Asking both the User's Brain and its Owner using Subjective and Objective Psychophysiological NeuroIS Instruments

Short Paper

Ricardo Buettner
Aalen University
73430 Aalen, Germany
ricardo.buettner@hs-aalen.de

Abstract

To empirically evaluate the ambitious NeuroIS visions in this paper we asked the brain and its owner, using 10 psychophysiological NeuroIS instruments and traditional subjective assessments during the execution of Microsoft Excel tasks that took place within a realistic large-scale experimental setup. To simultaneously elicit perceptive, psychophysiological and if possible objective data, we chose a multi-method research approach. By strictly following the NeuroIS guidelines we found evidence that NeuroIS measures are more objective and that a combination of various NeuroIS tools increases validity – supporting the corresponding NeuroIS claims. In addition we found that worst performers had a much greater workload during their task performance compared to top performers, which was coherently measurable by every single NeuroIS indicator – supporting the NeuroIS claim of effective triangulation.

Keywords: Human behavior, Human-computer interaction, NeuroIS, Cognitive workload, Eye-Tracking, Electrodermal activity, Electroencephalography, Facial recognition

Introduction

Increasingly human behavior scholars are equipped with more precise psychophysiological instruments such as electroencephalography (EEG) or electrodermal activity (EDA) devices (Dimoka et al. 2012), and some argue for "*directly asking the brain, not the person.*" (Dimoka et al. 2011, p. 687). This call to directly ask the brain comes with visionary claims from the NeuroIS community of offering a form of "*objective and unbiased measurement*" (Dimoka et al. 2011, p. 688), "*improving the validity of research findings ... by combining two (or more) [Neuro IS] methods*" (Riedl et al. 2010, p. 248), and "*helping to build IS theories*" (Dimoka et al. 2011, p. 688).

While touching on the philosophical mind-body problem we empirically investigate from a methodological perspective whether psychophysiological NeuroIS instruments are potentially more objective and more useful for information systems scholars than traditional subjective assessments.

The mind-body problem refers to open philosophical questions concerning the existence of a human mind and if so, the (causally) interaction of the mind and its physical body (Kim 2000). While the mind is concerned with mental processes, thoughts and consciousness, the body is concerned with the physical aspects of the brain and how the brain is structured. One major theoretical approach refers to the idea that the mind does not exist distinct from the brain (monism); the opposite approach comprises the idea that mind and body are different substances (substance dualism). While there is no final answer from (philosophical) science, more and more results from neuroscience support the monistic approach (i.e., physicalism). If physicalism is true, human perceptions – which IS scholars capture using self-rating instruments – simply reflect brain activity, which could be directly analyzed using psychophysiological

NeuroIS instruments. If substance dualism is true, it is interesting to see whether NeuroIS instruments can add complementary or contrary insights to subjective self-ratings.

That is why we asked the brain and its owner, using 10 NeuroIS instruments and established subjective assessments during the execution of Microsoft Excel tasks that took place within a realistic large-scale experimental setup. We simultaneously elicit perceptive, psychophysiological and if possible objective data in a multi-method research approach.

Our main construct for the investigation is Cognitive Workload since several psychophysiological instruments are available to study this concept. In a laboratory experiment we systematically compare the psychophysiological NeuroIS indicators of Cognitive Workload with the subjective self-ratings. It is not possible to assess Cognitive Workload directly using a pure objective instrument since this concept reflects the mind part of the mind-body problem, but we know from prior literature from various disciplines using survey and objective data, that Cognitive Workload is coherently negatively associated with Task Performance (Miller 1956; Goodhue and Thompson 1995; Stanovich and West 2000). That is why we additionally investigate the relationships between the body-related psychophysiological NeuroIS indicators of Cognitive Workload with objective and subjective Task Performance indicators to shed more light on Cognitive Workload as a psychological mind-concept.

We aim to contribute to the following research questions:

RQ1: Are NeuroIS measures the more objective ones (Dimoka et al. 2012, p. 679)?
RQ2: Does a combination of various NeuroIS tools increase validity (Riedl et al. 2010, p. 248)?
RQ3: Are NeuroIS measures suitable for triangulation (Dimoka et al. 2011, p. 692)?
RQ4: Which psychophysiological measures are the most predictive ones in terms of Task Performance?

At this stage of research the most important contributions from our work are:

1. We found evidence that NeuroIS measures are more objective and that a combination of various NeuroIS tools increases validity – supporting the corresponding NeuroIS claims (RQ1 and RQ2).
2. The worst performers had a much greater Cognitive Workload during their Task Performance compared to top performers, which was coherently measurable by every single NeuroIS indicator – supporting the idea of effective triangulation and the corresponding NeuroIS claim (RQ3).
3. Based on the effect sizes at this stage of research, the number of (long) gaze fixations and electrodermal variance seems to be most predictive in terms of Task Performance (RQ4).

The paper is organized as follows: Next we show the research background, the research model and subsequently derive the hypotheses. After that the research methodology, including the test procedure and the MS Excel tasks used as well as the description of the NeuroIS Cognitive Workload measurements are presented before the first results are shown and discussed.

Research Background, Research Model and Hypotheses

Since we are interested in finding out whether NeuroIS measures are more objective than self-ratings (see RQ1) and if these NeuroIS measures are suitable for triangulation (see RQ3), we used both subjective and psychophysiological Cognitive Workload indicators.

In order to assess Cognitive Workload using NeuroIS and subjective instruments we wove a web around the Cognitive Workload concept by studying the relevant literature and identified the most important influencing and effecting factors: Various theoretical perspectives (e.g. cognitive load theory, task technology fit, job demands-resources) conceptualized Cognitive Workload as an important information systems construct that strongly influences Task Performance. The related paradigm builds on the principle that user's cognitive capacity is limited. Despite the problem that measuring Cognitive Workload is very difficult (Payne 1982; Sweller et al. 1998), it can be characterized by the cost of thinking or the amount of elementary mental operations (Newell and Simon 1972). Since Task Demands and Users Resources were identified as the main antecedents of Cognitive Workload from information processing theory and all other subsequent or similar theories on memory processing (e.g. Evaristo et al. 1995), we integrated these factors in our research model (see figure 1).

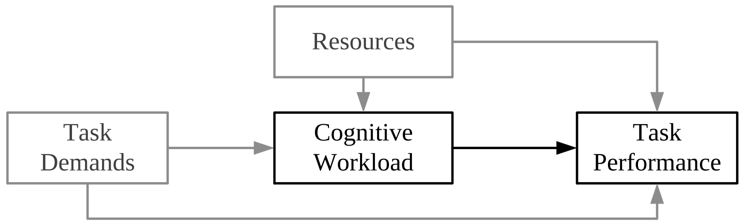

Figure 1. Research Model (grey – full model; black – reduced model for ICIS short paper)

To assess RQ1 to RQ4 we triangulate various NeuroIS Workload indicators with self-rated (Subjective) Cognitive Workload assessments. Based on the research questions and a philosophical science perspective (mind-body problem) we derived the following hypotheses.

If physicalism is true, subjective ratings simply reflect brain activity. If dualism is true, brain and mind possibly interact, but at least the mind adapts itself to the brain's processes. In both cases a Cognitive Workload assessment could be triangulated by NeuroIS if brain and mind respond with similar results, which increases validity (RQ2 and RQ3). Thus we hypothesize:

H1: NeuroIS Workload indicators are positively associated with Subjective Cognitive Workload.

In addition, the known negative relationship between Workload and Task Performance should also be true for both accesses – the mind and the brain (RQ2 and RQ3). Thus we hypothesize:

H2: NeuroIS Workload indicators are negatively associated with Subjective Task Performance.
H3: Subjective Cognitive Workload is negatively associated with Objective Task Performance.

From a purely materialistic body level perspective (which is true for physicalism and dualism) we hypothesize (RQ2 and RQ3):

H4: NeuroIS Workload indicators are negatively associated with Objective Task Performance.

If the NeuroIS assessments are more objective than subjective ones (see RQ1) which is in accordance with the physicalism theory approach, NeuroIS Workload indicators should be more strongly negatively associated with Objective Task Performance than Subjective Workload indicators are negatively associated with Objective Task Performance. Based on physicalism we hypothesize:

H5: NeuroIS Workload indicators are more strongly negatively associated with Objective Task Performance than Subjective Workload indicators are negatively associated with Objective Task Performance.

From a mind level perspective (which is true for both physicalism and dualism) we hypothesize:

H6: Subjective Cognitive Workload is negatively associated with Subjective Task Performance.

From physicalism we hypothesize:

H7: Subjective Task Performance is positively associated with Objective Task Performance.

Methodology

Test Procedure

The experiment was divided into 9 stages. Prior to all data collection, each test participant was welcomed by the experimenter (supervisor) who explained the key points of the experiment by reading aloud a detailed guide (stage 1). After that, the participant had to fill-out a consent form and the EDA device was applied (stage 2). In stage 3 a pre-interview was conducted in order to capture qualitative data concerning MS Excel skills and familiarity. In stage 4 the supervisor took the necessary precautions for the experiment, for which we made use of eye-tracking, EEG, and facial action analysis. Hence, all the psychophysiological systems were calibrated. In stage 5 the supervisor turned the task-sheet over, which included a short note describing that in what followed the participant would be asked to solve 7 MS Excel tasks within a time limit of approximately 5 minutes per task. If the participant signaled their understanding of the task-sheet, the supervisor started the computer task system synchronized with the psychophysiological systems (stage 6). At this stage the computer initially presented a questionnaire with demographics and a general attitude and emotion questionnaire. After that the 7 MS Excel tasks were presented in a balanced way. After each MS Excel task the computer presented the NASA TLX assessment. After finishing the MS Excel tasks the computer presented the general attitude and emotion questionnaire again (stage 8). At the end of stage 8 the EEG was taken down. In stage 9 a qualitative post-interview was conducted.

Please note that in this short research-in-progress paper we only show the results from stages 4 to 8 and not the results from the qualitative interviews or from the general attitude and emotion questionnaire, all of which will be part of the extended manuscript.

Microsoft Excel Spreadsheet Tasks

According to the cognitive load theory based approach of Cerpa et al. (1995) and based on Yi and Davis' (2003) procedure we systematically developed seven MS Excel spreadsheet tasks with different task complexity to induce various intrinsic cognitive loads (see table 1). Task presentation was balanced in terms of task difficulty.

No	Task description (shortened)	Difficulty level (1-7)	Success rate
1	Simple calculation and labeling	1	74 %
2	Completion of a number sequence	2	66 %
3	Filter function within a given data table	3	62 %
4	Creation of a bar chart with given data table	4	53 %
5	Some standard calculations (+, -, *, sum)	5	45 %
6	Rounded mean of given data column	6	34 %
7	Complex formatting (thousands separator, date, Roman number format)	7	11 %

Table 1. Task descriptions, difficulty level [1-very easy .. 7-very hard], and users success rate (percentage of successfully solved tasks within sample)

Psychophysiological NeuroIS Cognitive Workload Measurements

We asked a participants' brain using 10 NeuroIS Cognitive Workload indicators. Thereby we follow guideline 2 of vom Brocke and Liang (2014) and carefully applied the neuroscience standards (Eye-Tracking: Duchowski (2007), Eyegaze Edge™ manual; EDA: Boucsein (2012), MentalBioScreen K3 manual; EEG: Delorme and Makeig (2004), Emotiv EPOC manual; Facial actions: Ekman et al. (2002), Noldus FaceReader™ 6 manual).

<u>Pupillary diameter</u>: Pupillary assessment is an unobtrusive nonreactive research method well established in psychology (Goldinger and Papesh 2012; Laeng et al. 2012) and physiology (Loewenfeld 1999; Steinhauer et al. 2004) and also used in IS research (Cegarra and Chevalier 2008). It was coherently found that an increase of Cognitive Workload directly leads to a pupillary dilation. Prior research found that the pupillary diameter standard deviation (S.D.) is the proper Cognitive Workload measure within realistic task environments where sequences of sub-tasks (e.g. information search, reading, recall of information, writing) lead to a recurring demand for mental activity – resulting in a variability of the pupils (Buettner 2013). To capture the pupillary diameter, eye-tracking was performed using the binocular double 60 Hz Eyegaze Edge™ System with a 19" LCD monitor (86 dpi).

<u>Gaze fixations</u>: Information can only be perceived during gaze fixations and not 75msec before saccades start, during saccades, and 50msec after saccades (Leigh and Kennard 2004). Increases in Cognitive Workload correlates with an increase of fixations >250 msec and long fixations >500 msec (Just and Carpenter 1976, 1980; Rayner 1998; Van Orden et al. 2001). These (long) fixations indicate a (deep) cognitive processing. We used the number of fixations >250 msec and long fixations >500 msec as Cognitive Workload indicators. In addition we measured the interquartile range of fixation duration Q.75-Q.25 as a robust dispersion measure. Gaze fixation capturing was performed by the eye-tracking system described above.

EDA: According to Boucsein (2012) we measured Cognitive Workload by the standard deviation (S.D.) of the EDA signal using the MentalBioScreen K3 device, which traces participants' EDA each second. Therefore, two sensors (ECG electrodes, 10-W 55 GS) were applied to the non-dominant hand of the participant (1 channel).

EEG: Fairclough et al. (2005) found increased θ and β power during demanding task activity. Gevins et al. (1979) found increased θ, β and α power during Task Performance while the effect for α is not of as much importance. Ray and Cole (1985) found increased α activity reflects attentional demands and higher β power cognitive processes. That is why we used θ, β and α power as Cognitive Workload indicators from a wireless high resolution Emotiv EPOC system (14 channels: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4).

Facial action analysis: Stone and Wei (2011) found that facial actions measured by the Facial Action Coding System (FACS, cf. Ekman and Friesen (1978); Ekman et al. (2002)) correlates with Cognitive Workload. In order to analyze 20 FACS-based action units (AUs: inner brow raiser, brow lowerer, etc.) of the participants we used an Axis IP HD 720p camera which was installed in front of the participant above the monitor in combination with the Noldus FaceReader™ 6 software – calculating the AU-based Cognitive Workload measure [0..1].

Subjective and Objective Measurements

All constructs of the research model (figure 1) were also operationalized by subjective measures (table 2). Subjective Task Demand (STD) is the individual self-perception of how mentally demanding the task was (Hart and Staveland 1988). STD, Subjective Cognitive Workload (SCW) and Subjective Task Performance (STP) were assessed using the NASA TLX assessment by Hart and Staveland (1988); see table 2.

Construct	Abbr	Item text	Load	M	S.D.	Ref.
Subjective Skill Level (α= .876, CR = .924)	SSL-1	My knowledge of Microsoft Excel is quite good.	.805	2.77	1.540	Self-dev. based on Marakas et al. (2007, 1998)
	SSL-2	I don't have much experience with Microsoft Excel. [reversed coded]	.913	3.36	1.887	
	SSL-3	I am experienced in dealing with Microsoft Excel.	.963	3.02	1.745	
Subjective Familiarity Level (α = .881, CR = .925)	SFL-1	I use Microsoft Excel on a daily basis.	.875	2.49	2.172	Self-dev. based on Venka-tesh et al. (2012)
	SFL-2	I do not use Microsoft Excel. [reversed coded]	.891	3.88	2.205	
	SFL-3	I frequently use Microsoft Excel.	.924	2.86	1.937	
Subj. Task Demand	STD	How mentally demanding was the task?	single	3.57	1.866	Hart and Staveland (1988)
Subj. Cogn. Workload	SCW	How hard did you have to work to accomplish your level of performance?	single	3.63	1.916	
Subj. Task Performan.	STP	How successful were you in accomplishing what you were asked to do?	single	4.44	2.216	

Table 2. Items used for subjective and objective measurements (7 point Likert-scale)

The resource related constructs 'Skill Level' and 'Familiarity Level' were self-developed based on the computer self-efficacy argumentation by Marakas et al. (2007, 1998) for the Skill Level concept and the computer usage argumentation by Venkatesh et al. (2012) for the Familiarity Level concept. A user's Skill Level means the level of competence in terms of using MS Excel. The Familiarity Level describes how familiar and experienced a user is with handling MS Excel due to a certain level of usage in the past. We carefully evaluated both new measures in terms of internal consistency reliability, which is given for the new concepts as both values, Cronbach's α and Composite Reliability CR, were above .7 (table 2). Indicator reliability is also given since all factor loadings are above .707 (table 2) and they were all significant at a .001 level (nonparametric bootstrapping with 5,000 samples). Concerning convergent validity we can report that average variance extracted values are above .5 (.8067 for Familiarity Level and .8035 for Skill Level). Furthermore we successfully evaluated discriminant validity in terms of cross loadings. All items load substantially higher on the respective construct than on the other unrelated construct (least difference: .197) and the Fornell-Larcker criterion is also fulfilled for each concept since the square roots of the average variance extracted values of the constructs (.8982 and .8964) are greater than the correlation of both constructs (.7187).

Since the survey was conducted in Germany, items were translated and adjusted to meet the specific requirements of the German language according to Brislin (1970). Objective Task Performance was measured by the number of successfully solved tasks [0..7].

First Results and Discussion

Sample Characteristics

We recruited the first 53 working professionals to take part in a computer experiment. Please note that the whole experiment is still running and we expect to recruit >200 participants. Here we report the results of the first 53 participants (27 male, 26 female) aged from 19 to 67 years (M=29.9, S.D.=14.9) having from .5 to 42 years working experience (M=9.4, S.D.=12.9). All results exclusively concerning the hypotheses with objective and subjective constructs (none with psychophysiological NeuroIS indicators) are based on this n=53 sample.

To briefly sketch the promising research-in-progress results concerning the NeuroIS indicators, we draw on those participants who were either very successful (6 or 7 correct task solutions) or very unsuccessful (0 or 1 correct task solution) and compare these both groups concerning their psychophysiological NeuroIS Workload indicators. We controlled the selection for age since age often influences physiological responses. As the selection result, the very unsuccessful group consists of the 8 worst performers aged from 19 to 28 (M 22.6, S.D. 3.2). The very successful group consists of the 6 top performers aged from 19 to 29 (M 23.7, S.D. 3.9). As expected, the top-performer group were more skilled (SSL 4.9 vs. 2.1) and more familiar (SFL 4.1 vs. 2.0) with MS Excel compared to the worst-performer group (self-rating).

Psychophysiological NeuroIS based Assessment of Cognitive Workload

Pupillary data were cleaned by eye blinks (Verney et al. 2001). Pupillary, gaze fixation and EDA indicators were calculated with SPSS 17.0. We controlled all EDA values for room temperature (all n.s.). EEG data were filtered and preprocessed after removing baselines according to the EEGLAB manual (Delorme and Makeig 2004). Based on the EEGLAB toolbox 13.4.4b in conjunction with MATLAB R2015a we computed the Cognitive Workload related power spectra (θ, α, β II) over all 14 scalp electrodes (including bandpass filtering, and ICA). The facial AU-indicator was computed using the Noldus FaceReader™ 6. For the worst- and top-performer groups all NeuroIS Workload indicators are shown in table 3. Please note the excellent effect sizes (table 3). As a result we found significantly increased pupil diameter and EDA variation (S.D.) within the worst-performer group – both indicating more Cognitive Workload in this group (Boucsein 2012; Buettner 2013). In addition we found that worst-performers showed a greater time spread of gaze fixation and had much more gaze fixation >250msec/500msec which also indicates higher Cognitive Workload (Rayner 1998; Van Orden et al. 2001).

Furthermore, we found significantly increased θ, α and β EEG power within the worst-performer group – also indicating more Cognitive Workload (Fairclough et al. 2005; Gevins et al. 1979; Ray and Cole 1985). Finally, we found that the worst-performers had much more facial activity than the top-performers which also shows higher Cognitive Workload (Ekman et al. 2002; Stone and Wei 2011). These coherent results support the neural-efficiency hypothesis which states that more skilled and experienced individuals display lower (more efficient) brain activation while performing cognitive tasks (Neubauer and Fink 2009). As noted earlier our top-performers were much more skilled and more familiar with MS Excel.

Source	NeuroIS indicator	Ø Worst-Perform.	Ø Top-Perform.	Signific.	Cohen's d
Eye-Tracking	S.D. (s) Left Pupil	.273 mm	.200 mm	$p < .1$	.72
	S.D. (s) Right Pupil	.270 mm	.192 mm	$p < .1$	.89
	Fixation Duration Q.75-Q.25	.159 sec	.141 sec	$p < .1$	.75
	No. Fixations > 250msec	1,943.5	1,263.5	$p < .05$	.85
	No. Fixations > 500msec	471.0	336.8	$p < .1$	.76
EDA	S.D. (s) EDA signal	34.12 µS	21.13 µS	$p < .1$	.83
EEG	Theta θ (4-8 Hz) Power	622.422	101.038	$p < .001$	.56
	Alpha α (8-13 Hz) Power	249.443	70.665	$p < .001$	.50
	Beta β II (15-21 Hz) Power	75.521	36.479	$p < .01$	.34
Face	Facial Action Unit Indicator	.318	.291	$p < .1$	.83

Table 3. NeuroIS indicators of worst vs. top-performers. [MATLAB, FaceReader, SPSS]

Concerning the relationships between NeuroIS indicators and subjective as well as objective measures we calculated the correlations over all n=14 worst- and top-performers (table 4).

Neuro IS indicator	Obj Task Perf	Subj Task Perf	Subj Skills	Subj Familiarity	Subj Task Demand	Subj Cogn Workload
S.D. (s) Left Pupil	-.038^{ns}	-.249^{ns}	-.015^{ns}	-.357^{ns}	-.011^{ns}	-.011^{ns}
S.D. (s) Right Pupil	-.288^{ns}	-.436*	-.222^{ns}	-.389*	.231^{ns}	.189^{ns}
Fixation Duration Q.75-Q.25	-.365*	-.605**	-.461**	-.097^{ns}	.465**	.605**
No. Fixations > 250msec	-.527**	-.311^{ns}	-.610**	-.152^{ns}	.240^{ns}	.203^{ns}
No. Fixations > 500msec	-450*	-.407*	-.652***	-.163^{ns}	.209^{ns}	-231^{ns}
S.D. (s) EDA signal	-.486**	-.509**	-.192^{ns}	-.265^{ns}	.389*	.476**
Theta θ (4-8 Hz) Power	-.378*	-.112^{ns}	-.251^{ns}	.055^{ns}	.178^{ns}	.086^{ns}
Alpha α (8-13 Hz) Power	-.367*	-.174^{ns}	-.233^{ns}	.218^{ns}	.218^{ns}	.110^{ns}
Beta β II (15-21 Hz) Power	.002^{ns}	.097^{ns}	-.040^{ns}	.474**	-.068^{ns}	-.101^{ns}
Facial Action Unit Indicator	-.395*	-.222^{ns}	-.077^{ns}	-.060^{ns}	.371*	.273^{ns}
NeuroIS Correlation Share	7 of 10	4 of 10	3 of 10	2 of 10	3 of 10	2 of 10

Table 4. Correlation matrix NeuroIS – objective/subjective measures. $p < .1^{*}$, $.05^{**}$, $.01^{***}$

Results from table 4 show that 7 of 10 NeuroIS Workload indicators significantly correlate with Objective Task Performance at a medium-large effect level. In contrast, only 4 of 10 NeuroIS Workload indicators significantly correlate with Subjective Task Performance. This result is interesting because for the first time it shows evidence that the NeuroIS Workload indicators – Objective Task Performance relationships are stronger than the NeuroIS Workload indicators – Subjective Task Performance relationships. In conjunction with the result that Subjective Cognitive Workload correlates stronger with Subjective Task Performance than Objective Task Performance (-.527), we have some evidence that psychophysiological NeuroIS Workload indicators are actually more objective (RQ1). Since nearly all (20 of 21) significant NeuroIS Workload indicator correlations point in the same direction (table 4), we also found indications that a combination of various NeuroIS tools increases validity (RQ2). Results from table 4 show that NeuroIS measures are suitable for triangulation (RQ3). Based on the effect sizes at this stage of the research, the number of (long) gaze fixations and EDA variance seems to be most predictive in terms of TP (RQ4).

Hypotheses Evaluation

To evaluate the hypotheses we built a multitrait-multimethod matrix (Campbell and Fiske 1959) for Cognitive Workload and Task Performance, studied from both perspectives: objective and subjective. In this short paper we show the matrix results for the NeuroIS Workload indicator 'number of long gaze fixations > 250 msec' as one of the ten applied NeuroIS indicators.

			SUBJECTIVE		OBJECTIVE	
	Concept		Cogn Workload	Task Perf	Cogn Workload	Task Perf
		Measure	NASA TLX	NASA TLX	Fix > 250 msec	Correct solutions
SUBJECTIVE	Cogn WL	TLX	(0.604)	-0.527** H6	0.203ns H1	-0.393** H3/H5
	Task Perf	TLX	-0.527** H6	(0.952)	-0.311ns H2	0.653** H7
OBJECTIVE	Cogn WL	Fix250	0.203ns H1	-0.311ns H2	(0.786)	-0.527* H4/H5
	Task Perf	Corr Sol	-0.393** H3/H5	0.653** H7	-0.527* H4/H5	(1.000)

Table 5. Multitrait-Multimethod Matrix (diagonal shows reliability). p < .05*, .0001**

H1 and H2 cannot be supported at the current stage of research (n=14) since there is no statistically relevant relationship between the objective NeuroIS Workload indicator 'fixations > 250msec' and the subjective user evaluations of Workload and Task Performance. That is interesting, because it seems that subjective ratings do not simply reflect brain activity.

H3 and H4 are supported and, interestingly, Objective Task Performance correlates more strongly negatively with the objective NeuroIS Workload indicator 'fixations > 250msec' (-0.527) than with the Subjective Cognitive Workload assessment by NASA TLX (-0.393), which leads to support for H5. This underlines the result above, i.e. psychophysiological NeuroIS Workload indicators are actually more objective (RQ1).

H6 is supported since Cognitive Workload and Task Performance were subjectively coherently rated by the participants with a substantial negative relationship between them (-.527).

H7 is supported (.653). While it was hard for the participants to evaluate the mind-related Cognitive Workload concept, it seems to be easier to self-assess the manifest concept of Task Performance. In addition and as shown in table 5, from the subjective as well the objective perspective, the reliability of the mental concept 'Cognitive Workload' is lower than that of the more manifest concept 'Task Performance'. This result emphasizes that the mental concept 'Cognitive Workload' is more error-/bias-prone and hard to assess in comparison to manifest concepts. Because of this and since NeuroIS is costly it is often not useful to assess / triangulate manifest concepts such as Task Performance using NeuroIS methods.

However, as shown above, the NeuroIS Workload indicator 'fixations > 250msec' is more closely related to Objective Task Performance than Subjective Cognitive Workload is. That means that NeuroIS instruments studying the brain directly open an avenue to more objective evaluations. That is why NeuroIS can add fruitful (complementary or contrary) insights to subjective self-ratings of hardly accessible and error-/bias-prone concepts such as Cognitive Workload (RQ2 and RQ3).

Future Research

Determining a user’s Cognitive Workload is often mentioned as a fundamental problem in information systems research (Evaristo et al. 1995; Johannsen et al. 1992; Stassen et al. 1990), particularly in NeuroIS (de Guinea et al. 2014; Dimoka et al. 2012, 2011; Riedl et al. 2010). It is remarkable that scholars have traditionally investigated a Cognitive Workload and its derivatives (Cain 2007) primarily based on user-perceived/nonobjective measures (Ayyagari et al. 2011; Gupta et al. 2013; Ragu-Nathan et al. 2008; Tarafdar et al. 2010) or even discussed the need for more objective Cognitive Workload measurements without any measurement proposal (Evaristo et al. 1995; Wastell 1999). The discourse on this topic has shown the need to quantify Cognitive Workload based on objective physiological parameters (Dimoka et al. 2011, 2012; Riedl et al. 2010).

Despite Cognitive Workload being hard to assess, its evaluation by NeuroIS instruments could be fruitful for information systems scholars and could offer new data to evaluate workload-related information systems and psychological theories, e.g. information processing theory (Miller 1956), cognitive load theory (Sweller 1988; Sweller et al. 1998), flow theory (Csíkszentmihályi 1975), job demands-resources theory (Bakker and Demerouti 2006), Task-Technology Fit Theory (Goodhue and Thompson 1995), psychobiological theory (Kock 2004), and dual process theory (Stanovich and West 2000; Evans 2003). In addition, studies using NeuroIS instruments may even help to uncover new constructs.

In the extended manuscript we will also report on the question of whether we can localize neural Workload correlates and whether we can shed light on the muddle of various types of Workload by assessing these neural correlates both by using differentiating various tasks 1-7 and EEG scalp locations.

References

Ayyagari, R., Grover, V., and Purvis, R. 2011. "Technostress: Technological Antecedents and Implications," *MIS Quarterly* (35:4), pp. 831–858.

Bakker, A. B., and Demerouti, E. 2006. "The Job Demands-Resources model: state of the art," *Journal of Managerial Psychology* (22:3), pp. 309–328.

Boucsein, W. 2012. Electrodermal Activity, Springer, 2nd ed.

Brislin, R. W. 1970. "Back-Translation for Cross-Cultural Research," *Journal of Cross-Cultural Psychology* (1:3), pp. 185–216.

Buettner, R. 2013. "Cognitive Workload of Humans Using Artificial Intelligence Systems: Towards Objective Measurement Applying Eye-Tracking Technology," in *Proceedings 36th Annual German Conference on AI,* Koblenz, Germany, I. J. Timm, and M. Thimm (eds.), vol. 8077 of LNAI, pp. 37-48.

Campbell, D. T., and Fiske, D. W. 1959. "Convergent and discriminant validation by the multitrait-multimethod matrix". *Psychological Bulletin* (56:2), pp. 81–105.

Cain, B. 2007. "A Review of the Mental Workload Literature," Report, NATO, Toronto.

Cegarra, J., and Chevalier, A. 2008. "The use of Tholos software for combining measures of mental workload: Toward theoretical and methodological improvements," *Behavioral Research Methods* (40:4), pp. 988–1000.

Cerpa, N., Chandler, P., and Sweller, J. 1995. "Some Consequences of Training Strategies when Learning a Computer Application," in *Proceedings of the Australian Association for Research in Education Conference 1995*, Hobart, Tasmani.

Csíkszentmihályi, M. 1975. *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*, San Francisco: Jossey-Bass.

de Guinea, A. O., Titah, R., and Léger, P.-M. 2014. "Explicit and Implicit Antecedents of Users' Behavioral Beliefs in Information Systems: A Neuropsychological Investigation," *Journal of Management Information Systems* (30:4), pp. 179–210.

Delorme, A., and Makeig, S. 2004. "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods* (134:1), pp. 9–21.

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., and Weber, B. 2012. "On the Use of Neurophysiological Tools in IS Research: Developing a Research Agenda for NeuroIS," *MIS Quarterly* (36:3), pp. 679–A19.

Dimoka, A., Pavlou, P. A., and Davis, F. D. 2011. "NeuroIS: The Potential of Cognitive Neuroscience for Information Systems Research," *Information Systems Research* (22:4), pp. 687–702.

Duchowski, A. T. 2007. Eye Tracking Methodology: Theory and Practice, London, UK: Springer, 2nd ed.

Ekman, P., and Friesen, W. V. 1978. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., and Hager, J. C. 2002. *The Facial Action Coding System*, London: Weidenfeld & Nicolson, 2nd ed.

Evans, J. S. B. 2003. "In two minds: dual-process accounts of reasoning," *Trends in Cognitive Sciences* (7:10), pp. 454–459.

Evaristo, R., Adams, C., and Curley, S. 1995. "Information Load Revisited: A Theoretical Model," in *Proceedings of the 16th International Conference on Information Systems, G. Ariav, C. M. Beath, J. I. DeGross, R. Hoyer, and C. F. Kemerer (eds.),* Amsterdam, The Netherlands, Paper 18.

Fairclough, S. H., Venables, L., and Tattersall, A. 2005. "The influence of task demand and learning on the psychophysiological response," *International Journal of Psychophysiology* (56:2), pp. 171–184.

Gevins, A. S., Zeitlin, G. M., Doyle, J. C., Yingling, C. D., Schaffer, R. E., Callaway, E., and Yeager, C. L. 1979. "Electroencephalogram Correlates of Higher Cortical Functions," *Science* (203:4381), pp. 665–668.

Goldinger, S. D., and Papesh, M. H. 2012. "Pupil Dilation Reflects the Creation and Retrieval of Memories," *Current Directions in Psychological Science* (21:2), pp. 90–95.

Goodhue, D. L., and Thompson, R. L. 1995. "Task-Technology Fit and Individual Performance," *MIS Quarterly* (19:2), pp. 213–236.

Gupta, A., Li, H., and Sharda, R. 2013. "Should I send this message? Understanding the impact of interruptions, social hierarchy and perceived task complexity on user performance and perceived workload," *Decision Support Systems* (55:1), pp. 135–145.

Hart, S. G., and Staveland, L. E. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, pp. 139–183, in P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*. Amsterdam: North Holland Press.

Johannsen, G., Levis, A. H., and Stassen, H. G. 1992. "Theoretical Problems in Man-machine Systems and Their Experimental Validation," *Automatica* (30:2), pp. 217–231.

Just, M., and Carpenter, P. 1976. "Eye fixations and cognitive processes," *Cognitive Psychology* (8:4), pp. 441–480.

Just, M. A., and Carpenter, P. A. 1980. "A theory of reading: From eye fixations to comprehension," *Psychological Review* (87:4), pp. 329–354.

Kim, J. 2000. *Mind in a Physical World*, Cambridge, MA, U.S.: MIT Press.

Kock, N. 2004. "The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication Based on Darwinian Evolution," *Organization Science* (15:3), pp. 327–348.

Laeng, B., Sirois, S., and Gredebäck, G. 2012. "Pupillometry: A Window to the Preconscious?" *Perspectives on Psychological Science* (7:1), pp. 18–27.

Leigh, R. J., and Kennard, C. 2004. "Using saccades as a research tool in the clinical neurosciences," *Brain* (127:3), pp. 460–477.

Loewenfeld, I. E. 1999. *The Pupil: Anatomy, Physiology, and Clinical Applications*, Butterw.-Heinem., 2nd ed.

Marakas, G., Johnson, R., and Clay, P. F. 2007. "The Evolving Nature of the Computer Self-Efficacy Construct: An Empirical Investigation of Measurement Construction, Validity, Reliability and Stability Over Time," *Journal of the Association for Information Systems* (8:1), pp. 16–46.

Marakas, G. M., Yi, M. Y., and Johnson, R. D. 1998. "The Multilevel and Multifaceted Character of Computer Self-Efficacy: Toward Clarification of the Construct and an Integrative Framework for Research," *Information Systems Research* (9:2), pp. 126–163.

Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review* (63:2), pp. 81–97.

Neubauer, A. C., and Fink, A. 2009. "Intelligence and neural efficiency," *Neuroscience and Biobehavioral Reviews* (33:7), pp. 1004–1023.

Newell, A., and Simon, H. A. 1972. Human Problem Solving, Englewood Cliffs, N.J.: Prentice-Hall.

Payne, J. W. 1982. "Contingent Decision Behavior," *Psychological Bulletin* (92:2), pp. 382–402.

Ragu-Nathan, T. S., Tarafdar, M., Ragu-Nathan, B. S., and Tu, Q. 2008. "The Consequences of Technostress for End Users in Organizations: Conceptual Development and Empirical Validation," *Information Systems Research* (19:4), pp. 417–433.

Ray, W. J., and Cole, H. W. 1985. "EEG Alpha Activity Reflects Attentional Demands, and Beta Activity Reflects Emotional and Cognitive Processes," *Science* (228:4700), pp. 750–752.

Rayner, K. 1998. "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin* (124:3), pp. 372–422.

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., and Weber, B. 2010. "On the Foundations of NeuroIS: Reflections on the Gmunden Retreat 2009," *Communications of the Association for Information Systems* (27), pp. 243–264.

Stanovich, K. E., and West, R. F. 2000. "Individual differences in reasoning: Implications for the rationality debate," *Behavioral and Brain Sciences* (23:5), pp. 645–665.

Stassen, H. G., Johannsen, G., and Moray, N. 1990. "Internal representation, internal model, human performance model and mental workload," *Automatica* (26:4), pp. 811–820.

Steinhauer, S. R., Siegle, G. J., Condray, R., and Pless, M. 2004. "Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing," *International Journal of Psychophysiology* (52:1), pp. 77–86.

Stone, R. T., and Wei, C.-S. 2011. "Exploring the Linkage between Facial Expression and Mental Workload for Arithmetic Tasks," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (55:1), pp. 616-619.

Sweller, J. 1988. “Cognitive Load During Problem Solving: Effects on Learning,” *Cognitive Science* (12:2), pp. 257–285.

Sweller, J., van Merrienboer, J. J. G., and Paas, F. G. W. C. 1998. “Cognitive Architecture and Instructional Design,” *Educational Psychology Review* (10:3), pp. 251–296.

Tarafdar, M., Tu, Q., and Ragu-Nathan, T. S. 2010. “Impact of Technostress on End-User Satisfaction and Performance,” *Journal of Management Information Systems* (27:3), pp. 303–334.

Van Orden, K. F., Limbert, W., Makeig, S., and Jung, T.-P. 2001. “Eye Activity Correlates of Workload during a Visuospatial Memory Task,” *Human Factors* (43:1), pp. 111–121.

Venkatesh, V., Thong, J. Y. L., and Xu, X. 2012. “Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology,” *MIS Quarterly* (36:1), pp. 157–178.

Verney, S. P., Granholm, E., and Dionisio, D. P. 2001. “Pupillary responses and processing resources on the visual backward masking task,” *Psychophysiology* (38:1), pp. 76–83.

vom Brocke, J., and Liang, T.-P. 2014. “Guidelines for Neuroscience Studies in Information Systems Research,” *Journal of Management Information Systems* (30:4), pp. 211–234.

Wastell, D. G. 1999. “Learning Dysfunctions in Information Systems Development: Overcoming the Social Defenses With Transitional Objects,” *MIS Quarterly* (23:4), pp. 581–600.

Yi, M. Y., and Davis, F. D. 2003. “Developing and Validating an Observational Learning Model of Computer Software Training and Skill Acquisition,” *Information Systems Research* (14:2), pp. 146–169.