
7. Gwizdka, J.: Characterizing relevance with eye-tracking measures. In: Proceedings of IiiX’2014, pp. 58–67. ACM, New York (2014)
8. Gwizdka, J.: Tracking information relevance. In: Gmunden Retreat on NeuroIS 2014, p. 3. Gmunden, Austria (2014)

9. Gwizdka, J.: Differences in reading between word search and information relevance decisions: evidence from eye-tracking. In: Davis D.F., Riedl R., vom Brocke J., Léger P.-M., Randolph B. A. (eds.) *Information Systems and Neuroscience: Gmunden Retreat on NeuroIS 2016*, pp. 141–147. Springer, Berlin (2017)
10. Ajanki, A., Hardoon, D., Kaski, S., Puolamäki, K., Shawe-Taylor, J.: Can eyes reveal interest? Implicit queries from gaze patterns. *User Model User-Adapt. Interact.* **19**, 307–339 (2009)
11. Buscher, G., Dengel, A., Biedert, R., Elst, L.V.: Attentive documents: eye tracking as implicit feedback for information retrieval and beyond. *ACM Trans. Interact. Intell. Syst.* **1**, 9:1–9:30 (2012)
12. Simola, J., Salojärvi, J., Kojo, I.: Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cogn. Syst. Res.* **9**, 237–251 (2008)
13. Gwizdka, J., Shouyi, W., Cole, M., Hosseini, R.: Temporal dynamics of eye-tracking and EEG during reading and relevance decisions. *J. Assoc. Inf. Sci. Technol.* (in press)
14. Onorati, F., Barbieri, R., Mauri, M., Russo, V., Mainardi, L.: Characterization of affective states by pupillary dynamics and autonomic correlates. *Front Neuroeng.* **6**, 9 (2013)
15. Kahneman, D., Beatty, J.: Pupil diameter and load on memory. *Science* **154**, 1583–1585 (1966)
16. Krugman, H.E.: Some applications of pupil measurement. *JMR J. Mark. Res. Pre-1986.* **1**, 15 (1964)
17. Preusschoff, K., Hart, B.M., Einhäuser, W.: Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Front. Decis. Neurosci.* **5**, 115 (2011)
18. Einhäuser, W., Koch, C., Carter, O.L.: Pupil dilation betrays the timing of decisions. *Front Hum. Neurosci.* **4**, 18 (2010)
19. Hoeks, B., Levelt, W.J.M.: Pupillary dilation as a measure of attention: a quantitative system analysis. *Behav. Res. Methods Instrum. Comput.* **25**, 16–26 (1993)
20. Wierda, S.M., van Rijn, H., Taatgen, N.A., Martens, S.: Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *PNAS* **109**, 8456–8460 (2012)
21. Oliveira, F.T.P., Aula, A., Russell, D.M.: Discriminating the relevance of web search results with measures of pupil size. In: *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, pp. 2209–2212. ACM, Boston, MA (2009)
22. Gwizdka, J., Zhang, Y.: Differences in eye-tracking measures between visits and revisits to relevant and irrelevant web pages. In: *Proceedings of SIGIR'2015*, pp. 811–814. ACM, New York (2015)
23. Abbott, W.W., Faisal, A.A.: Ultra-low-cost 3D gaze estimation: an intuitive high information throughput compliment to direct brain-machine interfaces. *J. Neural Eng.* **9**, 046016 (2012)
24. Bobrov, P., Frolov, A., Cantor, C., Fedulova, I., Bakhnyan, M., Zhavoronkov, A.: Brain-computer interface based on generation of visual images. *PLoS ONE* **6**, e20674 (2011)
25. Khushaba, R.N., Greenacre, L., Kodagoda, S., Louviere, J., Burke, S., Dissanayake, G.: Choice modeling and the brain: a study on the electroencephalogram (EEG) of preferences. *Expert Syst. Appl.* **39**, 12378–12388 (2012)
26. Ramirez, R., Palencia-Lefler, M., Giraldo, S., Vamvakousis, Z.: Musical neurofeedback for treating depression in elderly people. *Audit. Cogn. Neurosci.* 354 (2015)
27. Wang, S., Gwizdka, J., Chaovalitwongse, W.A.: Using wireless EEG signals to assess memory workload in the n-back task. *IEEE Trans. Hum.-Mach. Syst.* **46**, 424–435 (2016)
28. Wal, C.N. van der, Irrmischer, M.: Myndplay: measuring attention regulation with single dry electrode brain computer interface. In: Guo, Y., Friston, K., Aldo, F., Hill, S., and Peng, H. (eds.) *Brain Informatics and Health*. pp. 192–201. Springer, Berlin (2015)
29. Gwizdka, J., Hosseini, R., Cole, M., Wang, S.: Temporal dynamics of eye-tracking and EEG during reading and relevance decisions. *J. Assoc. Inform. Sci. Tech.* **68**(10), 2299–2312 (2017)
30. Borlund, P.: The concept of relevance in IR. *J. Am. Soc. Inf. Sci. Technol.* **54**, 913–925 (2003)

31. Gwizdka, J.: I can and so I search more: effects of memory span on search behavior. In: Proceedings of CHIIR'2017, pp. 341–344. ACM, New York (2017)
32. Wei, X., Zhang, Y., Gwizdka, J.: YASFIIRE: yet another system for IIR evaluation. In: Proceedings of IiiX'2014, pp. 316–319. ACM, New York (2014)
33. O'Connell, R.G., Dockree, P.M., Robertson, I.H., Bellgrove, M.A., Foxe, J.J., Kelly, S.P.: Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J. Neurosci.* **29**, 8604–8611 (2009)

# Measuring Biosignals of Overweight and Obese Children for Real-Time Feedback and Predicting Performance

Nurten Öksüz, Russa Biswas, Iaroslav Shcherbatyi  
and Wolfgang Maass

**Abstract** Child obesity is a serious problem in our modern world and shows an increase of 60% since 1990. Due to time and cost intensity of traditional therapy programs, scientists started to focus on IT-based interventions. Our paper focuses on measuring biosignals (e.g. heart rate) of obese children during fittest including different physical activities (e.g. running). We investigate whether it is possible to predict the performance of obese children during running tests based on static (e.g. BMI) as well as dynamic (e.g. heart rate) parameters. Here, we focused on heart rate-related parameters from the inverted U-shaped heart rate response of obese children during running tests. For future research, we plan to consider physical activity (e.g. step count) of the children at home. Our approach is a NeuroIS service, which uses low-cost devices making prediction on an individual's future development and is also applicable to other domains (e.g. business information systems).

**Keywords** Heart rate · Obesity · Children · Fitness · Prediction

---

N. Öksüz (✉) · R. Biswas · I. Shcherbatyi · W. Maass  
Information and Service Systems, Department of Law and Economics, A5 4,  
Saarland University, 66123 Saarbrücken, Germany  
e-mail: nurten.oeksuez@iss.uni-saarland.de

R. Biswas  
e-mail: biswasrussa@gmail.com

I. Shcherbatyi  
e-mail: iaroslav.shcherbatyi@dfki.de

W. Maass  
e-mail: wolfgang.maass@iss.uni-saarland.de; wolfgang.maass@dfki.de

I. Shcherbatyi · W. Maass  
Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH,  
Campus D3 2, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany

## 1 Introduction

Child obesity has become a serious problem in our modern world with an increasing trend. According to Ogden et al. [1], the percentage of obese children aged 6–11 years in the United States increased from 7% in 1980 to 18% in 2012. Obese children and adolescents aged 12–19 years are more affected with an increase from 5% in 1980 to 21% in 2012 [1]. It has been observed that, besides psychological and physiological impacts, obesity has a lot of serious implications for the public and private healthcare sectors, for example dramatically increasing public health costs and obese children having risk factors for cardiovascular diseases [2, 3]. Therefore, several efforts are needed to control the persistent epidemic of overweight and obesity [4]. Results of holistic therapy programs have shown positive effects on therapy outcomes of obese children [5–8]. However, these interventions are time and cost consuming for both the patients and physicians. To counteract this problem, scientists started to focus on IT-based interventions by using pervasive and smart technologies. There are several mobile solutions, which measure vital signs and use sensors for daily use to help control obesity [9–11]. While most of the existing solutions measure physical activity, there are a few which use methods from the NeuroIS field. These NeuroIS tools and methods can be used to create efficient time and cost saving solutions tailored to patients' individual needs. However, there are, to the best of the authors' knowledge, no papers focusing on NeuroIS solutions which make predictions on the performance (e.g. running, push-up, curl-up, trunk lift) of obese children based on static (e.g. BMI, age, gender) as well as dynamic (e.g. heart rate, skin conductance, blood pressure) parameters. The goal of this paper is to investigate whether the performance of obese children during running tests can be predicted using static as well as dynamic parameters. In order to obtain the dynamic parameters, we conducted a fittest, which included a 6-minute running test. Several parameters including heart rate are measured during the fittest. In our study, we take *BMI* and *gender* as static parameters and *average heart rate* during the 6-minute running test as well as the *heart rate recovery* after the exercise as the dynamic parameters. We measured performance by counting the number of laps made during the 6-minute running test. The research question is as follows:

Is it feasible to predict the performance of overweight and obese children with the help of the static parameters BMI and gender as well as the dynamic parameters average heart rate during running test and heart rate recovery?



## 2 State of the Art

### 2.1 Relationship of Body Mass Index and Fitness Level

Several study results indicate a relationship between the *Body Mass Index* (BMI) and the fitness of an individual [12–15]. The study of Joshi et al. with a sample size of  $n = 7000$  school children doing a physical fitness assessment called *Fitnessgram* concludes that the fitness level of children having healthy BMIs is the highest, followed by those of overweight and obese children [12]. The results show that the higher the BMI, the less likely a child tends to be physically fit [12]. Physical fitness was measured by considering the number of exercises scored in the *healthy fitness zone* (HFZ) [12]. The study of Aires et al. [15] also strengthened these results by finding out that obese children between 11 and 18 years old performed a decreased number of tests in the HFZ compared to the normal-weight children, indicating a reduced performance in both physical strength and cardiovascular fitness.

### 2.2 Heart Rate During and After Exercise and Its Relationship to Fitness

Exercise heart rate as well as the post-exercise heart rate can give information about an individual's fitness level [16]. Physical activity or exercising elevates the heart rate for the duration of physical activity and slows it down during the cool down after the physical activity [17, 18]. The fitter an individual is, the lower the heart rate will be during training, the lower it will be during cool down and the faster it will return to the pre-exercise level [16]. Repeating the test after a certain period of time will create a comparable set of results that can be used to detect changes of an individual's fitness [16]. The *heart rate recovery* (HRR) depends on, amongst others, the intensity of the exercise and the cardiorespiratory fitness of an individual [19, 20]. Obese children and adolescents tend to have lower cardiorespiratory fitness and physical abilities when compared to normal-weight children and young adults, mainly due to increased effort required to carry the large amount of body fat and to move their larger body mass [21]. Furthermore, Singh et al. conducted a study using a maximal treadmill exercise to compare the HRR of normal-weight and overweight children [22]. The results show that children with higher BMI, especially those who are overweight, have slower 1-minute HRR after exercise [22].

### 2.3 *Pervasive and Smart Technologies for Controlling Obesity*

There are several approaches, which focus on measuring vital signs and applying sensors to help control obesity. BALANCE is able to automatically calculate the calories spent in the everyday activities by using inertial sensors, which are worn on the body. Nevertheless, the patient has to manually enter the calorie content of the single food items [9]. HealthAware uses GPS and accelerometer embedded in a smartphone to monitor activities and a camera to additionally analyze food items intake. The user needs to manually enter name of the single food items and the system will calculate the calorie based on the collected data [10]. UbiFit Garden uses classifiers trained to differentiate walking, running, and cycling using a stairs machine as well as an elliptical trainer by means of barometer and 3-d accelerometer to encourage physical activity [11]. TripleBeat is a NeuroIS service, which consists of accelerometer to measure movements during run as well as *Electrocardiogram* (ECG) sensors to monitor the heart rate [23]. ExerTrek monitors exercise heart rate during exercise and give real-time online feedback about the user's heart status and any occurring abnormalities. Furthermore, there are many commercial solutions (e.g. RunKeeper, Sportypal, and Runtastic PRO) available that track the activities of the users to help them lose weight.

## 3 Research Methodology

Our study is conducted at a Swiss children's hospital in St. Gallen in cooperation with four universities. In total, 20 children aged between 11 and 17 years with higher BMI values ( $25 < \text{BMI} < 37$ ) participated in the fittest (7 female and 13 male). The participants have taken the fittest in a sports hall, which consisted of exercise elements of the Dordel Koch Fittest and the EuroFit Fitness Test. The fittest included a running test in the last 6 min. For every child, the number of laps was counted. Besides BMI, gender, number of steps and number of laps the children ran during the 6 min. running test, the exercise heart rate (about 25 min) as well as the post-exercise heart rate (cool down period of 3 min) were measured. To measure the heart rate, the participants were equipped with a Scosche Rhythm + heart rate monitor and a Samsung Galaxy S6 smartphone, in which the app *PathMate2* is installed. The app *PathMate2* collects the heart rate data from the heart rate monitor when the *Exercise Button* or the *Cool Down Button* is pressed and sends the data to the server, where the data is processed for further explorative and predictive analysis. Before the participants start the exercise, the *Exercise Button* was pressed to measure the initial heart rate as well as the heart rate during the exercise. Right after the exercise, the *Cool Down Button* was pressed to separately measure the heart rate of the participants during the cool down.

For the purpose of predictive analysis, we calculated the average heart rate during steady state as the *average heart rate* during the running test (see Fig. 1). Furthermore, the heart rate difference between the start of the cool down and the average of the last 10 values of the cool down is taken as the *heart rate recovery*.

### 4 Results

We intend to predict the number of laps made during the 6-minute running test using the linear regression model Least Absolute Shrinking and Selection Operator (LASSO). The features used to train the model are *BMI*, *gender*, *average heart rate* during the running and *heart rate recovery*. We used k-fold cross validation with  $k = 3$  to select the best parameter ( $\alpha$ ) for the model. The data set is divided into train and test set using Leave One Out Cross Validation (LOOCV). Each data point (child) in the data set is used once as a test set (singleton) while the remaining data is used as the train set.

The results show that the average difference between the actual number of laps and the number of laps predicted by our model is 2.185 with an overall average error of 7.1%. Our model makes better prediction of the number of laps made during the 6-minute running test compared to the constant value model (baseline method). The constant value model considers the average number of laps in the train set as the predicted number of laps for the test set. The Mean Squared Error (MSE) of our model (MSE = 6.892) is better than the MSE of the constant value model (MSE = 10.052).

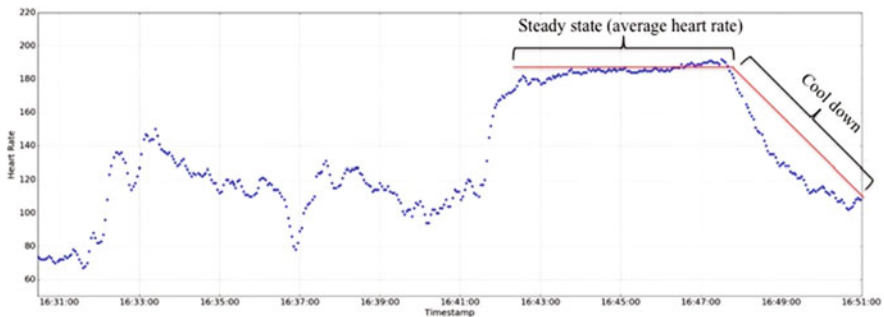


Fig. 1 Heart rate during the fittest

## 5 Discussion

The goal of this paper was to investigate the feasibility of predicting the performance of overweight and obese children by the number of laps made during a 6-minute running test using the static parameters *BMI* and *gender* as well as the dynamic parameters *average heart rate* during the running test and *heart rate recovery*. As mentioned above, the study of Joshi et al. concludes a casual relationship between the BMI and the fitness level of the children [12]. The fitness level of a specific child was measured in terms of performance during several exercises (e.g. push-up, curl-up, running). In our study, we measured performance by counting the number of laps made during a 6-minute running test. Furthermore, the heart rate recovery of children after exercise has a causal relationship with their fitness level (see Sect. 2.2). Therefore, in our predictive analysis study we included the static parameter BMI as well as the dynamic parameter *heart rate recovery* as features to predict the performance of the children during the running test. We also included *average heart rate* during the running test for the analysis since heart rate during the exercise is another dynamic parameter, which serves as an indicator of an individual's fitness level (see Sect. 2.2). The quantitative results as mentioned in Sect. 3 depicts that our method works better than the baseline method, which is the constant value model. Thus, our method provides a tentative prediction on the performance of the observed children. Despite the fact that the MSE of our model is better than that of the baseline method, the accuracy can still be improved. The accuracy of our model is influenced by several factors. First, our data set might be biased due to the fact that the study is done only on overweight and obese children. This implies that all the children in the data set exhibit almost similar health characteristics. Second, the data set is very small leading to chances of overtraining the model. To overcome it in the best possible way we used LOOCV to divide the data into train set and test set. Third, the children exerted themselves to different extents during the 6-minute running test. Nevertheless, despite of these drawbacks the linear model used in our study is capable of predicting the number of laps with an overall error of 7.1%.

## 6 Conclusion and Outlook

It can be concluded from our analysis that pre-exercise and post-exercise heart rate as well as BMI and gender can be leveraged to predict the number of laps made by the children during the running test. Therefore, low-cost wearable devices along with predictive analysis methods allow predicting health conditions reducing the cost of the traditional therapy programs. In future work, we intend to focus on applying the method to a large data set including obese, overweight and normal children to improve the accuracy of prediction. Furthermore, we plan to use a standardized fitness test (e.g. treadmill running test) to provide all the children with

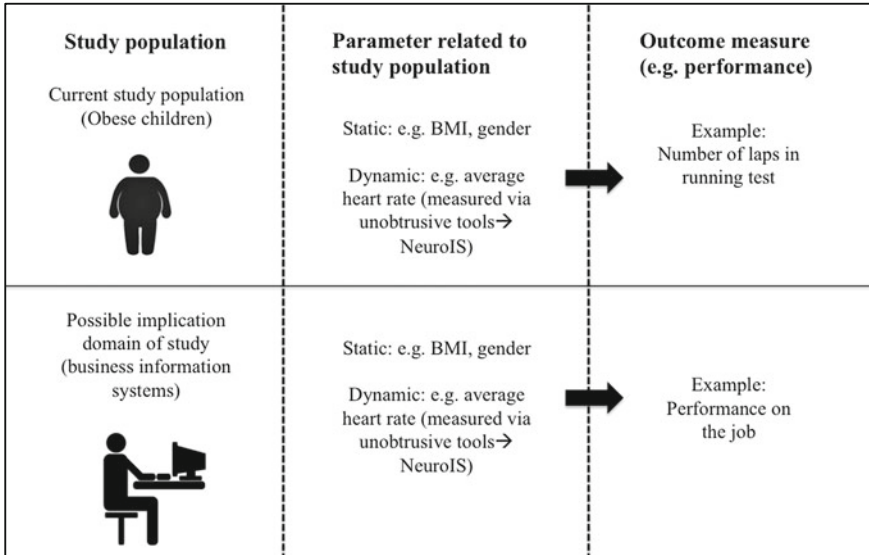


Fig. 2 Current and possible study population and predicted performance

the same fittest environment. However, taking other static (e.g. age) and dynamic parameters (e.g. blood pressure) into account can also lead to other interesting predictions. In future, similar predictive analysis methods could open up new areas of remote patient monitoring and interventions as well as other domains using low-cost devices such as smartphones and smartwatches (see Fig. 2).

For instance, this approach is not only interesting for child obesity, but also applicable to fields such as business information systems domain, since biosignals of employees on the job is a current topic. Kowatsch [24] for example suggests measuring physiological arousal of employees on the job to detect a relationship between job strain and task performance. For our purpose, measuring the heart rate of employees on the job to predict the task performance on the job could be the prime focus.

## References

- Ogden, C.L., Carroll, M.D., Kitt, B.K., Flegal, K.M.: Prevalence of childhood and adult obesity in the United States, 2011–2012. *JAMA* **311**, 806–814 (2014)
- Hänggli, B., Laimbacher, J., Gutzwiller, F., Fäh, D.: Adipositas therapie bei Kindern und Jugendlichen in der Schweiz: Wo liegen die Hürden. *Schweizerische Ärztezeitung* **89**, 277–281 (2008)
- Wang, Y.C., McPherson, K., Marsh, T., Gortmaker, S.L., Brown, M.: Health and economic burden of the projected obesity trends in the USA and the UK. *The Lancet* **378**, 815–825 (2011)

4. Murer, S.B., Saarsalu, S., Zimmermann, M.B., Aeberli, I.: Pediatric adiposity stabilized in Switzerland between 1999 and 2012. *Eur. J. Nutr.* **53**, 865–875 (2014)
5. Sempach, R., Farpour-Lambert, N., L'Allemand, D., Laimbacher, J.: Therapie des adipösen Kindes und Jugendlichen: Vorschläge für multi-professionelle Therapieprogramme. *Paediatrica* **18**, 33–37 (2007)
6. Ho, M., Garnett, S.P., Baur, L., Burrows, T., Stewart, L., Neve, M., Collins, C.: Effectiveness of lifestyle interventions in child obesity: systematic review with meta-analysis. *Pediatrics* **130**, e1647–e1671 (2012)
7. Oude, L.H., Baur, L., Jansen, H., Shrewsbury, V.A., O'Malley, C., Stolk, R.P., Summerbell, C.D.: Interventions for treating obesity in children. *Cochrane Database Syst. Rev.* **21**, CD001872 (2009)
8. L'Allemand, D., Kirchhoff, E., Bolten, M., Zumbrunn, A., Martin, X., Sempach, R., Farpour-Lambert, N.: Evaluation of therapy for overweight children and adolescents in Switzerland: Therapy in multiprofessional group programs—part 2 of KIDSSSTEP—collection and analysis of data. Final report, pp. 1–83. Federal Office for Public Health, Bern (2014)
9. Denning, T., Andrew, Hartung, C., Lester, J., Borriello, G., Duncan, G.E.: BALANCE: Towards a usable pervasive wellness application with accurate activity inference. In: *Proceedings of the 10th workshop on Mobile Computing Systems and Applications*, p. 5. ACM (2009)
10. Gao, C., Kong, F., Tan, J.: Healthaware: Tackling obesity with health aware smart phone systems. In: *Proceedings of IEEE International Conference on Robotics and Biomimetics*, pp. 1549–1554. IEEE Press, Guilin (2009)
11. Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J.A.: Activity sensing in the wild: A field trial of Ubifit Garden. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1797–1806. ACM, New York (2008)
12. Joshi, P., Bryan, C., Howat, H.: Relationship of body mass index and fitness levels among schoolchildren. *J Strength Cond. Res.* **26**, 1006–1014 (2012)
13. Leyk, D., Rohde, U., Gorges, W., Ridder, D., Wunderlich, M., Dinklage, C., Sievert, A., Rütther, T., Essfeld, D.: Physical performance, body weight and BMI of young adults in Germany 2000–2004: Results of the physical-fitness-test study. *Int. J. Sports Med.* **27**, 247–242 (2006)
14. Rauner, A., Mess, F., Woll, A.: The relationship between physical activity, physical fitness and overweight in adolescents: a systematic review of studies published in or after 2000. *BMC Pediatr.* **13**, 1–19 (2013)
15. Aires, L., Silvia, P., Santos, R., Santos, P., Ribeiro, J.C., Mota, J.: Association of physical fitness and body mass index in youth. *Minerva Pediatr.* **60**, 397–405 (2008)
16. Bird, S.R., Smith, R.A., James, K.: *Exercise Benefits and Prescription*. Nelson Thornes, Cheltenham (1998)
17. Swain, D.P., Leutholtz, B.C.: *Exercise Prescription: A Case Study Approach to the ACSM Guidelines*. Human Kinetics, Champaign (2007)
18. Draper, N., Marshall, M.: *Exercise Physiology: For Health and Sports Performance*. Routledge, New York (2014)
19. Mazzeo, R.S., Marshall, P.: Influence of plasma catecholamine on the lactate threshold during graded exercise. *J. Appl. Physiol.* **67**, 1319–1322 (1989)
20. Ohuchi, H., Tasato, H., Sugiyama, H., Arakaki, Y., Kamiya, T.: Responses of plasma norepinephrine and heart rate during exercise in patients after Fontan operation and patients with residual right ventricular outflow tract obstruction after definitive reconstruction. *Pediatr. Cardiol.* **19**, 408–413 (1998)
21. Dupuis, J.M., Vivant, J.F., Daudet, G., Bouvet, A., Clement, M., Dazord, A., et al.: Personal sports training in the management of obese boys aged 12 to 16 years. *Arch. Pediatr.* **7**, 1185–1193 (2000)
22. Singh, T.P., Rhodes, J., Gauvreau, K.: Determinants of heart rate recovery following exercise in children. *Med. Sci. Sports Exerc.* **40**, 601–605 (2008)

23. Oliveira, R. De, Oliver, N.: TripleBeat: Enhancing exercise performance with persuasion. In: Proceedings of the 10th International conference on Human computer interaction with mobile devices and services, pp. 255–264. ACM (2008)
24. Kowatsch, T.: Emotions in ubiquitous information systems: an empirical investigation of electrodermal activity and its relation to service breakdowns, perceived ease of use, and task performance. Dissertation, University St. Gallen (2016)

# Stationarity of a User's Pupil Size Signal as a Precondition of Pupillary-Based Mental Workload Evaluation

Ricardo Buettner, Ingo F. Scheuermann, Christian Koot, Manfred Rössle and Ingo J. Timm

**Abstract** We discuss the concept of stationarity as a precondition of pupillary-based assessments of a user's mental workload and report results from an experiment differentiating stationarity and non-stationarity pupillary size signals.

**Keywords** NeuroIS · Eye-tracking · Mental workload · Pupil diameter  
Stationarity · Augmented Dickey-Fuller test

## 1 Introduction

While a user's mental workload can be evaluated by pupillary-based eye-tracking [1–8], environmental (e.g. luminescence [9]) and factors other than workload-related mental processes (e.g. emotional arousal [10]) also influence a user's pupil size.

In order to detect luminescence changes, which were reflected in sustainable shifts of the user's pupil size, we apply the concept of stationarity analysis. A stationary process is a stochastic process  $X_t$  whose probability distribution does not change when shifted in time  $t$ :  $E(X_t) = E(X_{t-1}) = \mu$ ;  $\text{Var}(X_t) = \gamma_0 < \infty$ ,  $\text{Cov}(X_t, X_{t-1}) = \gamma_k$ . While a stationary pupil size signal has been discussed as a precondition for assessing mental workload [11, 12], no pupillary signal related guideline exists.

---

R. Buettner (✉) · I.F. Scheuermann · C. Koot · M. Rössle  
Aalen University, Aalen, Germany  
e-mail: ricardo.buettner@hs-aalen.de

I.J. Timm  
Trier University, Trier, Germany



## 2 Methodology

### 2.1 Applying the NeuroIS Guidelines

In order to clearly contribute to NeuroIS research and show strong methodological rigor, we followed the NeuroIS guidelines established by Brocke and Liang [13]. In particular, to assess prior research in the field of measuring mental workload as an important IS concept, a comprehensive literature review was conducted (cf. [14]). To base the experimental design adequately on solid research in related fields of neuroscience [15] we reviewed the fundamental anatomic mechanism of the pupillary dilation controlled by the vegetative nervous system and the key role of the Edinger-Westphal nucleus that is inhibited by mental workload and directly leads to a pupillary dilation. The methodology uses eye-tracking-based pupillometry as a well-established approach in physiology and psychology for “widening the ‘window’ of data collection” [15 p. 93]. With this method, bio-data (i.e. pupil diameter) can be used to better understand mental workload as an IS construct (cf. guideline 4 of [14]). In comparison to other neuroscience tools, eye-tracking-based pupillometry is the contact-free and efficient method of choice [16]. We applied the guidelines and standards from Duchowski [17] and the Eyegaze Edge<sup>TM</sup> manual.

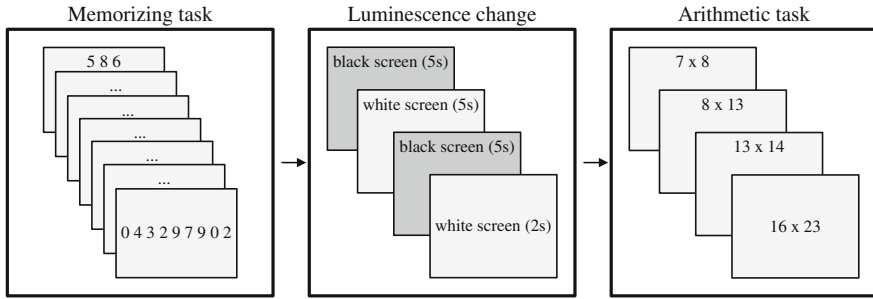
### 2.2 Measurements

To capture the pupillary diameter, eye-tracking was performed using the binocular double Eyegaze Edge<sup>TM</sup> System eye-tracker paired with a 19" LCD monitor (86 dpi) set at a resolution of  $1280 \times 1024$ , whereby the eye-tracker samples the pupillary diameter at a rate of 60 Hz for each eye separately.

### 2.3 Stimuli and Test Procedure

Following Beatty [18] and Hess and Polt [19] we manipulated mental workload using two well-documented experimental settings in psychology. In the first stage our participants had to memorize and reproduce numbers consisting of three to nine digits [18]. In this stage the luminescence changes on the computer screen were small, with only numbers on a bright background presented, which were interrupted by bright and light green break slides. In the second stage, we showed a fixed order of dark and bright screens [black (5 s) → white (5 s) → black (5 s) → white (2 s)] without any mental task. In the third stage the participants had to solve four arithmetic multiplication problems representing a high cognitive demand level as documented [19].

Prior to all data collection, each test participant was welcomed by the experimenter (supervisor of the experiment). After that the participant was asked to fill out



**Fig. 1** Test procedure

a consent form and also a questionnaire with demographics. After that, we took the necessary precautions for the experiment, in which we make use of the eye-tracking system. Hence, the eye-tracker was calibrated. In the next stage, the experiment began with the memorizing task, followed by the luminescence change stage without mental workload before the participants were invited to solve the last mental workload task (arithmetic) (Fig. 1).

## 2.4 Augmented Dicker-Fuller Test

The Augmented Dickey-Fuller (ADF) test evaluates if a time series variable follows a unit-root process. The null hypothesis is that the variable contains a unit root, and the alternative is that the variable is generated by a stationary process.

To calculate the ADF statistics we used the `tseries` package within the R `x64 3.3.3` environment [20].

## 3 Results

### 3.1 Sample Characteristics

Twelve volunteers (six females) aged from 21 to 38 ( $M = 26.2$ ,  $SD = 4.1$ ) participated.

### 3.2 Stationarity Test Results

Table 1 shows the Augmented Dickey-Fuller t-statistic test results while small  $p$ -values suggest that the pupillary size time series is stationary. Vice versa, large  $p$ -values indicate non-stationarity.

Table 1 Stationarity test results

Participant	Mental workload task (memorizing)		Luminescence change (no mental workload)		Mental workload task (arithmetic)	
	Left eye	Right eye	Left eye	Right eye	Left eye	Right eye
1	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.48$ )	Non-stationary ( $p = 0.52$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
2	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	NA (to less data points)	NA (to less data points)	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
3	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.59$ )	Non-stationary ( $p = 0.33$ )	Stationary ( $p < 0.05$ )	Stationary ( $p < 0.1$ )
4	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.50$ )	Non-stationary ( $p = 0.69$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.05$ )
5	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.62$ )	Non-stationary ( $p = 0.47$ )	Stationary ( $p < 0.05$ )	Stationary ( $p < 0.1$ )
6	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.39$ )	Non-stationary ( $p = 0.73$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
7	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.54$ )	Non-stationary ( $p = 0.66$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
8	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.50$ )	Non-stationary ( $p = 0.74$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.05$ )
9	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.28$ )	Non-stationary ( $p = 0.43$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
10	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.37$ )	Non-stationary ( $p = 0.31$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
11	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.44$ )	Non-stationary ( $p = 0.42$ )	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )
12	Stationary ( $p < 0.01$ )	Stationary ( $p < 0.01$ )	Non-stationary ( $p = 0.69$ )	Non-stationary ( $p = 0.80$ )	Stationary ( $p < 0.05$ )	Stationary ( $p < 0.1$ )

## 4 Discussion, Limitations and Future Research

As shown in Table 1, the ADF test identified with 100% accuracy whether the related pupillary size time series is stationary or not.

The results can be used in NeuroIS research evaluating the stationarity of pupillary signals to exclude unidirectional luminescence changes, typically caused by sunrise, sunset or monitor brightness changes—which regularly occurred in non-laboratory environments [1, 21]. The method could be applied as a precondition of mental workload assessment excluding some biases in pupil size. However, it should be mentioned that short-time balanced pupil changes, for instance caused by a user's emotions [10] or the pupils' light reflexes [22] cannot be detected using this method. Spectral analysis based procedures such as the calculated Index of Cognitive Activity [22–24] could subsequently be applied to exclude the effect of pupils' light reflexes.

Since we only compared two extreme scenarios (either a mental workload task without a substantial luminescence change or a substantial luminescence change without a mental workload task) future work should verify the ADF results through a two times two factor design (mental workload task yes/no times substantial luminescence change yes/no).

## References

1. Buettner, R.: The relationship between visual website complexity and a user's mental workload: A NeuroIS Perspective. In *Information Systems and Neuro Science*, vol. 10 of LNISO, pp. 107–113. Gmunden, Austria (2016)
2. R. Buettner, Analyzing mental workload states on the basis of the Pupillary Hippus, in *NeuroIS'14 Proceedings*, p. 52 (2014)
3. Buettner, R., Sauer, S., Maier, C., Eckhardt, A.: Towards ex ante Prediction of user performance: A novel NeuroIS methodology based on real-time measurement of mental effort, in *HICSS-48 Proceedings*, pp. 533–542 (2015)
4. Buettner, R.: Cognitive workload of humans using artificial intelligence systems: Towards objective measurement applying eye-tracking technology, in *KI 2013 Proceedings*, ser. LNAI, vol. 8077, pp. 37–48 (2013)
5. Buettner, R., Daxenberger, B., Eckhardt, A., Maier, C.: Cognitive workload induced by information systems: Introducing an objective way of measuring based on pupillary diameter responses, in *Pre-ICIS HCI/MIS 2013 Proceedings*, 2013, p 20 (2013)
6. Dimoka, A., Pavlou, P.A., Davis, F.D.: NeuroIS: The potential of cognitive neuroscience for information systems research. *ISR* 22(4), 687–702 (2011)
7. Dimoka, A., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P.H., Pavlou, P.A., Müller-Putz, G., Riedl, R., vom Brocke, J., Weber, B.: On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS. *MISQ* 36(3), 679–A19 (2012)
8. Riedl, R., Banker, R.D., Benbasat, I., Davis, F.D., Dennis, A.R., Dimoka, A., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Müller-Putz, G., Pavlou, P.A., Straub, D.W., vom Brocke, J., Weber, B.: On the foundations of NeuroIS: Reflections on the Gmunden Retreat 2009. *CAIS* 27, 243–264 (2010)

9. Steinhauer, S.R., Siegle, G.J., Condray, R., Pless, M.: Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *Int. J. Psychophysiol.* **52**(1), 77–86 (2004)
10. Bradshaw, J.L.: Pupil size as a measure of arousal during information processing. *Nature* **216** (5114), 515–516 (1967)
11. Knapen, T., de Gee, J.W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., Theeuwes, J.: Cognitive and ocular factors jointly determine pupil responses under equiluminance. *PLoS ONE* **11**(5), e0155574 (2016)
12. Czajka, A.: Pupil dynamics for iris liveness detection. *IEEE Trans. Inf. Forensics Secur.* **10** (4), 726–735 (2015)
13. vom Brocke, J., Liang, T.-P.: Guidelines for neuroscience studies in information systems research. *JMIS* **30**(4), 211–234 (2014)
14. vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A.: Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *CAIS* **37**(1) article 9 (2015)
15. Goldinger, S.D., Papesh, M.H.: Pupil dilation reflects the creation and retrieval of memories. *Curr. Dir. Psychol. Sci.* **21**(2), 90–95 (2012)
16. Laeng, B., Sirois, S., Gredebäck, G.: Pupillometry: A window to the preconscious? *Perspect Psychol Sci* **7**(1), 18–27 (2012)
17. Duchowski, A.T.: *Eye tracking methodology: Theory and practice*, 2nd edn. Springer, London (2007)
18. Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* **91**(2), 276–292 (1982)
19. Hess, E.H., Polt, J.M.: Pupil size in relation to mental activity during simple problem-solving. *Science* **143**(3611), 1190–1192 (1964)
20. R. Core Team. *A language and environment for statistical computing*. Austria: R Foundation for statistical computing Vienna (2017)
21. Buettner, R.: A user’s cognitive workload perspective in negotiation support systems: An eye-tracking experiment. In *PACIS 2016 Proceedings* (2016)
22. Marshall, S.P.: The index of cognitive activity: measuring cognitive workload, In: *Proceedings of the 7th IEEE Conference on Human Factors and Power Plants*, Scottsdale Arizona, pp. 7-5–7-9. 15–19 Sept (2002)
23. Marshall, S.P.: Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity, U.S. Patent 6,090,051 (2000)
24. Marshall, S.P.: Identifying cognitive state from eye metrics. *Aviat. Space Environ. Med.* **78** (Supplement 1), B165–B175 (2007)

# Towards Reconceptualizing the Core of the IS Field from a Neurobiological Perspective

Lars Taxén

**Abstract** The IS discipline has so far been unable to define the meaning of its foundational concepts “information” and “system”. As a consequence, the core of the IS field and its relation to the IT artifact, materiality, organization, and so on, are extensively discussed without reaching closure. To this end, this paper proposes a set of conceptual stepping stones towards reconsidering the core of the IS field from a neurobiological perspective. The analysis suggests that *information* can be defined as intrinsically related to individual, neural abilities for acting; and *system* as a dialectical relation between the individual and the IT artifact. As a consequence, *information system* is seen as having both an individual and social facet. These results indicate that a neurobiological perspective may open up new avenues for revitalizing the IS field.

**Keywords** Information · System · Activity modalities · Functional systems  
Neurobiology · Anokhin · Vygotsky · Luria

## 1 Introduction

According to Beath et al. [1, p. v] IS (Information System) research needs to attend simultaneously “the technical and the human (social) side of IT in its organizational context ... and it is precisely this combination that gives IS research its distinctive value”. However, research initiatives are aggravated by the failure to establish a well-defined core of the IS field. Bedrock IS concepts such as “information” and “system” remain undefined: “Virtually all the extant IS literature fails to explicitly specify meaning for the very label that identifies it. This is a vital omission, because without defining what we are talking about, we can hardly know it” [2, p. 338].

---

L. Taxén (✉)

Department of Computer and Information Science, Linköping University, Linköping, Sweden

e-mail: lars.taxen@gmail.com

Further, little progress has been made since the 1990s in conceptualization of the central entity of the field—the IT artifact [3, 4].

This state of play provides the motivation for the paper, which is to reconsider the foundation of the IS field from *neurobiological* perspective. This is in line with the ambitions of the NeuroIS initiative to build “superior IS theories with assumptions and constructs that better correspond to the brain’s functionality” [5, p. 2]. However, NeuroIS contributions have so far “seldom applied specific neuroscience theories in concrete IS research studies” [6, p. 83].

As a first step towards reconsidering the core of the IS field, alternative definitions of the key IS concepts “information” and “system” are suggested. The line of argument proceeds as follows. First, two fundamental assumptions for a neurobiological perspective of the IS core are proposed. From these assumptions, a set of conceptual stepping stones are devised: requisite neurobiological predispositions enabling action, the dynamics of action, the structure of mental functions, the social formation of the brain, and the inclusion of brain components in mental functions. Together, these stepping stones suggest that *information* can be defined as intrinsically related to individual, neural abilities for acting. Likewise, *system* is seen as a dialectical relation between the individual and the IT artifact. As a consequence, the *Information System* needs to be reconceptualized as having both an idiosyncratic, individual facet and a communal, social facet.

In this way, the neurobiological perspective enables completely new conceptualizations of core IS constructs, which is the knowledge contribution of the paper. Hence, a first stepping stone towards building an alternative foundation for the IS field is achieved. In conclusion, such findings, inchoate as they may be, are promising enough for launching a more extensive research initiative, aimed at revitalizing the IS field from a neurobiological perspective.

## 2 Fundamental Neurobiological Assumptions

Any research program needs to proceed from some fundamental, “hardcore assumptions”, which are not questioned as long as the program progresses [7]. A first assumption from a neurobiological point of view is that *brains evolved to control the activities of bodies in the world*. The “mental is inextricably interwoven with body, world and action: the mind consists of structures that operate on the world via their role in determining action” [8, p. 527]. A second assumption is that *individuals cannot be understood* without taking their social environment into account. The opposite is also true: *the social environment cannot be understood* without understanding the individual [9]. Accordingly, the neural and social realms form a unity, which parts “cannot be separated or isolated without destroying the phenomenon that is studied” [9, pp. 336–337].

### 3 Conceptual Stepping Stones

From the fundamental assumptions, the following conceptual stepping stones are suggested towards reconsidering the IS core.

#### 3.1 *Neurobiological Predispositions Enabling Action*

The purpose of this stepping stone is to identify phylogenetically evolved, neurobiological predispositions for action, which “are universal and inherent for all humans, independent of language and environmental conditions” [10, p. 43]. Metaphorically, such predispositions can be seen as a neurobiological “infrastructure” that the individual is endowed with at birth. While providing a full account of such predispositions is indeed a prodigious task, it is nevertheless possible to consider *requisite* predispositions, that is, necessary albeit not sufficient ones. One proposal for such predispositions is as follows [11, 12]:

- *Objectivating*: attending something towards which actions are directed
- *Contextualizing*: foregrounding that which is relevant for acting
- *Spatializing*: orienting in the environment
- *Temporalizing*: anticipating and carrying out actions
- *Stabilizing*: habitualizing appropriate actions
- *Transitional*: refocusing attention.

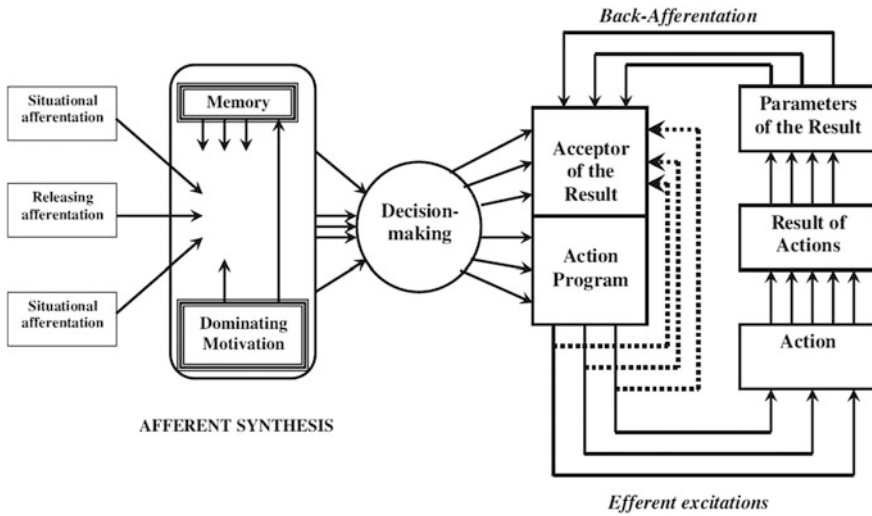
These predispositions are termed *activity modalities*, and were devised from long-term observations and reflections in practice [13]. Importantly, all modalities are needed. A brain lesion destroying any modality will inevitably obstruct the individual from acting.

#### 3.2 *The Dynamics of Action*

The focus of this stepping stone is the dynamics of an action. To this end, Anokhin [14] has proposed the model in Fig. 1.

The various stages in this model involve two kinds of functions depending on which kind of nerves are actuated [14]: afferent (going from the periphery of the body to the brain), and efferent (going from the brain to effectors such as muscles or glands). The stage “afferent synthesis” performs “space-time integration on the multisensory percept, a Gestalt” [15]. Based on this Gestalt, a decision is taken of “*what to do, how to do, and when to do*” [14, p. 114, italics in original]. “Decision making” involves two functions—characterization of the expected result (“acceptor of the result”), and formation of an “action program”. Functions in “efferent





**Fig. 1** The dynamics of action (reproduced from [14, p. 115]; with permission)

excitation” enable action, after which the result modifies and stores the “acceptor of the result” via “back-afferentation”.

### 3.3 The Structure of Mental Functions

The purpose of this stepping stone is to model the structure of mental functions. According to Luria [16], such functions are *complex functional systems*. No specific function is ever connected with the activity of one single brain center: “It is always the product of the integral activity of strictly differentiated, hierarchically interconnected centers” [17]. In Fig. 2, a model for a functional system enabling action is illustrated.

This model shows dependencies between individual functions (which in turn may be functional systems) contributing to the overall functional system; from basic ones and progressing upwards. As such, the model illustrates how the neural system “comes alive” after being shut down, for example, during sleep. In a metaphoric sense, this can be likened with starting up a car to its idling state; thus preparing it for subsequent action. The gist of the model in Fig. 2 is to show how the functional system as a whole is impacted if a particular function is inhibited by a brain damage in its contributing components. The components realizing the functions are subdued in order to focus on simplicity and critical functional dependencies.

As can be seen, the two models in Figs. 1 and 2 are related. The stages in the Anokhin model correspond to functional groups in the structural model. Together, these models capture the architecture of the individual brain at a level suitable for

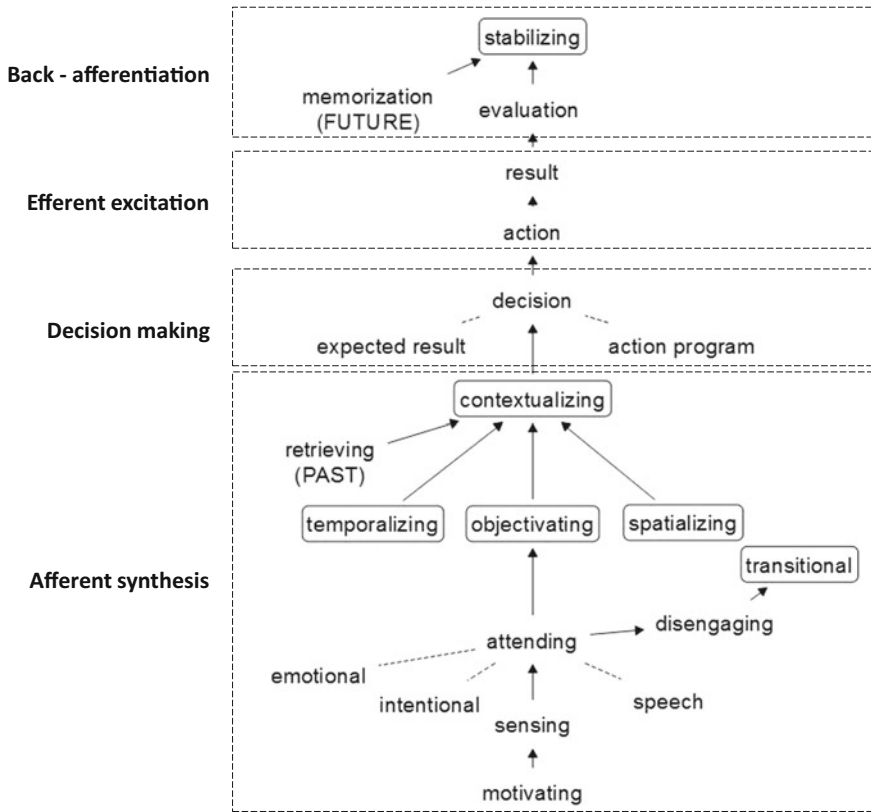


Fig. 2 A functional system enabling action (the activity modalities are emphasized)

further inquiries into the relation between the neural and social realms. This is done in the subsequent section.

### 3.4 The Social Formation of the Brain

The purpose of this stepping stone is to conceptualize how the individual and the environment mutually constitute each other. The predispositions in the neurobiological “infrastructure” will develop into neurobiological *abilities* in interaction with the cultural and historical environment the individual is immersed in. Importantly, neural predispositions need to be distinguished from neural abilities. For example, many contemporaries of Julius Caesar certainly had predispositions to become pianists, but were never able to develop the corresponding ability because the pianoforte had not yet been invented [14]. As a consequence, “external aids or

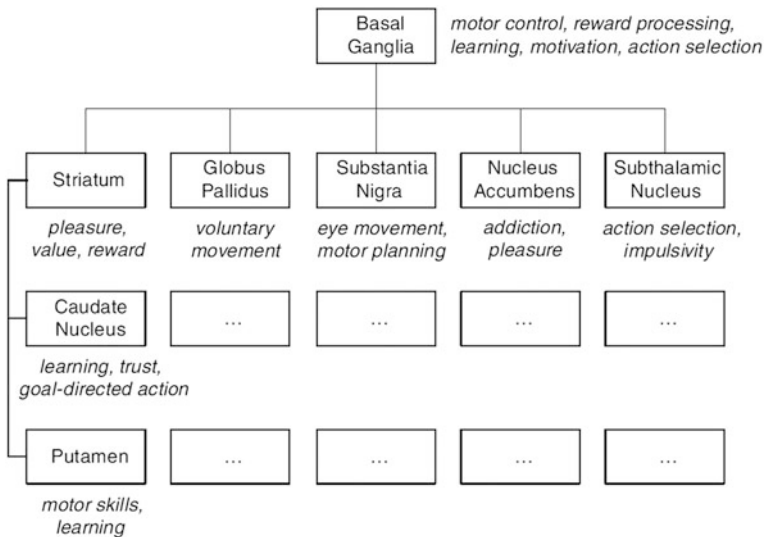
historically formed devices are *essential elements in the establishment of functional connections between individual parts of the brain*” [16, p. 31; italics in original].

The essence of this insight is that the development of the neural system necessitates two pre-existing “infrastructures”—a neural one and a social one. The neural one provides opportunities for the individual to develop neural abilities, but these opportunities are constrained and enabled by the social one. Every action changes both infrastructures, albeit on different timescales from millisecond (neural), cultural-historical (social), and evolutionary (neural predispositions).

### 3.5 On Inclusion of Brain Components in Mental Functional Systems

The purpose of this stepping stone is to indicate how functions of individual brain components relate to functional systems as conceived by Luria, Vygotsky, and other scholars [9, 14, 16]. In order to illustrate this we may consider the basal ganglia and its sub-components [6, pp. 86–87] (see Fig. 3).

For example, the function “goal-directed action” is associated in Fig. 3 with the caudate nucleus sub-component. However, the caudate nucleus is not realizing this function on its own. Rather, “goal-directed action” needs to be considered as a *functional system* [16], possibly structured as in Fig. 2. From this model, we may conclude that several sub-components of the basal ganglia, besides the caudate



**Fig. 3** The basal ganglia, major sub-components, and important functions (from [6, p. 86]; reproduced with permission)

nucleus, are involved in “goal-directed action”; such as the subthalamic nucleus (action selection); the substantia nigra (motor planning); the globus pallidus (movement); and the putamen (motor skills, learning). We can also see that the same component may contribute to several functions in the functional system, for example the putamen. As a consequence, explanative theoretical knowledge about functions of individual components needs to be complemented by functional system models in order to fully specify functions of neural components.

## 4 IS Implications

### 4.1 Information is *Intrinsically Related to the Individual Abilities for Acting*

According to Boland [18], the essence of information is revealed in its name: “Information is an inward-forming”. This view complies well with the neurobiological approach. In the stage “afferent synthesis” (see Figs. 1 and 2), sensations emanating from the environment are integrated into a multisensory percept as a prerequisite for acting. This integration takes place *entirely in the brain*. The result is *informative* for the individual. Since the activity modalities are posited as requisite for action, this means that *information* may be conceptualized as the totality of objectivating, contextualizing, spatializing, temporalizing, stabilizing, and transitional information. Consequently, action is alleviated if the environment is congruent with these modalities, which implies that we strive to construct our environment accordingly. So, for example, we have maps alleviating spatialization, clock alleviating temporalization, and so on.

### 4.2 System is *Comprised of Individual Neural Abilities and the IT Artifact*

Paul [19, p. 379] has suggested defining Information Systems as “Information Technology in Use”. Since we have posited that information is inherently individual, Paul’s definition indicates that the *system* can be seen as the entity made up by the individual user in interaction with the IT artifact. The development of such a system is manifested as neurobiological abilities in the individual, and most certainly as an adaption of the IT artifact to suit the needs of the social context such as an organization. In principle, this means that we need to conceptualize *Information System* as having both an idiosyncratic, individual facet, and a communal, social facet.

This conceptualization of the IS makes it possible to address several outstanding IS issues from a new vantage point. For example, the IT artifact is seen as a regular

physical artifact based on technology, which means that we “do not need to put humans inside the boundary of the IT artifact in order to make these artifacts social” [20, p. 94]. The specificity of the IT artifact lies in its designation to support the integration of information and subsequent action in *all dimensions given by the activity modalities*. Further, the definition of IS as a dialectical unity of the individual and IT artifact enables a clear ontological separation of them, while still maintaining their inescapable, mutual constitution. This is in stark contrast to the ontological foundation of the prevalent IS research stream of sociomateriality, which claims that any “distinction of humans and technologies is analytical only” [21, p. 456].

## 5 Concluding Discussion

IS research progressing from the current foundation of the IS field is unable to address long-standing, die-hard issues, such as the nature of the IT artifact and the IS [4], the question about “materiality” [22], and the status of the IS discipline [1]. This paper proposes to investigate such issues from a new foundation for the IS field, based on a neurobiological perspective. Needless to say, results achieved so far are in a nascent state; a kind of *prescience* “discerning or anticipating what we need to know and, equally important, of influencing the intellectual framing and dialogue about what we need to know” [22, p. 13]. To advance this state, a comprehensive research initiative is called for. As an established IS sub-discipline, the NeuroIS initiative is in a unique position to pilot such an initiative, thus opening up qualitatively new avenues for IS research. With the availability of committed NeuroIS scholars and access to NeuroIS methods and tools [6], the time is ripe to engage this stock of knowledge in an urgently needed revitalization of the IS field.

## References

1. Beath, C., Berente, N., Gallivan, M.J., Lyytinen, K.: Expanding the frontiers of information systems research: introduction to the special issue. *J. Assoc. Inf. Syst.* **14**(4), Article 4 (2013). <http://aisel.aisnet.org/jais/vol14/iss4/4>
2. Lee, A.S.: Retrospect and prospect: information systems research in the last and next 25 years. *J. Inf. Technol.* **25**(4), 336–348 (2010)
3. Akhlaghpour, S., Wu, J., Lapointe, L., Pinsonneault, A.: The ongoing quest for the IT artifact: looking back, moving forward. *J. Inf. Technol.* **28**(2), 150–166 (2013)
4. Alter, S.: The concept of “IT artifact” has outlived its usefulness and should be retired now. *Inf. Syst. J.* **25**(1), 47–60 (2015)
5. Dimoka, A., Pavlou, P.A., Davis, F.: Neuro-IS: the potential of cognitive neuroscience for information systems research. In: *Proceedings of the 28th International Conference on Information Systems*. Montreal, Canada, 11–14 Dec (2007)
6. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS—Information Systems and the Brain*. Springer, Berlin (2016)

7. Lakatos, I.: Falsification and the methodology of scientific research programs. In: Lakatos, I., Musgrave, A. (eds.) *Falsification and the Methodology of Scientific Research Programs*, pp. 91–132. Cambridge University Press, New York (1970)
8. Love, N.: Cognition and the language myth. *Lang. Sci.* **26**(6), 525–544 (2004)
9. Toomela, A.: There can be no cultural-historical psychology without neuropsychology. And vice versa. In: Yasnitsky, A., van der Veer, R., Ferrari, M. (eds.) *The Cambridge Handbook of Cultural-Historical Psychology*, pp. 315–349. Cambridge University Press, Cambridge (2014)
10. Kotik-Friedgut, B.: Development of the Lurian approach: a cultural neurolinguistic perspective. *Neuropsychol. Rev.* **16**(1), 43–52 (2006)
11. Taxén, L.: The activity modalities: a priori categories of coordination. In: Liljenström, H. (ed.) *Advances in Cognitive Neurodynamics (IV)*, pp. 21–29. Springer Science + Business, Dordrecht (2015)
12. Taxén, L.: An investigation of the nature of information systems from a neurobiological perspective. In: Davis, F., Riedl, R., vom Brocke, J., Léger, P.-M., Randolph, A. (eds.) *Information Systems and Neuroscience, Gmunden Retreat on NeuroIS 2015*, pp. 27–34. Springer Science + Business Media, Dordrecht (2015)
13. Taxén, L.: *Using Activity Domain Theory for Managing Complex Systems*. Information Science Reference. Information Science Reference (IGI Global), Hershey, PA (2009)
14. Toomela, A.: Biological roots of foresight and mental time travel. *Integr. Psychol. Behav. Sci.* **44**, 97–125 (2010)
15. Freeman, W.: How and why brains create meaning from sensory information. *Int. J. Bifurc. Chaos* **14**(2), 515–530 (2004)
16. Luria, A.R.: *The Working Brain*. Penguin Books, London (1973)
17. Akhutina, T.V., Vygotsky, L.S., Luria, A.R.: Foundations of neuropsychology. *J. Rus. East Eur. Psychol.* **41**(3/4), 159–190 (2003)
18. Boland, R.J.: The in-formation of information systems. In: Boland, R.J., Hirschheim, R.A. (eds.) *Critical Issues in Information Systems Research*, pp. 363–379. Wiley, New York (1987)
19. Paul, R.: Loose change. *Eur. J. Inf. Syst.* **19**, 379–381 (2010)
20. Goldkuhl, G.: The IT artefact: an ensemble of the social and the technical?—a rejoinder. *Syst. Sign. Actions* **7**(1), 90–99 (2013)
21. Orlikowski, W.J., Scott, S.V.: Sociomateriality: challenging the separation of technology, work and organization. *The Acad. Manag. Ann.* **2**(1), 433–474 (2008)
22. Hassan, N.R.: Editorial: a brief history of the material in sociomateriality. *ACM SIGMIS Database: The DATABASE Adv. Inf. Syst.* **47**(4), 10–22 (2016)

# Using EEG Signal to Analyze IS Decision Making Cognitive Processes

Nabila Salma, Bin Mai, Kamesh Namuduri, Rasel Mamun,  
Yassir Hashem, Hassan Takabi, Natalie Parde and Rodney Nielsen

**Abstract** In this paper, we demonstrate how electroencephalograph (EEG) signals can be used to analyze people's mental states while engaging in cognitive processes during IS decision-making. We design an experiment in which participants are required to complete several cognitive tasks with various cognitive demands and under various stress levels. We collect their EEG signals as they perform the tasks and analyze those signals to infer their mental state (e.g., relaxation level and stress level) based on their EEG signal power.

**Keywords** EEG · Decision making · Signal processing · Cognitive process

---

N. Salma (✉) · B. Mai · K. Namuduri · R. Mamun · Y. Hashem · H. Takabi · N. Parde ·  
R. Nielsen  
University of North Texas, Denton, USA  
e-mail: Nabila.Salma@unt.edu

B. Mai  
e-mail: Bin.Mai@unt.edu

K. Namuduri  
e-mail: Kamesh.Namuduri@unt.edu

R. Mamun  
e-mail: MdRasel.Mamun@unt.edu

Y. Hashem  
e-mail: Yassir.Hashem@unt.edu

H. Takabi  
e-mail: Hassan.Takabi@unt.edu

N. Parde  
e-mail: Natalie.Parde@unt.edu

R. Nielsen  
e-mail: Rodney.Nielsen@unt.edu

## 1 Introduction

In this paper, we investigate the properties of EEG signals when people engage in cognitive processes during Information Systems (IS) decision-making. Decision-making is fundamental to most human behaviors [1], and can be classified into four categories: intuitive, empirical, heuristic and rational. Among them, rational decision-making is more easily defined and explained with cognitive psychology, and will be the focus of this paper. Rational decision-making is a method where the brain develops a criterion of functions representing potential choices and processing available information to find the good choice among others [1]. The two subcategories for rational decisions are static and dynamic. Static decisions are made based on statistically viable information such as loss and gain, cost–benefit, practicality and functionality. Dynamic decisions are based on alternatives, present situation and past knowledge of similar situations. In this paper, we focus on the EEG signals analysis of rational decision-making.

This paper discusses how rational decision-making can be analyzed with help from EEG (electroencephalogram) signal power variance generated by humans who are making those decisions. EEG has been traditionally used as a diagnostic application for diseases such as Epilepsy [2], and more recently has become a popular tool for NeuroIS studies (e.g., [3, 4]), and decision-making research (e.g., [5, 6]). In this paper, we measure and analyze the EEG signals from decision makers in an experimental setting, and investigate what the EEG signals can tell us about those decision-makers' behaviors.

The human brain releases EEG signals in various frequency bands, usually categorized into five bands: Alpha (8–13 Hz), Beta (14–31 Hz), Delta (4 Hz or less), Gamma (greater than 32 Hz) and Theta (4–7.5 Hz). Among them, Alpha, Beta, Delta and Theta are most widely used for EEG signal analysis, especially for various cognitive functions. These will also be the EEG signals we focus on in this paper.

Neuroscience literature has established the various and distinct roles for each of these EEG signal bands in human cognitive functions [7]. It has been shown in studies that in subjects who are awake, Delta waves can relate to cognitive concentration. Several experiments have demonstrated that there is a clear relation between cognitive concentration and increased activity in the Delta frequency band [8]. Theta is an indicator of stress. The study presented in [9] shows that EEG Theta/Beta ratio as a potential biomarker for effects of stress on attention. The study confirmed a negative relation between Theta/Beta ratio and stress-induced attentional control [9]. Statistical analyses in literature have also shown a positive relationship between stress and theta power spectrum density value [10]. Beta waves are associated with cognitive processing. Activity in this frequency band will increase when there is cognitive challenge and increased demand for a cognitive task [11]. According to [12], increasing Beta activity has been identified with high concentration and attention. Alpha waves are well known for their correlation with a relaxed state. During a resting period, the Alpha frequency band is seen to have



maximum magnitude. The magnitude of Alpha waves is higher when eyes are closed compared to when eyes are open [11]. According to [13] decreasing Alpha activity is consistent with higher cognitive demand in decision making. In addition, cognitive activity typically suppresses alpha and elevates beta activity [14], and frontal theta signal may serve as an index for mental effort [14, 15].

We conduct an experiment in which participants are asked to perform tasks of various levels of cognitive processing (data entry vs. application programming) under various levels of stress (no time limit for the task vs. with time limit). Through the analysis of the processed EEG signals from the participants, we replicate the results that Alpha band signal power is higher when the task requires lower cognitive demands, and Theta band signal power is higher when the task involves higher stress. Surprisingly, our data also indicate that when performing a high-cognitive-demand task, the participants' Alpha signal power is higher when the stress level is higher. We look into the experiment design and propose some possible explanations for this surprising observation.

The rest of the paper is organized as follows: in Sect. 2, we discuss the experiment design and signal processing techniques, followed by data analyses and discussion in Sect. 3. We conclude in Sect. 4.

## 2 Experiment

### 2.1 Method

We recruited 25 participants to participate in our experiments; 15 were male and the other 10 were female. All participants were between the ages of 18 and 34 years old and were graduate or undergraduate students in the Department of Computer Science and Engineering in University of North Texas. The EEG recordings of 5 of the participants are incomplete and for this research we used the EEG signals from the other 20 participants. Participants were asked to perform 6 tasks in total. EEG signals generated from four of the tasks are used in this study (the other two tasks are not relevant to our research questions). For Task 1 participants were asked to perform data entry (login to a database system and update the student records using the information provided). Task 2 was a similar data entry task (update the same student information but with twice the number of student records) but with a time limit. Task 3 was to perform a computer programming exercise (complete a coding project for designing a calculator, using a language they felt most comfortable with) and Task 4 was a similar programming exercise (complete the same calculator application but using a different language) with a time limit. All four tasks are typical representatives for IS activities that require rational cognitive decision making in completing those tasks. The experiments were conducted in a dedicated EEG laboratory, and the room was set up to keep the same environmental conditions for all tasks and all participants. The experiments were conducted for each

participant separately and at different times during the day. The participants were seated in a comfortable chair. After the relevant areas on the face and mastoids were cleaned, the Geodesic Sensor Net (GSN) was positioned on the participant's head. The examiner checked for signal impedances, applying additional saline solution and readjusting sensors as needed to ensure minimal impedance and optimal signal quality between each electrode and the participant's scalp. The examiner then explains the task and what the participant had to do step-by-step using a predefined script located on the computer desktop. The participants were given five minutes to read the script before each task and to feel comfortable with the test environment.

The second and fourth tasks were conducted at the end of the work day, so that the participants would have attended classes, exams or labs during the day prior to coming to participate in the experiment. And the experiment was time constrained to induce further stress among the participants.

To measure the participants' brain activities, we used EGI's Geodesic EEG System 400 [16], with a 256-channel HydroCel GSN. We used a sampling rate of 1000 Hz. The device has been widely adopted by the clinical and research community because of its ease of use, comfort, and ability to produce high-quality and high-resolution data.

## 2.2 EEG Recording

EEG recordings from all sensors were used for analysis. Signal analysis was performed in LabVIEW. Recordings were analyzed in 100 s segments. Recordings were processed to remove artifacts from muscle movements such as eye blinks. A fast Fourier transform using hamming window with 50% overlap was used. A digitized version of an analog signal is an approximation of the analog signal. This signal analysis platform designed in LabVIEW decomposes the signal into the approximated frequency component of the original signal. However, EEG is not a stationary signal. During analysis the components change in frequency and amplitude at every window as transient waveforms appear intermittently. Choosing short window duration minimizes the effect of being non-stationary and generates a smoother PSD plot [16]. In this case a window length of 1024 was used. The overlap in this design is set to 50% which is half the window length. This means each sample will make equal impact on the spectrum. The design was verified with simulated EEG to confirm that design input meets design output. For verification testing, 100 s of simulated EEG recording was used at 1000 Hz sampling rate. For experimental recordings, signal power in frequency band activity of Delta (0.5–4 Hz), Theta (4–7.5 Hz), Alpha (8–13 Hz) and Beta (14–26 Hz) were calculated. Mean signal power of each frequency band for each recording (for each task, for each participant) was computed. Ratios of these mean signal power values across tasks were used for data analysis to draw conclusions.

## 2.3 *Workflow of the Design*

Each EEG recording is uploaded into Read Biosignal VI in LabVIEW. The entire design is placed inside a single while loop. This tool reads bio-signals block by block. The block sizes are in seconds and they can vary depending on the length of the EEG recording. In this case, the block size is set to 100 s. The loop stop condition is wired with the End of File (EOF) terminal of the block. The loop stops when it reaches the end of the uploaded EEG recording. The signal powers and percentages are calculated for each loop and saved in the respective arrays for Alpha, Beta, Theta and Delta. Each additional loop adds a new calculated value to the array for the subsequent 100 s of recording until the end of the recording. The mean values are calculated as the loops are iterated and the final mean values reflect results for the entire recording. EEG FFT Spectrum VI is used to separate the frequency bands (Alpha, Beta, Delta and Theta) from the raw signal. This VI computes the single-sided power spectral density (PSD). The time series is then divided into overlapping subcategories of signal elements. Periodograms of these subcategories are then averaged to plot the PSD. For this design the VI returns the PSD values in linear scale. Frequency bands for EEG are defined in the VI to be extracted accordingly. An unbundling function is used to extract the elements. It obtains the FFT spectrum as a cluster and creates terminals for respective frequency bands for the measured value to be used independently. The signal power value returned is the absolute value of power in each frequency band. The percentage of each frequency band shows the distribution of power in respective frequency bands. Signal power and percentage values for each iteration are saved in an  $n$ -dimensional array. Each time the while loop runs and a new value results from EEG FFT Spectrum VI, this function enters the value into its respective array. The feedback node attached to its output to input stores data from one iteration to another. Therefore, at the end of the final loop the array contains values from all iterations. This array is an input to the Mean VI which then takes the values from all iterations in consideration in order to compute the final mean value. The signal power mean values and power distribution percentages are then recorded in a data sheet for all the frequency bands for further analysis.

## 3 Results and Analyses

### 3.1 *The Impact of Cognitive Demand and Stress Level*

To investigate the impact of the tasks' cognitive demand on brain signal power, we calculate the Alpha/Beta ratio generated from the four tasks. We then perform a paired  $t$ -test comparing the ratio for the data entry task vis-à-vis the ratio for the application programming task. When the tasks are performed without a time limit, the Alpha/Beta ratio is shown to be significantly higher for the data entry task than

**Table 1** The paired *t*-test result for Alpha/Beta ratio between data entry and programming tasks without time constraints and with time constraint

Pair 1a	Mean	Std. Dev.	<i>t</i>	df	Sig.
A/B: data entry—programming (without time limit)	0.24520	0.41985	2.612	19	0.017
Pair 1b	Mean	Std. Dev.	<i>t</i>	df	Sig.
A/B: data entry—programming (with time limit)	-0.15576	0.63030	-1.119	19	0.277

for the application programming task (see Pair 1a in Table 1). When the tasks are performed with time limit, the ratio difference is not significant between the two tasks (see Pair 1b in Table 1).

These results are aligned with the literature that Alpha signals are positively related to relaxation. While engaging in a more cognitive demanding task, people tend to be less relaxed, thus generating lower levels of Alpha signals. The lack of statistical significance in the results from the time-constrained tasks seems to suggest that the relaxation state is quite vulnerable to stress level.

To investigate the impact of stress level on the brain signal power, we calculate the Theta/Beta ratio generated from the four tasks. We then perform a paired *t*-test between the ratio when performing low stress tasks (in this case, the tasks with no time constraint) vis-à-vis the ratio when performing high stress tasks (in this case, the tasks with time limit). When the participants perform the application programming task, their Theta/Beta ratio is shown to be significantly higher under time constraint compared to without time constraint (see Pair 2a in Table 2). When they perform the data entry task, the ratio difference is not significant (see Pair 2b in Table 2).

These results are aligned with the literature that Theta signals are positively related to stress level. While engaging in cognitive tasks with time constraints when the participants are mentally/physically tired, they tend to experience higher stress levels compared to engaging in tasks with no time constraints and when they are relatively fresh, thus the participants generate higher levels of Theta signals. The lack of statistical significance in the results from the data entry tasks seems to

**Table 2** The paired *t*-test result for Theta/Beta ratio between tasks without time constraint versus with time constraint

Pair 2a	Mean	Std. Dev.	<i>t</i>	df	Sig.
T/B: no time limit—time limit (programming task)	-1.49532	2.63847	-2.535	19	0.020
Pair 2b	Mean	Std. Dev.	<i>t</i>	df	Sig.
T/B: no time limit—time limit (data entry task)	0.36237	1.98533	0.816	19	0.424

**Table 3** The paired *t*-test result for Alpha/Beta ratio between programming tasks without time constraint versus with time constraint

Pair 3	Mean	Std. Dev.	<i>t</i>	df	Sig.
A/B: no time limit—time limit (programming task)	-0.33655	0.64671	-2.327	19	0.031

suggest that the low cognitive demand of the task may override the impacts from the stress level induced by the time constraints.

One surprising result we obtain while comparing the Alpha/Beta ratio is that when performing the programming task, the participants show significantly higher Alpha/Beta ratio when there is a time constraint versus when there is no time constraint (see Table 3).

This seems to suggest that for the programming task, the participants are more relaxed (higher Alpha/Beta ratio) when there is a time constraint than when there is no time constraint. One possible explanation to this counter-intuitive result is that in our experiment, all participants are computer science students, who may be well versed in creating applications in various programming languages. Thus the required task (creating a calculator application) is an easy task for the participants. Therefore, the time constraint did not impede their relaxation level. In addition, in our experiment design, their time-constrained task is after their no-time-constrained task, and they are the same task except that they need to use another programming language in the time-constrained task, thus they are already familiar with the task requirements, and as a result, they show a higher relaxation level for the second implementation (the time-constrained task), perhaps their familiarity with the task overrides the impact of their lower familiarity with the programming language.

## 4 Conclusion

In this paper, we demonstrate how EEG signals can be used to analyze people's mental states while engaging in cognitive processes during decision making. We design an experiment in which participants are required to complete several cognitive tasks with various cognitive demands and under various stress levels. We collect their EEG signals during their task performance and analyze the signal to infer their mental state such as relaxation level and stress level based on their EEG signal power. We find that when people engage in decision-making cognitive process, higher cognitive demand from the decision-making processes results in lower Alpha/Beta signal ratio, which indicates a lower level of relaxation; and higher stress level usually results in a higher Theta/Beta ratio. For future work, we plan to refine our experiment and conduct cross-factor analyses of the impacts of various factors that influence brain signals during decision-making cognitive processes.

## References

1. Wang, Y., Ruhe, G.: The cognitive process of decision making. *Int. J. Cog. Inf. Nat. Intell.* 73–85 (2007)
2. Smith, S.: EEG in the diagnosis, classification, and management of patients with epilepsy. *J. Neurol. Neurosurg. Psychiatry*, **76**, ii2–ii7 (2005)
3. Riedl, R., Léger, P.-M.: *Fundamentals of NeuroIS—Information Systems and the Brain*. Springer, Berlin (2016)
4. Dimoka, A., Pavlou, P., Davis, F.: NeuroIS: The potential of cognitive neuroscience for information systems research. *Inf. Syst. Res.* **22**(4) (2011)
5. Larsen, T., O’Doherty, J.P.: Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI–EEG study. *Philos. Trans. R. Soc. B Biol. Sci.* **369**(1655) (2014)
6. Douglas, P.K., Lau, E., Anderson, A., Head, A., Kerr, W., Wollner, M., Moyer, D., Li, W., Durnhofer, M., Bramen, J., Cohen, M.S.: Single trial decoding of belief decision making from EEG and fMRI data using independent components features. *Front. Hum. Neurosci.* **7**, 392 (2013)
7. Matthews, G., Reinerman-Jones, L.E., Barber, D.J., Abich IV, J.: The psychometrics of mental workload: multiple measures are sensitive but divergent. *Hum. Factors* **57**(1), 125–143 (2015)
8. Putman, P., Verkuil, B., Arias-Garcia, E., Pantazi, I., van Schie, C.: EEG theta/beta ratio as a potential biomarker for attentional control and resilience against deleterious effects of stress on attention. *Cogn. Affect. Behav. Neurosci.* **14**(2) (2014)
9. Baldauf, D., Desimone, R.: Neural mechanisms of object-based attention. *Science* **344**(6182), 424–427 (2014)
10. Awang, S.A., Pandiyan, P.M., Yaacob, S., Ali, Y.M., Ramidi, F., Mat, F.: Spectral density analysis: theta wave as mental stress indicator. In: *Communications in Computer and Information Science Signal Processing, Image Processing and Pattern Recognition*, pp. 103–112 (2011)
11. Thomas, A., Budzynski, H., Evans, J., Abarbanel, A.: Utility of segmentation of EEG frequency bands by function. In: *Introduction to Quantitative EEG and NeuroFeedback*. Elsevier Inc., pp. 89–93 (2009)
12. Alfimtsev, A.N., Basarab, M.A., Devyatkov, V.V., Levanov A.A.: A new methodology of usability testing on the base of the analysis of user’s electroencephalogram. *J. Comput. Sci. App.* **3**(5), 105–111 (2015)
13. Davis, E., Hauf, J., Wu, Q., Everhart, E.: Brain function with complex decision making using electroencephalography. *Int. J. Psychophysiol.* **79**(2), 175–183 (2010)
14. Borghini G., Vecchiato G., Toppi J., Astolfi L., Maglione A., Isabella R., Babiloni F., Assessment of mental fatigue during car driving by using high resolution EEG activity and neurophysiologic indices. In: *Proceedings of Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 6442–6445 (2012)
15. Gevins, A., Smith, M.E.: Neurophysiological measures of cognitive workload during human-computer interaction. *Theor. Issues Ergon. Sci.* **4**, 113–131 (2003)
16. Campbell, G.: EEG recording and analysis for sleep research. *Curr. Protoc. Neurosci.* 10.2 (2009)