Identifying Cognitive Assistance with Mobile Electroencephalography: A Case Study with In-Situ Projections for Manual Assembly

THOMAS KOSCH, LMU Munich, Munich, Germany
MARKUS FUNK, TU Darmstadt, Darmstadt, Germany
ALBRECHT SCHMIDT, LMU Munich, Munich, Germany
LEWIS L. CHUANG, Max Planck Institute for Biological Cybernetics, Tübingen, Germany

11

Manual assembly at production is a mentally demanding task. With rapid prototyping and smaller production lot sizes, this results in frequent changes of assembly instructions that have to be memorized by workers. Assistive systems compensate this increase in mental workload by providing "just-in-time" assembly instructions through *in-situ* projections. The implementation of such systems and their benefits to reducing mental workload have previously been justified with self-perceived ratings. However, there is no evidence by objective measures if mental workload is reduced by *in-situ* assistance. In our work, we showcase electroencephalography (EEG) as a complementary evaluation tool to assess cognitive workload placed by two different assistive systems in an assembly task, namely paper instructions and *in-situ* projections. We identified the individual EEG bandwidth that varied with changes in working memory load. We show, that changes in the EEG bandwidth are found between paper instructions and *in-situ* projections, indicating that they reduce working memory compared to paper instructions. Our work contributes by demonstrating how design claims of cognitive demand can be validated. Moreover, it directly evaluates the use of assistive systems for delivering context-aware information. We analyze the characteristics of EEG as real-time assessment for cognitive workload to provide insights regarding the mental demand placed by assistive systems.

CCS Concepts: • **Human-centered computing → User studies**; *User models*; *Usability testing*;

Additional Key Words and Phrases: Cognitive Workload; Working Memory; Assistive Technology; Electroencephalography; Workload-Aware Computing; Cognition-Aware Computing

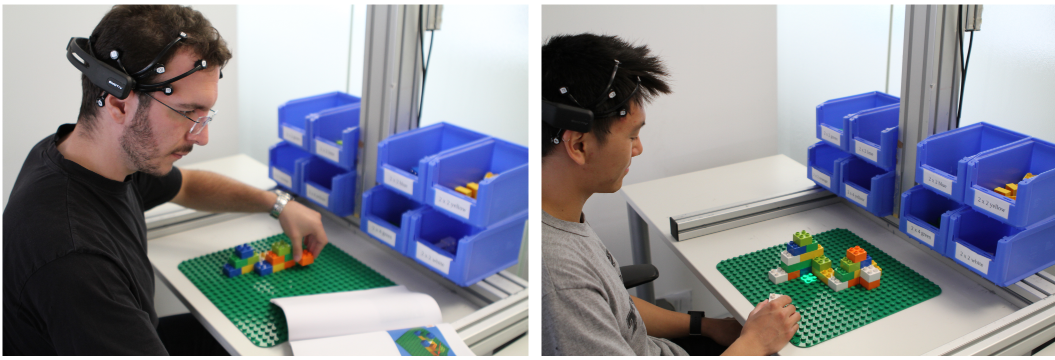
ACM Reference Format:
Thomas Kosch, Markus Funk, Albrecht Schmidt, and Lewis L. Chuang. 2018. Identifying Cognitive Assistance with Mobile Electroencephalography: A Case Study with In-Situ Projections for Manual Assembly. *Proc. ACM Hum.-Comput. Interact.* 2, EICS, Article 11 (June 2018), 20 pages. https://doi.org/10.1145/3229093

1 INTRODUCTION

Modern manufacturing processes are increasingly defined by smaller lot sizes of bespoke designs. Gone are the days that require workers to act purely on rote learning. Instead, novel assembly instructions must frequently be committed to memory as soon as new designs and components enter into the production pipeline [28]. Thus, assembly is increasingly defined by its cognitive

Authors' addresses: Thomas Kosch, LMU Munich, Munich, Germany, thomas.kosch@ifi.lmu.de; Markus Funk, TU Darmstadt, Darmstadt, Germany, funk@tk.tu-darmstadt.de; Albrecht Schmidt, LMU Munich, Munich, Germany, albrecht.schmidt@ifi.lmu.de; Lewis L. Chuang, Max Planck Institute for Biological Cybernetics, Tübingen, Germany, lewis.chuang@tuebingen.mpg.de.

2573-0142/2018/6-ART11 $15.00
https://doi.org/10.1145/3229093



(a) (b)

Fig. 1. **(a)** Paper-based and **(b)** projected *in-situ* instructions at a manual assembly workplace. An EEG headset measures the level of working memory. This provides objective insights about cognitive processes during the usage of assistive technology in manual assembly processes.

instead of physical demands. Manual assembly places workload on cognitive processes that underlie executive working memory [3, 29]. Given that cognitive workload impacts the individual task performance [18, 37], it follows that profit margins may suffer as a result of reduced production throughput and error-prone manufacturing.

Assembly instruction systems have been introduced to ameliorate this. Such assistance consists of instructions on printed paper, external displays, or *in-situ* instructions projected directly on the assembly workplace. Such systems are designed to simultaneously reduce cognitive demands and optimize performance. Indeed, previous research has suggested that design manipulations of the visual representation and modality of assembly instructions can have a significant influence on self-reported measures of cognitive workload [28, 44]. Strong arguments have been made for the use of *in-situ* instructions that present just-in-time assembly instructions [13, 25]. Furthermore, visual *in-situ* instructions integrated into the workplace showed a large increase in performance [27, 48].

Novel user interfaces – for instance, assistive instruction systems – are often evaluated for their improvements regarding performance or subjective workload in order to justify their implementation over the *status quo*. Performance is typically assessed in terms of error rate and assembly completion time, while cognitive workload is often assessed using self-rated measures or semi-structured interviews [22]. Commonly used self-rated measures include the NASA-TLX [35] or the simpler RSME [7, 86]. However, these metrics are susceptible to individual differences in subjective reporting. For instance, extroverted confident individuals might be less likely to indicate workload to the same degree as introspective modest individuals [39]. Furthermore, novel assistive systems and visualizations might introduce a novelty effect. This means, that participants provide self-reported scores which are influenced by the excitement over a new technology [81]. Additionally, subjective reports can only be performed after test completion and rely on cognitive processes, namely working memory. Such methods are limited in their objective and real-time assessment of cognitive workload when evaluating assistive technologies.

Therefore, physiological methods are increasingly employed as a means to estimate mental states, such as cognitive workload [34, 45, 47]. For example, Brain-Computer Interfaces [2, 79] (BCIs) have been used successfully to assess the complexity of presented information in a variety of scenarios [2, 46] (see Figure 1). A BCI records brain activity in real-time and computationally derives estimates of a targeted mental state to either provide feedback about the physiological state of the user or enable device control through neuronal activity. Increasingly, advances in brain-recording technology and neuroscience findings contribute towards a vision of ubiquitous BCI deployment in everyday places of work and play [10, 42, 80]. In the current context, a BCI for cognitive workload sensing could be used in collaborative work settings by allocating tasks to workers who are less fatigued [21, 22] or provide adaptive assistance whenever necessary [23]. However, for this future to be realized, it is necessary to ensure that recorded estimates are robust and valid for its given setting.

BCIs may record brain activity via Electroencephalography (EEG), which has been preferred for its lightweight and mobile hardware, high temporal resolution, and non-invasiveness nature [33]. EEG measures neuronal signals from scalp electrodes, relative to a reference electrode [20, 32]. Previous research has utilized EEG as a measure for cognitive workload to evaluate task difficulties in real-time [30, 31, 75]. To some extent, it is possible to rely on EEG to classify cognitive workload in real-time, particular those that involve working memory processes [33]. This makes EEG to a viable alternative or complementary metric for mental workload measures.

In this work, we employ a lightweight mobile EEG headset to evaluate cognitive workload, specifically executive working memory load, to provide a direct estimation of the cognitive benefits of implementing *in-situ* projection systems. Such systems are expected to achieve early market penetration [16, 49, 76] and allow to be used as an instantaneous evaluation tool for novel user interfaces as well as for assistive technologies. In a user study (N=12) we investigate differences in workload measures during manual assembly with two different instruction systems, namely printed paper instructions and projected *in-situ* instructions. We analyze EEG power in the individual alpha-frequency bands to infer cognitive workload on working memory processes. Previous research measured a decrease of frequencies in the alpha-band the more working memory participants experienced [30, 31]. We find that *in-situ* instruction systems significantly reduce executive working memory. This is complemented by a discussion considering our results and implications to assess working memory in real-time, where novel assistive technologies can be evaluated regarding the required mental demand using EEG. Finally, we describe a physiologically-based methodology about how to evaluate assistive technologies for manual assembly tasks and present an experimental protocol to assess assistive systems regarding working memory.

2 BACKGROUND

Previous research has investigated effort in examining EEG data while subjects perform specific tasks. We summarize relevant past research regarding (1) the use of EEG as a non-invasive method to obtain electrical potentials generated by the human brain, (2) how working memory is determined by EEG, and (3) how EEG measures are deployed in real-world scenarios.

2.1 Brain Sensing

The human brain represents the information and control unit of humans. It consists of approximately 100 billion neurons [36] which underlie the human cognition. This includes information processing, where neuronal activity can be measured non-invasively with the use of scalp electrodes. Neurons communicate by exchanging electrical activity using neurotransmitters, a chemical transferred between neurons. In this work, we focus on measuring this electrical activity via an EEG headset. EEG is commonly leveraged in clinical application and yields a non-invasive method to estimate brain activity [56, 63, 83]. By placing conductive electrodes on a scalp, electrical potentials between $1\mu v$ and $100\mu v$ (microvolts) are measured. These measures are relative to a reference point, which is another electrode attached to the scalp or earlobe [20, 55]. Typically, changes in electrical potentials are observed by analyzing frequency bands. For example, previous work found a drop in frequencies of alpha (8 - 12 Hz) and an increase of theta (4 - 8 Hz) [43] when subjects have to raise mental capacities. An alternative approach is the assessment of Event-Related Potentials (ERPs) to infer mental workload [9, 70]. However, ERPs are averaged over many trials to filter out noise which makes it unsuitable for real-time assessment.

Unfortunately, EEG measurements are prone to noise during measurement. Head movements, muscle contractions, eye movements, or eye blinks influence the electromagnetic field on the scalp. Researchers are concerned about this and invest great effort to reduce the number of measurement artifacts [14]. However, artifacts cannot be fully avoided. The usage of BCIs often requires a controlled environment comprising minimal body movements by the user. Such conditions are often impractical for end users due to its high experimental control. However, recent technical advances ameliorate these disadvantages [61]. Furthermore, BCIs have found access to the consumer market, priced between $249[1] and $1600[2]. Affordable open source solutions can be acquired within the OpenEEG Project[3]. In our work, we use such a low-cost EEG headset to investigate its feasibility regarding the assessment of working memory.

2.2 Quantifying Cognition using EEG

Much research has been performed to define and quantify cognitive workload. Sweller et al. [77, 78] defined three key components of cognitive workload comprising intrinsic, germane, and extraneous cognitive workload. Intrinsic workload describes the inherent complexity of the task itself, and can therefore not be manipulated by external sources. Germane workload describes the cognitive effort subjects need to comprehend and process new information. Extraneous workload describes the cognitive demand to understand and process the visual representation of the underlying information. Experiments often evaluate extraneous workload by influencing the visualization of information. Intrinsic and germane workload is not easy to manipulate since they are task related or depend on the individual mental capabilities.

More precisely, cognitive workload defines the mental effort being used in working memory, which is responsible for fast information processing and limited within its capacities [3]. Gevins et al. [31] invested effort to quantify the occurrence and amount of working memory using EEG. They found a decrease in alpha frequencies and an increase in theta frequencies during cognitively demanding tasks. Jensen et al. [38] replicated their results using a memory demanding task. Scharinger et al. [72] also showed a variety of other tasks which correlate with increased working memory. However, working memory strongly depends on individual attributes, such as age or neuronal health state [43].

Machine learning became popular to classify mental states using EEG data [52–54] and has been used to estimate workload in different contexts [60]. Lee et al. [51] collected EEG data of participants to classify cognitive demanding tasks according to their difficulty. Participants were asked to do different cognitive tasks, including mental rotation and mental arithmetic. Instead of using previously recorded data associated with similar tasks [1, 41], data was collected from scratch. Using machine learning, a classification accuracy of 92.4% was achieved. However, the experiment has not focused on classifying working memory itself. An approach to classify working memory was undertaken by Grimes et al. [33]. Participants performed a N-back task, which is a popular task to demand the short-term memory [58, 66]. Their study included four different difficulties, where their results yielded a classification accuracy of 99% for binary classification and an accuracy of 88% for classification between four classes of difficulty. Nonetheless, a medical EEG headset was

[1] www.choosemuse.com - last access 2018-05-28
[2] www.futurehealth.org/bm_at1.htm - last access 2018-05-28
[3] www.openeeg.sourceforge.net - last access 2018-05-28

(a) (b) (c)

Fig. 2. **(a):** Snippet from the used Lego Duplo paper instruction. The necessary brick can be seen in the upper left corner. A red arrow denotes the final assembly position of the brick. **(b):** Highlighted box using *in-situ* projections. **(c):** Projection of the target position of a selected brick.

utilized for their experiment, which is hard to deploy on the fly. Furthermore, an artificial task was used to measure the level of working memory instead of using a real-world scenario.

2.3 Evaluating and Adapting User Interfaces

Besides using EEG for quantifying and classifying working memory researchers have employed this measurement to enable the development of cognition-aware systems. Zander et al. [85] investigated how passive brain activity measurements can be leveraged to handle take-over tasks during autonomous driving. Prinzel et al. [50] researched how task allocation can be carried out efficiently using EEG-related metrics. Finally, El-Komy et al. [17] used physiological sensing devices in an assembly environment comprising additional artificial tasks to measure the workers' stress levels. However, only results from the emotion and arousal scores retrieved by the integrated BCI interface were reported.

To the best of our knowledge, no prior work concerning the evaluation of EEG to derive cognitive workload during manual assembly and with different instruction systems has been done. In this work, we close this gap by utilizing standardized manual assembly tasks with two different instruction systems. We record and compare EEG data in a repeated measures experimental design to evaluate EEG as a valid assessment tool for cognitive workload induced by assistive technologies during manual assembly.

3 ASSEMBLY INSTRUCTION SYSTEMS

We identified two different instruction systems from related research, which are different regarding subjective perception [25]. Within both instruction systems, Lego Duplo bricks are assembled. This task resembles a full replacement for a real assembly task and enables to change the complexity of the task without changing the task itself [25, 48].

3.1 Assembly Instruction Visualizations

Informed by related work, two instruction visualizations are identified which differ in overall interpretation complexity [25]. We have chosen paper instructions, as they represent the current state of the art when it comes to transfer assembly instructions in manual assembly lines [24]. This is compared to projected *in-situ* instructions, where assembly instructions are projected on the workplace. We have chosen these two assembly instruction modalities as subjective measures suggest an alleviation of workload for *in-situ* projected instructions compared to printed paper instructions [25, 27]. In the following paragraphs, we describe both instruction systems used in the study in detail.

3.1.1 Paper Instructions. We printed single-sided instructions on an A4 sheet of paper. Each work step was printed on a single page, such that the position and size of every step were the same. The paper instructions were put together in the correct order using a folder and positioned to the left or right of the users, depending on their handedness. The folded assembly instruction remained the same position relative to the user's position. The instruction shows a brick on the upper left corner which is required to select for completing the current work step. Marked by an arrow, the assembly instruction shows the final position of the brick (see Figure 2a).

3.1.2 In-Situ Instructions. We compare the presented paper instructions with *in-situ* projections displaying assembly instructions. We use a similar system as shown by Funk et al. [26]. A projector mounted above the work table displays the next assembly step on the workspace. A Kinect v2 validates each work step of the Lego Duplo construction. This includes the verification of correct item selections from bins by observing the hand movements of the participant (see Figure 2b) and assembly steps by comparing the (see Figure 2c). The next work step is displayed when the current work step is performed correctly. The system is waiting until the current work step is carried out correctly and does not proceed if the user makes an error.

4 MANIPULATING COGNITIVE WORKLOAD

We manipulate working memory to assess the validity of our setup. We use a visual N-back task [58] with two levels of task difficulty (N= 0 and N= 2) to induce cognitive workload, namely executive working memory. In the N-back task a series of numbers is presented (i.e., numbers). Each symbol appears at a fixed position. Upon symbol representation, participants have to decide if the current symbol is equal to the symbol shown N steps ago. Participants have to keep a sequence of N symbols constantly in their memory, decide if a match occurred, and then update the sequence in their memory.

For example, the *0*-back task requires participants to compare each displayed symbol with the first one seen in the series. Since the currently displayed number matches with the first one in the series during the *0*-back task, no memory updates are required. However, the task difficulty can be manipulated by changing N [30, 33, 66]. For N= 2, participants have to memorize the last two symbols in the series while paying attention to matches when the currently displayed symbol is the same symbol as shown two items ago. Table 1 shows an example of the N-back.

Displayed number	5	8	3	4	3	9	1
Press button (*0*-back task)	5	8	3	4	3	9	1
Press button (*1*-back task)		5	8	3	4	3	9
Press button (*2*-back task)			5	8	3	4	3
Press button (*3*-back task)				5	8	3	4

Table 1. Example of the N-back task task. Participants have to confirm by a button press, that the currently displayed number matches with the number seen N numbers ago.

The motoric requirements remain constant across difficulty levels during the N-back task. This enables comparisons between different difficulties as working memory load is measured while excluding reactions from external stimuli.

A smartphone app[4] is used to display a single matching N-back task. Throughout the experiment, we use a Nexus 5X[5] with a screen size of 5.2 inches to run the N-back trials. The displayed symbols ranged between 0 and 9, which appeared in a random order. Each number is displayed in the center of the smartphone screen for one second. Afterward, the screen remains blank for 2.5 seconds before the next number appears. Participants have to press a button on the screen if a match occurs.

5 STUDY: EEG AS INDICATOR FOR WORKLOAD

In the following study, we assess the validity of our setup by conducting two N-back tasks with different complexities as described before. Afterwards, we conduct two assembly tasks per participant, each with the two previously mentioned assembly instruction systems. The overall hypothesis is that *in-situ* projections induce less cognitive workload than printed paper instructions. This leads to the following hypotheses:

H1: Projected *in-situ* instructions will induce *higher alpha power*, relative to printed instructions.

H2: Projected *in-situ* instructions will produce *lower scores of subjective workload*, relative to printed instructions.

H3: Projected *in-situ* instructions will produce *fewer item selection errors*, relative to printed instructions.

H4: Projected *in-situ* instructions will produce *fewer assembly errors*, relative to printed instructions.

H5: Projected *in-situ* instructions will produce *faster completion times*, relative to printed instructions.

(a) (b)

Fig. 3. **(a):** The Emotiv Epoc wireless EEG headset featuring 16 electrodes including two reference electrodes. **(b):** Electrode placement layout of the 14 measurement electrodes [82].

[4] www.play.google.com/store/apps/details?id=cz.wie.p.nback - last access 2018-05-28
[5] www.gsmarena.com/lg_nexus_5x-7556.php - last access 2018-05-28

5.1 Methodology and Measures

We used the Emotiv Epoc as brain-sensing device during the whole experiment (see Figure 3a and Figure 3b). Since the alpha band varies between participants, we conducted an eyes opened and eyes closed task to estimate the individual alpha band [4]. The overall duration of this trial was one minute. The participants started with their eyes opened and were verbally instructed to close their eyes after 30 seconds to provoke a sudden increase in alpha power [4]. The participants kept their eyes closed for another 30 seconds. We use this peak as a reference point for extracting alpha power for later analysis. Furthermore, we took ±2 Hz around the peak frequency as a measure for the alpha band.

We continue then with the N-back task to verify the validity of our setup and accuracy of our EEG measures. Participants start with a *0*-back task to be induced with low workload, followed by a *2*-back task to be induced with high workload. The 2-back induces sufficient complexity to make the differences in working memory between resting and N-back task visible in EEG data [8]. A NASA-TLX questionnaire is filled out after each N-back condition to collect subjectively perceived workload.

We begin with the assembly task afterward. Inspired by several reference tasks proposed by previous research [24], we use a Lego Duplo task to evaluate the paper and projected *in-situ* instructions in terms of measured workload. We prepared two different assembly instructions, where each of them is modeled as paper instruction and projection as shown in Figure 2. For the assembly tasks, we used a repeated measures experimental design with the instruction visualization as single factor including the levels paper and *in-situ*. We counterbalanced the order of conditions according to the balanced Latin square. As identified in previous work [6, 12, 25, 71], we measure the number of errors and the task completion time per trial. The number of errors was divided into item selection and assembly errors. An item selection error is counted whenever participants put their hands into a box where incorrect bricks reside. An assembly error is counted when a brick is assembled in a wrong position. To reduce the number of head movements, we seated the participants before the assembly experiment at a comfortable distance to the assembly setup, so that boxes and the assembly plate can be reached with minimal effort. Participants filled out a NASA-TLX questionnaire after each trial to provide their subjectively perceived workload during the last assembly condition. Ultimately, we asked the participants verbally about their preference between both assembly instruction systems and noted their answers for later analysis.

5.2 Procedure

Participants signed a consent form and provided their demographics after we explained the course of research to them. We put the BCI on the participant's head and ensured a good connectivity between the scalp and all electrodes. We explained to the participants what EEG signals and noisy artifacts are. We asked participants to keep their head as still as possible and to avoid unnecessary eye blinks.

Participants started with a one-minute baseline resting task. Participants kept their eyes opened for 30 seconds. After the first 30 seconds elapsed, participants were instructed to close their eyes for another 30 seconds. The study continued with two N-back tasks, each comprising 20 numbers. The total runtime of each N-back was one minute and ten seconds. The experiment started with the *0*-back task. Participants had to press the match button every time they saw a number since the currently displayed numbers refer to the same number. Participants continued with the *2*-back task. The participants had to press the match button whenever the currently displayed number was equal to the number shown two numbers back. We recorded EEG data during all tasks. Participants were asked to fill out a NASA-TLX questionnaire after each N-back task.

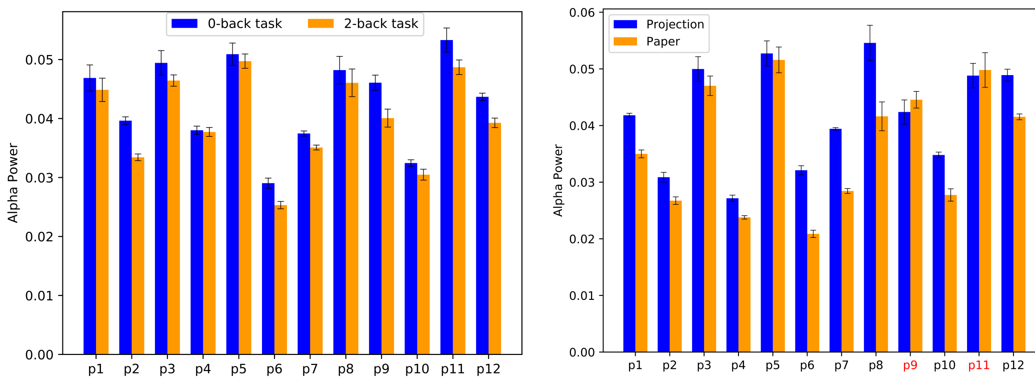


Fig. 4. **(a):** IAF power for working memory load, i.e., *0*-back, and *2*-back. The IAF power is higher for lower working memory load. The error bars depict the standard error of the sample mean. **(b):** Participant-wise comparison of IAF power when using *in-situ* projections and paper instructions. Except p9 and p11, the use of *in-situ* instructions results in higher IAF power compared to paper instructions. The error bars depict the standard error of the sample mean.

After the verification procedure, participants started with the assembly procedure. They began with either paper instructions or projected *in-situ* instructions based on the order of the balanced Latin square. Additionally, we shuffled the order of the Lego Duplo instructions itself between the conditions. During the assembly, we recorded raw EEG data, counted the number of errors separated into item selection or assembly errors, and the time used for assembly. During all conditions, participants were instructed to keep their head still and avoid unnecessary eye blinks. After every assembly, participants were asked to fill out a NASA-TLX questionnaire to assess the perceived the workload during the usage of the given instruction system. Qualitative comments about the preference of users regarding the instruction systems were collected in the end.

5.3 Data Processing

We apply the following data processing procedure: Data is filtered using a spatio-spectral decomposition method [64] with filter thresholds between 0.5 Hz and 20 Hz to remove unwanted frequencies caused by eye blinks or head movements. We average all 14 channels by calculating the element-wise mean of the signals. We remove the first and last four seconds of the signal to avoid unwanted artifacts caused by the beginning and end of the trial [51]. We divide the signal into one-second slices with an overlap of half a second. Instead of extracting the alpha band between 8 Hz and 12 Hz, we determine the maximum peak during the eyes opened and eyes closed task for each individual [19, 43]. The power spectra around the maximum peak (±2 Hz) is averaged and used as individual alpha power.

5.4 Results

We recruited twelve participants over our university mailing lists (8 male, 4 female). All participants were students and had normal or corrected-to-normal vision. None of the participants were affected by neurological disorders. The mean age was 23 years (SD = 2.22). Participants were compensated with 10 Euro for their participation.

5.4.1 Executive Working Memory Load and EEG. We identified alpha desynchronization as features that corresponded with an increased load in executive working memory. The alpha band is known to vary across individuals [15]. Therefore, we determined an individual alpha frequency (IAF) bandwidth for each participant based on their peak frequency value given their baseline EEG. We defined ±2 Hz around the peak frequency as the individual alpha band. Overall, the mean IAF band ranged between 6.5 Hz and 10.5 Hz ($SD = 2.14$).

The mean power of the IAF power for each participant was submitted to a two-tailed paired-samples t-test for the factor *executive working memory load* consisting of the *0*-back and *2*-back task. A Shapiro-Wilk test confirmed that the data was normally distributed. The results reveal a significant difference, $t(11) = 5.90$, $p < 0.001$, $Cohen's\ d = 1.70$, between the *0*-back and *2*-back task. Figure 4a shows the direct comparison of alpha activation between both conditions. To summarize, our chosen EEG feature was sensitive to load in an executive working memory task, namely there was less power in the IAF band when there was higher load in this task. Thus, we report a large effect size in support of H1.

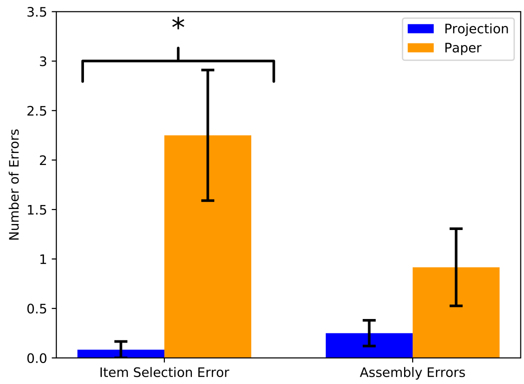
5.4.2 Assembly Performance and Alpha Power. The IAF power of each participant was submitted to a two-tailed paired-samples t-test for the factor *haptic assembly instructions* consisting of paper and *in-situ* instructions after a Shapiro-Wilk test confirmed that the data was normally distributed. The results reveal a significant difference, $t(11) = 3.86$, $p = 0.003$, $Cohen's\ d = 1.12$. Figure 4b shows the mean alpha activation per condition and participant. Similar to our findings with the executive working memory load, we find that IAF power is generally lower for paper instructions compared to when participants experienced projected *in-situ* instructions instead—note that this was not true for p9 and p11 at the individual level. Thus, we report a large effect size in support of H2. More importantly, we show that the IAF power responds for lower executive working memory load as it does for our projected *in-situ* instructions.

Additionally, we statistically compare the number of errors made by the participants during assembly as well as the time they required to finish the assembly between the different conditions. We classify errors into item selection errors and assembly errors. Overall, participants did in average 2.25 ($SD = 2.301$) item selection errors during the paper instruction condition and 0.083 ($SD = 0.289$) errors during the *in-situ* condition. 0.917 (SD = 1.379) assembly errors were performed when using paper instructions and 0.25 ($SD = 0.452$). A Shapiro-Wilk test did not show a normal distribution for the item selection and assembly errors. Thus, we have conducted a Wilcoxon signed-rank test. A significant difference was found for the number of item selection errors, $p = 0.009$. However, no significant difference was found for the number of assembly errors, $p = 0.188$. Figure 5a compares the number of item selection and assembly errors between both instruction systems.

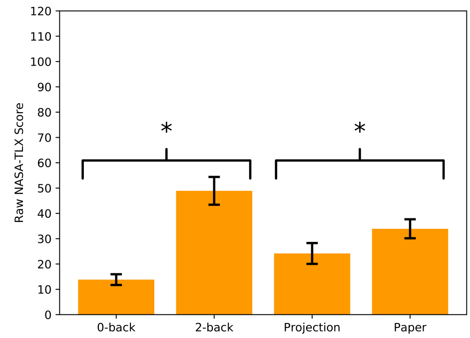
The task completion time averages to 217.08 seconds ($SD = 40.31$) for paper instructions and 124 seconds ($SD = 13$) for projected *in-situ* instructions. The task completion time shows a significant effect between both conditions regarding task completion time, $t(11) = -8.82$, $p = 0.001$, $Cohen's\ d = -2.55$.

5.4.3 Qualitative Assessment. We statistically analyze the subjectively perceived workload using the collected NASA-TLX questionnaires. We average the scores received from the six Likert scales per condition to calculate the mean raw NASA-TLX scores. The averaged NASA-TLX score amounts to 13.83 ($SD = 7.34$) for the *0*-back task, 48.92 ($SD = 19.01$) for the *2*-back task, 33.92 ($SD = 13$) for the assembly task using paper instructions, and 24.17 ($SD = 14.26$) during assembly using projected *in-situ* projections.

A Shapiro-Wilk test confirmed that the data is normally distributed. A t-test shows a statistical difference in NASA-TLX scores between the *0*-back and *2*-back condition, $t(11) = -6.41$, $p = 0.001$, $Cohen's\ d = -1.85$. Therefore, more workload was subjectively perceived in the *2*-back task than



(a)



(b)

Fig. 5. **(a):** Mean item selection and assembly errors. The error bars indicate the standard error. Brackets indicate significant differences. **(b):** Mean raw NASA-TLX scores per condition. The error bars depict the standard error. Brackets indicate significant differences.

in the *0*-back task. A significant difference between paper and *in-situ* instructions, $t(11) = -3.86$, $p = 0.003$, $Cohen's\ d = -1.12$, was found. Therefore, lower alpha activation was measured during the *in-situ* conditions compared to the printed paper instruction conditions. Figure 5b shows the mean NASA-TLX scores.

Additionally, we collected comments from the participants by asking for their preference regarding the assembly in instruction system. Most participants provided us with positive feedback, such as

> *"[...] projected instructions were easier and faster to understand than paper instructions. The light was a good guidance."* (p1, p2, p3, p7)

> *"I felt that I was faster using projections since I did not have to flip the paper. Having both hands free enhanced the overall assembly."* (p5)

> *"It was easier to follow the light than to follow the paper instructions itself. However, I felt like a robot during assembly."* (p4)

However, some participants stated, that

> *"[...] the higher assembly speed using projections was stressful. Paper instructions provided a short relieve in workload when flipping the page."* (p9, p11, p12)

or that they were

> *"[...] unused to it, but [I] could familiarize with it after some time."* (p3, p6, p9).

It is interesting to note that p9 and p11 perceived the assembly as stressful due to its fast pace. Both participants also show less alpha activation during the projected *in-situ* condition than in the paper condition (see Figure 4b). Instructions were provided immediately by the *in-situ* instruction system, which was perceived as stressful two participants (p9, p11). Paper instructions provided cognitive alleviation when flipping the page to mentally prepare for the next work step. Previous research also supports a positive correlation between stress and cognitive workload [65, 74]. The discomfort of the provided assembly instruction system can also be elicited from our results.

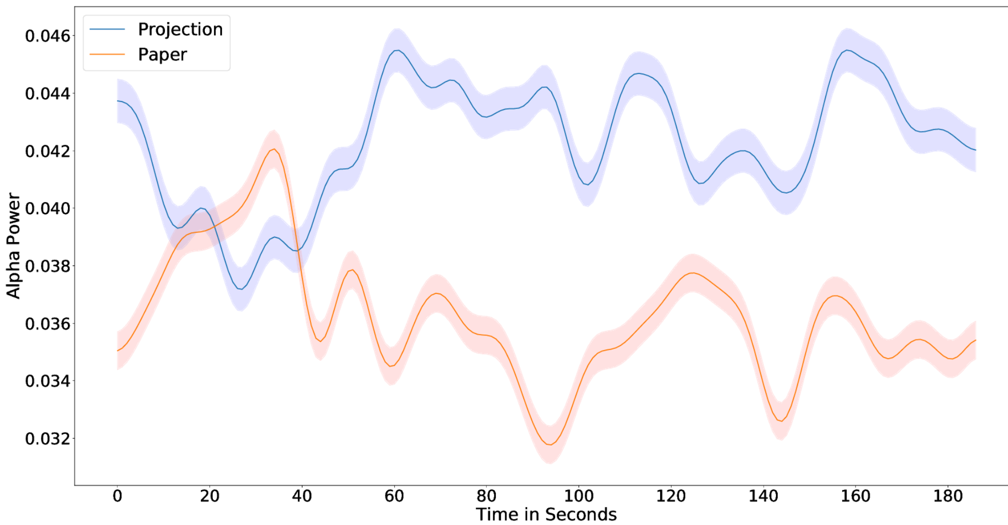


Fig. 6. Mean alpha power fluctuation of all participants over time. Alpha power increases when using projected *in-situ* instructions and decreases for paper instructions. The shadowed area describes the standard deviation of each data point per participant.

5.5 Assessing Cognitive Workload in Real-Time

Our results show a significant difference in alpha activation between paper and projected *in-situ* instructions. Using the collected data, we analyze the real-time applicability of EEG for cognitive workload assessment. We achieved this by calculating the alpha fluctuation over time for both conditions.

Similar to the previous analysis steps, we preprocess the data using the spatio-spectral decomposition method [64] with filter thresholds between 0.5 Hz and 20 Hz and calculate the mean of all channels. We remove the first and last four seconds of the signal to remove unwanted response artifacts. We calculate the individual alpha band per participant and averaged the alpha power overall participants for both conditions. Figure 6 shows the alpha fluctuation for both conditions.

For the paper instruction, there is an increase followed by a decrease of alpha activation over time for the paper instructions. Workload starts to differentiate with time when information from paper instructions have to be held continuously in the short-term memory. This could be a reason for the decrease of alpha power with time.

Projected *in-situ* instructions show a decrease followed by an increase in alpha power. We assume, that the novelty of the system for the participants is responsible for the decrease in alpha power at the beginning of the condition. Alpha power increases when the participants become familiar with the system and less information has to be kept in the short-term memory. A Shapiro-Wilk test showed a non-normal distribution. A Wilcoxon signed-rank test resulted in a statistical difference between both conditions, $p = 0.001$.

6 DISCUSSION

We evaluate the viability of a commercially available EEG device for estimating mental workload of two different instruction systems to augment a manual assembly task. We discuss the validity of our hypotheses and present a framework to evaluate interactive assistive technologies based on EEG measures.

6.1 Validating EEG Setup prior to the Experiment

To validate the correctness of our EEG setup and measures, we manipulated the difficulty of a N-back task to vary working memory load and investigated its impact on EEG measurements. Specifically, we focused on alpha power in an individually determined frequency bandwidth, proximal to 10 Hz. Previous research has consistently demonstrated lower alpha power in participants who experience high working memory load compared to low working memory load [31]. Thus, we expected lower alpha power during the *2*-back task compared to a *0*-back task. Indeed, we found lower alpha power for higher task difficulty and higher alpha power for lower task difficulty during the N-back conditions. In other words, participants experienced less cognitive workload during the *0*-back task compared to the *2*-back task. This supports the applicability of using commercial devices to infer cognitive workload and converges with the results found in previous work [33, 51].

Before evaluating novel interactive assistive systems using EEG, our findings suggest to conduct an eyes opened and eyes closed task to elicit the individual's alpha peak. Afterward, a N-back task with $N = 0$ and $N > 1$ should be conducted to infer the validity of EEG measures in terms of lower alpha activation for higher task complexities. The individual alpha peak is elicited when participants close their eyes [43] enabling an individual IAF analysis for upcoming EEG trials.

6.2 Evaluating Assembly Instruction Systems

Statistical comparisons of alpha power confirmed that projected *in-situ* instructions generated significantly higher alpha power than the use of printed paper instructions. In other words, the projected *in-situ* instructions induced less cognitive workload. Thus, H1 is statistically supported. This finding that is based on EEG measurements agrees with the subjective self-reporting measures based on NASA-TLX questionnaires [22]. A significant difference in subjectively perceived workload was found between paper and projected *in-situ* instructions (H2). In general, our results support the idea that alpha-band frequency power of EEG measurements is a valid metric for estimating the levels of cognitive workload induced by a specific instruction system or visualization. Altogether, this agrees with the motivation of *in-situ* projections, which is to reduce cognitive load by providing situated information at the appropriate times.

Regarding assembly performance, we can partially confirm the outcomes of previous research [25]. We found a significant difference in the number of item selection errors. We confirm that fewer item selection errors were observed using projected in-situ instructions compared to printed paper instructions (H3). However, no significant difference between both instructions systems was found on assembly errors. Therefore, we cannot confirm H4. Assembling with projected *in-situ* instructions takes significantly less time than printed paper instructions. This agrees with previous research [25] and supports our final hypothesis (H5).

Our results encourage to measure and evaluate EEG to assess the mental demand of assistive technologies in manual assembly processes. However, the experiment has taken place under controlled conditions in lab environments where participants were restricted to a reduced number of eye blinks and head movements. We recommend the evaluation of EEG to evaluate assistive technologies in controlled environments before they will be deployed in real-world scenarios or for further scientific research inferring the mental demand during usage. This way, novelty biases can be detected by comparing correlations between EEG measures and self-rated assessments. System architects benefit from these evaluation steps as objectively measured workload can be considered into the design pipeline.

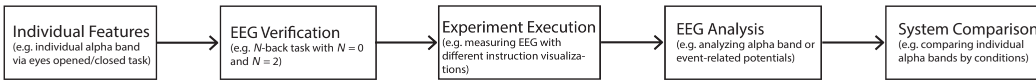


Fig. 7. Experimental protocol which enables system designers to evaluate mental demanding elements in their user interface design on a cognitive level.

6.3 EEG as Real-Time Evaluation Tool

Additionally, we investigated the applicability of EEG as a real-time evaluation tool for detecting working memory. Previous research used medical EEG systems to derive the current level of executive working memory in real-time [5, 57, 62]. We show that differences in cognitive workload are measurable for two different instruction visualization systems using a mobile consumer EEG headset.

Figure 6 shows, how the level of alpha power changes with time during assembly. The alpha power for paper instructions decreases with time as the current bin for an item selection has to be recalled with every assembly step. Furthermore, the final position of the brick has to be recognized and placed correctly according to the paper instruction.

Projected *in-situ* instructions show the opposite effect. After a decrease in alpha power, which could be attributed to a lack of familiarity with the system, an increase in alpha power is observed. This can be due to an alleviation of working memory load since projected instructions eliminate the need to maintain a set of manual assembly instructions in working memory. Projected instructions are updated in accordance with each step of manual assembly and at the appropriate time steps. Thus, unlike with the use of paper instructions, it is no longer necessary to maintain and recall instructions from working memory.

The stability in alpha fluctuation for both instruction systems support the use of commercial EEG devices for real-time workload estimation. This agrees with the subjective perception of workload through NASA-TLX questionnaires and verbal feedback provided by the participants. A real-time system for estimating cognitive load could benefit use-case scenarios such as in evaluating user interfaces, assessment of workload in safety-critical tasks, or real-time adaptation of user interfaces suited to the current level of cognitive workload. By detecting constant low or high alpha fluctuations the provided assistance can be adjusted depending on the measured workload levels. Therefore, we present an experimental protocol which evaluates the cognitive demand assistive systems require during runtime (see Figure 7). Designers can use this protocol as an assessment template to find distracting or cognitive demanding user interface elements.

Novel engineered assistive technologies can benefit from real-time insights into the mental resources by an operator. This enables user interface designers to test visualization adaptations for different measures of cognitive workload. However, the question of how to provide user interface adaptation for different levels of cognitive workload still remains and depends highly on the use case scenario in which assistive technologies are deployed.

6.4 Other EEG Metrics for Evaluating Mental Workload

We relied on an established EEG metric (i.e., alpha power) to confirm previous claims that projected *in-situ* instructions can reduce working memory. Frequency domain measures of cognitive load are well-established and proved to be viable [29, 30]. However, frequency-based measures often lack the discrimination of functional interpretations, afforded by time-domain measures. Event-related potentials (ERPs) [69] refer to the time-domain waveform that is contingent upon the occurrence of a critical event such as the presentation of a stimulus.

Unlike frequency domain measures, the time-varying EEG activity of subsequent voltage deflections can be individually attributed to functional mechanisms that underlie information processing. For example, an early negative deflection in the ERP waveform around 100–200 ms is often associated with stimulus detection while a later positive deflection between 250–600 ms is associated with stimulus recognition and working memory updating [67]. Clearly, this allows for stronger functional discriminability in the type of cognitive work that is experienced by the user, compared to the all-encompassing term 'mental workload'. In fact, signal processing and classification approaches have been proposed to utilize ERP features for mental workload classification [8, 40].

The use of ERPs for estimating mental workload is limited by the fact that ERPs are event-related. This means that ERPs can only be extracted if the event that triggers them is known beforehand. These are typically the items presented in N-back tasks. Unfortunately, not all interfaces that are of interest will consist of a N-back component. One approach around this could be to introduce task-irrelevant probes for ERPs, such as an environmentally sound. More recent research has shown that involuntary ERPs to task-irrelevant stimuli vary in their amplitudes depending on the cognitive demands of the primary task that they are engaged in [11, 59, 73]. Application cases include estimating the workload of users playing Tetris across different difficulty levels to estimating the immersion of users in high-fidelity driving simulators.

7 LIMITATIONS

We are aware of certain limitations of the study. The study was conducted in a controlled laboratory setting, whereby participants were instructed to blink as little as possible and to avoid unnecessary head movements in order to minimize artifacts in the EEG signal. Users are unlikely to abide by such instructions in mobile real-world settings. Nonetheless, we point out that the manual assembly task is one that involves substantial activity. Therefore, our EEG recordings were likely to have contained a significant degree of motoric noise and continued to be robust for our current estimation of cognitive load. Cortical activity related to the planning of motor actions can also result in lower frequency power in a bandwidth that overlaps with alpha (i.e., mu-power: 8 Hz to 12 Hz [68]). Therefore, the current results could have been affected by motor planning during item selection and placement within the assembly task, rather than cognitive load per se. Nonetheless, we assume that our results are valid for cognitive load and not motoric activity, given that we validated our EEG measure with a corresponding N-back task. Furthermore, we extracted the individual alpha band, which is not necessarily located in the mu frequency range.

8 CONCLUSION AND FUTURE WORK

In this work, we investigated whether a projected *in-situ* system for presenting instructions during a manual assembly task could serve its intended purpose in reducing cognitive load. We employed a commercial EEG headset to derive direct measurements of neural activity that varied in accordance with working memory load in a N-back task, namely alpha power [43]. This same measurement (i.e., alpha power) was larger for projected *in-situ* instructions than the traditional approach of paper instructions, demonstrating that cognitive load was lower when a projected *in-situ* system was employed. To the best of our knowledge, this is the first study that provides direct evidence for assistive technologies, and specifically in its role to alleviate working memory demands. To date, only subjective questionnaire estimates have been collected. The current work demonstrates the viability of using a commercial EEG device for evaluation purposes, even in a setting that involves a large amount of user activity (i.e., manual assembly).

In future work, we intend to investigate if functional distinctions of cognitive workload can be achieved with a commercially available EEG device. In addition, we seek to address the question of what constitutes an optimal level of cognitive workload. It has been argued that low workload could result in boredom and passive fatigue, which is just as undesirable as high workload [84]. Based on this, we plan to create a workload-aware environment where projected *in-situ* instructions can be presented just-in-time for cognitive alleviation by responding to the current actions, changing task demands, and the workload level experienced by the brain. Assessing cognitive workload in real-time using EEG will support the establishment of a physiological-based benchmark for assistive instruction systems and user interfaces. Such a benchmark will strongly contribute towards the principled and systematic evaluation of general novel engineered user interfaces, besides those employed in industrial assembly settings. To foster and encourage research in this area, we publish the dataset and used assembly instructions on our institutes' website[6].

9 ACKNOWLEDGEMENTS

This work was supported by the German Federal Ministry of Education and Research as part of the project KoBeLU (Grant No. 16SV7599K) and German Research Foundation (DFG) within the projects C02, C03 and task force TF3 of SFB/Transregio 161. Work on this project was partially funded by the Bavarian State Ministry of Education, Science and the Arts in the framework of the Centre Digitization, Bavaria (ZD.B) and by the Deutsche Forschungsgemeinschaft (DFG) (Grant No. AL 1899/2-1). Part of this research was supported by the European Union's Horizon 2020 Programme under ERCEA grant no. 683008 AMPLIFY.

REFERENCES

[1] Charles W Anderson and Zlatko Sijercic. 1996. Classification of EEG signals from four subjects during five mental tasks. In *Solving engineering problems with neural networks: proceedings of the conference on engineering applications in neural networks (EANN'96)*. Turkey, 407–414.

[2] E. W. Anderson, K. C. Potter, L. E. Matzen, J. F. Shepherd, G. A. Preston, and C. T. Silva. 2011. A User Study of Visualization Effectiveness Using EEG and Cognitive Load. *Computer Graphics Forum* 30, 3 (2011), 791–800. https://doi.org/10.1111/j.1467-8659.2011.01928.x

[3] Alan D Baddeley and Graham Hitch. 1974. Working memory. *Psychology of learning and motivation* 8 (1974), 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

[4] Robert J. Barry, Adam R. Clarke, Stuart J. Johnstone, Christopher A. Magee, and Jacqueline A. Rushby. 2007. EEG differences between eyes-closed and eyes-open resting conditions. *Clinical Neurophysiology* 118, 12 (2007), 2765 – 2773. https://doi.org/10.1016/j.clinph.2007.07.028

[5] Chris Berka, Daniel J. Levendowski, Milenko M. Cvetinovic, Miroslav M. Petrovic, Gene Davis, Michelle N. Lumicao, Vladimir T. Zivkovic, Miodrag V. Popovic, and Richard Olmstead. 2004. Real-Time Analysis of EEG Indexes of Alertness, Cognition, and Memory Acquired With a Wireless EEG Headset. *International Journal of Human–Computer Interaction* 17, 2 (2004), 151–170. https://doi.org/10.1207/s15327590ijhc1702_3

[6] Jonas Blattgerste, Benjamin Strenge, Patrick Renner, Thies Pfeiffer, and Kai Essig. 2017. Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '17)*. ACM, New York, NY, USA, 75–82. https://doi.org/10.1145/3056540.3056547

[7] Karel A Brookhuis and Dick de Waard. 2002. On the assessment of (mental) workload and other subjective qualifications. *Ergonomics* 45, 14 (2002), 1026–1030.

[8] Anne-Marie Brouwer, Maarten A Hogervorst, Jan BF Van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. 2012. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of neural engineering* 9, 4 (2012), 045008.

[9] Anne-Marie Brouwer, Maarten A Hogervorst, Jan B F van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. 2012. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering* 9, 4 (2012), 045008. http://stacks.iop.org/1741-2552/9/i=4/a=045008

[10] Andreas Bulling and Thorsten O Zander. 2014. Cognition-aware computing. *IEEE Pervasive Computing* 13, 3 (2014), 80–83. https://doi.org/10.1109/MPRV.2014.42

[11] Christopher G Burns and Stephen H Fairclough. 2015. Use of auditory event-related potentials to measure immersion during a computer game. *International Journal of Human-Computer Studies* 73 (2015), 107–114.

[6] www.hcilab.org/identifying_cognitive_assistance_data_set - last access 2018-05-28

[12] Sebastian Büttner, Henrik Mucha, Markus Funk, Thomas Kosch, Mario Aehnelt, Sebastian Robert, and Carsten Röcker. 2017. The Design Space of Augmented and Virtual Reality Applications for Assistive Environments in Manufacturing: A Visual Approach. In *Proceedings of the 10th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, New York, NY, USA. https://doi.org/10.1145/3056540.3076193

[13] Sebastian Büttner, Oliver Sand, and Carsten Röcker. 2015. Extending the Design Space in Industrial Manufacturing Through Mobile Projection. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '15)*. ACM, New York, NY, USA, 1130–1133. https://doi.org/10.1145/2786567.2794342

[14] Tim R.H. Cutmore and Daniel A. James. 1999. Identifying and reducing noise in psychophysiological recordings. *International Journal of Psychophysiology* 32, 2 (1999), 129 – 150. https://doi.org/10.1016/S0167-8760(99)00014-8

[15] Michael Doppelmayr, Wolfgang Klimesch, Th Pachinger, and B Ripper. 1998. Individual differences in brain dynamics: important implications for the calculation of event-related band power. *Biological cybernetics* 79, 1 (1998), 49–57.

[16] Matthieu Duvinage, Thierry Castermans, Mathieu Petieau, Thomas Hoellinger, Guy Cheron, and Thierry Dutoit. 2013. Performance of the Emotiv Epoc headset for P300-based applications. *Biomedical engineering online* 12, 1 (2013), 56.

[17] Mai ElKomy, Yomna Abdelrahman, Markus Funk, Tilman Dingler, Albrecht Schmidt, and Slim Abdennadher. 2017. ABBAS: An Adaptive Bio-sensors Based Assistive System. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA. https://doi.org/10.1145/3027063.3053179

[18] Randall W Engle, Stephen W Tuholski, James E Laughlin, and Andrew RA Conway. 1999. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General* 128, 3 (1999), 309. https://doi.org/10.1037/0096-3445.128.3.309

[19] A. Fink, R.H. Grabner, C. Neuper, and A.C. Neubauer. 2005. EEG alpha band dissociation with increasing task demands. *Cognitive Brain Research* 24, 2 (2005), 252 – 259. https://doi.org/10.1016/j.cogbrainres.2005.02.002

[20] Bruce J Fisch and Rainer Spehlmann. 1999. *Fisch and Spehlmann's EEG primer: basic principles of digital and analog EEG*. Elsevier Health Sciences.

[21] C. Ailie Fraser, Tovi Grossman, and George Fitzmaurice. 2017. WeBuild: Automatically Distributing Assembly Tasks Among Collocated Workers to Improve Coordination. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1817–1830. https://doi.org/10.1145/3025453.3026036

[22] Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In *Proceedings of the 10th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, New York, NY, USA. https://doi.org/10.1145/3056540.3056548

[23] Markus Funk, Tilman Dingler, Jennifer Cooper, and Albrecht Schmidt. 2015. Stop Helping Me - I'm Bored! Why Assembly Assistance needs to be Adaptive. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM. https://doi.org/10.1145/2800835.2807942

[24] Markus Funk, Thomas Kosch, Scott W Greenwald, and Albrecht Schmidt. 2015. A benchmark for interactive augmented reality instructions for assembly tasks. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. ACM, 253–257. https://doi.org/10.1145/2836041.2836067

[25] Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive Worker Assistance: Comparing the Effects of In-situ Projection, Head-mounted Displays, Tablet, and Paper Instructions. (2016), 934–939. https://doi.org/10.1145/2971648.2971706

[26] Markus Funk, Sven Mayer, and Albrecht Schmidt. 2015. Using In-Situ Projection to Support Cognitively Impaired Workers at the Workplace. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '15)*. ACM, New York, NY, USA, 185–192. https://doi.org/10.1145/2700648.2809853

[27] Markus Funk, Alireza Sahami Shirazi, Sven Mayer, Lars Lischke, and Albrecht Schmidt. 2015. Pick from Here! - An Interactive Mobile Cart using In-Situ Projection for Order Picking. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM. https://doi.org/10.1145/2750858.2804268

[28] Markus Funk and Albrecht Schmidt. 2015. Cognitive assistance in the workplace. *IEEE Pervasive Computing* 14, 3 (2015), 53–55. https://doi.org/10.1109/MPRV.2015.53

[29] Alan Gevins and Michael E Smith. 2003. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science* 4, 1-2 (2003), 113–131.

[30] Alan Gevins, Michael E Smith, Harrison Leong, Linda McEvoy, Susan Whitfield, Robert Du, and Georgia Rush. 1998. Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 40, 1 (1998), 79–91.

[31] Alan Gevins, Michael E Smith, Linda McEvoy, and Daphne Yu. 1997. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral cortex* 7, 4 (1997), 374–385. https://doi.org/10.1093/cercor/7.4.374

[32] Pierre Gloor. 1969. Hans Berger on Electroencephalography. *American Journal of EEG Technology* 9, 1 (1969), 1–8.

[33] David Grimes, Desney S. Tan, Scott E. Hudson, Pradeep Shenoy, and Rajesh P.N. Rao. 2008. Feasibility and Pragmatics of Classifying Working Memory Load with an Electroencephalograph. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 835–844. https://doi.org/10.1145/1357054.1357187

[34] Eija Haapalainen, SeungJun Kim, Jodi F. Forlizzi, and Anind K. Dey. 2010. Psycho-physiological Measures for Assessing Cognitive Load. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing (UbiComp '10)*. ACM, New York, NY, USA, 301–310. https://doi.org/10.1145/1864349.1864395

[35] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[36] Suzana Herculano-Houzel. 2009. The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in human neuroscience* 3 (2009), 31.

[37] Nina Hollender, Cristian Hofmann, Michael Deneke, and Bernhard Schmitz. 2010. Integrating cognitive load theory and concepts of human-computer interaction. *Computers in Human Behavior* 26, 6 (2010), 1278 – 1288. https://doi.org/10.1016/j.chb.2010.05.031 Online Interactivity: Role of Technology in Behavior Change.

[38] Ole Jensen, Jack Gelfand, John Kounios, and John E. Lisman. 2002. Oscillations in the Alpha Band (9-12 Hz) Increase with Memory Load during Retention in a Short-term Memory Task. *Cerebral Cortex* 12, 8 (2002), 877. https://doi.org/10.1093/cercor/12.8.877

[39] Ted J Kaptchuk. 2003. Effect of interpretive bias on research evidence. *Bmj* 326, 7404 (2003), 1453–1455.

[40] Ivo Käthner, Selina C Wriessnegger, Gernot R Müller-Putz, Andrea Kübler, and Sebastian Halder. 2014. Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biological psychology* 102 (2014), 118–129.

[41] Z. A. Keirn and J. I. Aunon. 1990. A new mode of communication between man and his surroundings. *IEEE Transactions on Biomedical Engineering* 37, 12 (Dec 1990), 1209–1214. https://doi.org/10.1109/10.64464

[42] Yoshifumi Kitamura, Yoshihisa Yamaguchi, Imamizu Hiroshi, Fumio Kishino, and Mitsuo Kawato. 2003. Things Happening in the Brain While Humans Learn to Use New Tools. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 417–424. https://doi.org/10.1145/642611.642684

[43] W. Klimesch, H. Schimke, and G. Pfurtscheller. 1993. Alpha frequency, cognitive load and memory performance. *Brain Topography* 5, 3 (1993), 241–251. https://doi.org/10.1007/BF01128991

[44] Thomas Kosch, Yomna Abdelrahman, Markus Funk, and Albrecht Schmidt. 2017. One Size does not Fit All - Challenges of Providing Interactive Worker Assistance in Industrial Settings. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (2017), 6. https://doi.org/10.1145/3123024.3124395

[45] Thomas Kosch, Mariam Hassib, Daniel Buschek, and Albrecht Schmidt. 2018. Look into My Eyes: Using Pupil Dilation to Estimate Mental Workload for Task Complexity Adaptation. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. ACM, New York, NY, USA, Article LBW617, 6 pages. https://doi.org/10.1145/3170427.3188643

[46] Thomas Kosch, Mariam Hassib, and Albrecht Schmidt. 2016. The Brain Matters: A 3D Real-Time Visualization to Examine Brain Source Activation Leveraging Neurofeedback. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1570–1576. https://doi.org/10.1145/2851581.2892484

[47] Thomas Kosch, Mariam Hassib, Paweł W. Woźniak, Daniel Buschek, and Florian Alt. 2018. Your Eyes Tell: Leveraging Smooth Pursuit for Assessing Cognitive Workload. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 436, 13 pages. https://doi.org/10.1145/3173574.3174010

[48] Thomas Kosch, Romina Kettner, Markus Funk, and Albrecht Schmidt. 2016. Comparing Tactile, Auditory, and Visual Assembly Error-Feedback for Workers with Cognitive Impairments. In *Proceedings of the 18th international ACM SIGACCESS conference on Computers and accessibility*. ACM. https://doi.org/10.1145/2982142.2982157

[49] Olave E. Krigolson, Chad C. Williams, Angela Norton, Cameron D. Hassall, and Francisco L. Colino. 2017. Choosing MUSE: Validation of a Low-Cost, Portable EEG System for ERP Research. *Frontiers in Neuroscience* 11 (2017), 109. https://doi.org/10.3389/fnins.2017.00109

[50] III Lawrence J. Prinzel, Pope Alan T., Freeman Frederick G., Scerbo Mark W., and Mikulka Peter J. 2001. *Empirical Analysis of EEG and ERPs for Pyschophysiological Adaptive Task Allocation.* Technical Report.

[51] Johnny Chung Lee and Desney S. Tan. 2006. Using a Low-cost Electroencephalograph for Task Classification in HCI Research. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST '06)*. ACM, New York, NY, USA, 81–90. https://doi.org/10.1145/1166253.1166268

[52] Fabien Lotte. 2014. *A Tutorial on EEG Signal-processing Techniques for Mental-state Recognition in Brain–Computer Interfaces.* Springer London, London, 133–161. https://doi.org/10.1007/978-1-4471-6584-2_7

[53] F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger. 2018. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering* 15, 3 (2018), 031005. http://stacks.iop.org/1741-2552/15/i=3/a=031005

[54] F Lotte, M Congedo, A Lecuyer, F Lamarche, and B Arnaldi. 2007. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering* 4, 2 (2007), R1. http://stacks.iop.org/1741-2552/4/i=2/a=R01

[55] Jaakko Malmivuo and Robert Plonsey. 1995. *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields.* Oxford University Press, USA.

[56] S. G. Mason and G. E. Birch. 2003. A general framework for brain-computer interface design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11, 1 (March 2003), 70–85. https://doi.org/10.1109/TNSRE.2003.810426

[57] R. Matthews, P. J. Turner, N. J. McDonald, K. Ermolaev, T. M. Manus, R. A. Shelby, and M. Steindorf. 2008. Real time workload classification from an ambulatory wireless EEG system using hybrid EEG electrodes. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 5871–5875. https://doi.org/10.1109/IEMBS.2008.4650550

[58] Bruce Mehler, Bryan Reimer, and Jeffery A Dusek. 2011. MIT AgeLab delayed digit recall task (n-back). *Cambridge, MA: Massachusetts Institute of Technology* (2011).

[59] Matthew W Miller, Jeremy C Rietschel, Craig G McDonald, and Bradley D Hatfield. 2011. A novel approach to the physiological measurement of mental workload. *International Journal of Psychophysiology* 80, 1 (2011), 75–78.

[60] Christian Mühl, Camille Jeunet, and Fabien Lotte. 2014. EEG-based workload estimation across affective contexts. *Frontiers in neuroscience* 8 (2014). https://doi.org/10.3389/fnins.2014.00114

[61] Tim R Mullen, Christian AE Kothe, Yu Mike Chi, Alejandro Ojeda, Trevor Kerth, Scott Makeig, Tzyy-Ping Jung, and Gert Cauwenberghs. 2015. Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Transactions on Biomedical Engineering* 62, 11 (2015), 2553–2567. https://doi.org/10.1109/TBME.2015.2481482

[62] Klaus-Robert Müller, Michael Tangermann, Guido Dornhege, Matthias Krauledat, Gabriel Curio, and Benjamin Blankertz. 2008. Machine learning for real-time single-trial EEG-analysis: From brain-computer interfacing to mental state monitoring. *Journal of Neuroscience Methods* 167, 1 (2008), 82 – 90. https://doi.org/10.1016/j.jneumeth.2007.09.022 Brain-Computer Interfaces (BCIs).

[63] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. 2012. Brain computer interfaces, a review. *Sensors* 12, 2 (2012), 1211–1279. https://doi.org/10.3390/s120201211

[64] Vadim V. Nikulin, Guido Nolte, and Gabriel Curio. 2011. A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition. *NeuroImage* 55, 4 (2011), 1528 – 1535. https://doi.org/10.1016/j.neuroimage.2011.01.057

[65] N. Y. L. Oei, W. T. A. M. Everaerd, B. M. Elzinga, S. van Well, and B. Bermond. 2006. Psychosocial stress impairs working memory at high loads: An association with cortisol levels and memory retrieval. *Stress* 9, 3 (2006), 133–141. https://doi.org/10.1080/10253890600965773 arXiv:http://dx.doi.org/10.1080/10253890600965773 PMID: 17035163.

[66] Adrian M. Owen, Kathryn M. McMillan, Angela R. Laird, and Ed Bullmore. 2005. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping* 25, 1 (2005), 46–59. https://doi.org/10.1002/hbm.20131

[67] R Parasuraman and J Beatty. 1980. Brain events underlying detection and recognition of weak sensory signals. *Science* 210, 4465 (1980), 80–83. https://doi.org/10.1126/science.7414324

[68] Anat Perry, Nikolaus F. Troje, and Shlomo Bentin. 2010. Exploring motor system contributions to the perception of social information: Evidence from EEG activity in the mu/alpha frequency range. *Social Neuroscience* 5, 3 (2010), 272–284. https://doi.org/10.1080/17470910903395767 arXiv:http://dx.doi.org/10.1080/17470910903395767 PMID: 20169504.

[69] Gert Pfurtscheller and FH Lopes Da Silva. 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology* 110, 11 (1999), 1842–1857.

[70] R N Roy, S Charbonnier, A Campagne, and S Bonnet. 2016. Efficient mental workload estimation using task-independent EEG features. *Journal of Neural Engineering* 13, 2 (2016), 026019. http://stacks.iop.org/1741-2552/13/i=2/a=026019

[71] Oliver Sand, Sebastian Büttner, Volker Paelke, and Carsten Röcker. 2016. *smARt.Assembly – Projection-Based Augmented Reality for Supporting Assembly Workers.* Springer International Publishing, Cham, 643–652. https://doi.org/10.1007/978-3-319-39907-2_61

[72] Christian Scharinger, Alexander Soutschek, Torsten Schubert, and Peter Gerjets. 2017. Comparison of the working memory load in n-back and working memory span tasks by means of EEG frequency band power and P300 amplitude. *Frontiers in human neuroscience* 11 (2017). https://doi.org/10.3389/fnhum.2017.00006

[73] Menja Scheer, Heinrich H Bülthoff, and Lewis L Chuang. 2016. Steering demands diminish the early-P3, late-P3 and RON components of the event-related potential of task-irrelevant environmental sounds. *Frontiers in human neuroscience* 10 (2016). https://doi.org/10.3389/fnhum.2016.00073

[74] Daniela Schoofs, Diana Preuß, and Oliver T. Wolf. 2008. Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology* 33, 5 (2008), 643 – 653. https://doi.org/10.1016/j.psyneuen.2008.02.004

[75] A. Stipacek, R.H. Grabner, C. Neuper, A. Fink, and A.C. Neubauer. 2003. Sensitivity of human EEG alpha band desynchronization to different working memory components and increasing levels of memory load. *Neuroscience Letters* 353, 3 (2003), 193 – 196. https://doi.org/10.1016/j.neulet.2003.09.044

[76] Kirill Stytsenko, Evaldas Jablonskis, and Cosima Prahm. 2011. Evaluation of consumer EEG device Emotiv EPOC. In *MEi: CogSci Conference 2011, Ljubljana*.

[77] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285. https://doi.org/10.1207/s15516709cog1202_4

[78] John Sweller. 1994. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction* 4, 4 (1994), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

[79] Anirudh Vallabhaneni, Tao Wang, and Bin He. 2005. *Brain—Computer Interface*. Springer US, Boston, MA, 85–121. https://doi.org/10.1007/0-306-48610-5_3

[80] Boris M. Velichkovsky and John Paulin Hansen. 1996. New Technological Windows into Mind: There is More in Eyes and Brains for Human-computer Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96)*. ACM, New York, NY, USA, 496–503. https://doi.org/10.1145/238386.238619

[81] John D Wells, Damon E Campbell, Joseph S Valacich, and Mauricio Featherman. 2010. The effect of perceived novelty on the adoption of information technology innovations: a risk/reward perspective. *Decision Sciences* 41, 4 (2010), 813–843. https://doi.org/10.1111/j.1540-5915.2010.00292.x

[82] Brenda Wiederhold and Giuseppe Riva. 2013. *Annual Review of CyberTherapy and Telemedicine: Positive Technology and Health Engagement for Healthy Living and Active Ageing.*

[83] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology* 113, 6 (2002), 767 – 791. https://doi.org/10.1016/S1388-2457(02)00057-3

[84] Mark S Young and Neville A Stanton. 2002. Malleable attentional resources theory: a new explanation for the effects of mental underload on performance. *Human factors* 44, 3 (2002), 365–375.

[85] Thorsten O. Zander, Lena M. Andreessen, Angela Berg, Maurice Bleuel, Juliane Pawlitzki, Lars Zawallich, Laurens R. Krol, and Klaus Gramann. 2017. Evaluation of a Dry EEG System for Application of Passive Brain-Computer Interfaces in Autonomous Driving. *Frontiers in Human Neuroscience* 11 (2017), 78. https://doi.org/10.3389/fnhum.2017.00078

[86] Ferdinand Rudolf Hendrikus Zijlstra. 1993. Efficiency in work behaviour: A design approach for modern tools. (1993).

Received February 2018; revised April 2018; accepted June 2018