Sensorimotor Processing in the Human Brain and in Cognitive Architectures

Dissertation
zur Erlangung des Grades
"Ph.D. in Cognitive Science"
im Fachbereich Humanwissenschaften der
Universität Osnabrück

vorgelegt von
Andrew Melnik

Osnabrück, 2017

Supervisor:

Prof. Dr. Peter König
University of Osnabrück, Germany

Additional supervisors:

Dr. W. David Hairston
U.S. Army Research Laboratory, USA

Prof. Dr. Daniel P. Ferris
University of Florida, Gainesville, USA

“Between stimulus and response there is a space. In that space is our power to choose our response. In our response lies our growth and our freedom.“
(Viktor E. Frankl)

Abstract

Sensorimotor processing is a critical function of the human brain with multiple cortical areas specialized for sensory recognition or motor execution. Although there has been considerable research into sensorimotor control in humans, the steps between sensory recognition and motor execution are not fully understood. This thesis investigates different aspects of sensorimotor processing in the human brain and proposes approaches to cognitive architectures. Here, I describe a series of six studies: an examination of sensorimotor processing in the human brain, evaluation of new mobile EEG systems for modern sensorimotor paradigms, investigation of a balance between memorization and active sampling of visual information in a sensorimotor task, and three studies on cognitive architectures for spatial reasoning and navigation in 2D environments.

In the first study (Chapter 2), we tested a hypothesis that some functional modules in the human brain participate in both sensory, as well as motor processing. We defined functional modules as independent components (ICs) yielded by an independent component analysis (ICA) of EEG data. We found EEG correlates of tightly coupled sensorimotor processing. Our results are compatible with frameworks like Embodied Cognition, Common Coding, and Sensorimotor Contingencies. This study also examined how different types of tasks and foci of attention influence sensorimotor processing, revealing variations in reaction time between conditions with the same set of stimuli. Observed differences in response time were based not on stimuli but higher cortical functions, which unfolds ‘top-down’ influences on sensorimotor processing.

In the second study (Chapter 3), we proposed an approach to evaluate new mobile EEG systems by means of ERP responses. A shift in the field of neuroscience to a greater appreciation of natural behavior, where participants can act and move in order to experience proprioceptive and vestibular sensations, coincided with the development of affordable and mobile EEG systems. Are the results obtained with new mobile EEG systems and research-grade EEG systems equivalent? In order to answer this important question, we developed a benchmark of six paradigms. Yielded results can be used for evaluation of new mobile EEG systems.

In the third study (Chapter 4), we investigated eye movements in a goal-directed task, and we described the balance between memorization and active sampling of visual information. A small change in a reward structure in the task altered the subjects' sensorimotor processing strategy. However, surprisingly not in the most efficient way regarding the time of performance. This result exposed the impact of behavioral aspects on sensorimotor processing. Participants' behavior significantly changed towards increased use of working memory when we introduced the price of saccades.

In the fourth study (Chapter 5), I proposed a connection between eye fixations in humans and breaking down of long term goals into a sequence of task-related episodes. Episodes represent nodes in a graphical model. The results suggest that spatial reasoning can work as a search mechanism based on a graphical model. These approaches correlated with the ideas of deictic codes and affordances.

In the fifth and sixth studies (Chapters 6-7), we proposed approaches to cognitive architectures for navigation in 2D environments. We used 2D computer-game environments, because, in comparison to robotic platforms, computer games are more convenient for training and testing of cognitive architectures. We proposed a way to deal with contextual uncertainty by embracing actions into a perceptual palette. We also proposed a cognitive architecture of an artificial agent with an egocentric view in a pathfinding task. We demonstrated an influence of inbuilt "instincts" and strategies onto the goal directed performance.

Studies presented in this thesis were motivated by the intention to extend and deepen knowledge of how sensory information transforms into motor actions. The results contributed to our understanding of sensorimotor processing in the human brain, revealed sources of variance in EEG data, showed a balance of memorization and active information sampling, and highlighted approaches to cognitive architectures.

Table of Content

Chapter 1. General Introduction

1.1 Approaches to studying sensorimotor processing in the context of the Stroop task

“Sensorimotor processing refers to a process by which sensory information or input is coupled or integrated to a related motor response in the central nervous system. This process underlies both involuntary or reflexive actions and voluntary acts” (http://registry.it.csiro.au/def/keyword/nature/subjects/_sensorimotor-processing:1). Sensorimotor processing is a critical function of the human brain with multiple cortical areas specialized for sensory recognition or motor execution. Although there has been considerable research into sensorimotor control in humans, the steps between sensory recognition and motor execution are not fully understood (Bear, Connors, & Paradiso, 2007).

Early studies, in the absence of modern brain-imaging techniques, were mostly focused on behavioral aspects, learning, and response measurements in humans and primates. In the early twentieth century, Pavlov’s studies of conditioned reflexes (Pavlov, 1927) demonstrated a learning procedure in which neutral sensory stimuli (e.g., a bell) elicited a response (e.g., salivation) similar to the one elicited by a biologically potent stimulus (e.g., food). Therefore, Pavlov’s studies divided sensorimotor responses into categories of ‘hard wired’ and learned with experience. Later, J. Ridley Stroop demonstrated an interference of different types of sensory information when presented simultaneously, which led to increases in sensorimotor processing time. In a task, when the name of a color (e.g., "blue", "green", or "red") is printed in a color that is not denoted by the name (e.g., the word "red" printed in blue ink instead of red ink), naming the color of the word takes longer and is more prone to errors than when the color of the ink matches the name of the color (Stroop, 1935). (Posner & Snyder, 1975) used the distinction between controlled and automatic processing, where color naming was a controlled process and a word reading an automatic one. (Cohen, Dunbar, & McClelland, 1990) developed a parallel distributed processing model of the Stroop effect, suggesting that “differences in practice lead to differences in the strength of processing, and this makes it possible to capture asymmetries of performance such as those observed in the Stroop task”. The model also indicates that Stroop-like effects can arise from the competition between two qualitatively similar processes, which differ only in their strength. The Stroop effect highlights that, although sensory processing is highly parallel, there are sensorimotor-processing steps where attentional resources are limited and interference of relevant and irrelevant sensory information for a motor output may occur.

Modern brain-imaging techniques (fMRI, PET) have provided important insights into sensorimotor processing in human and primate brains. Several studies indicated an involvement of the anterior cingulate cortex in the Stroop interference effect (Botvinick, Cohen, & Carter, 2004; Pardo, Pardo, Janer, & Raichle, 1990). The conflict monitoring model was proposed to explain the activity of ACC in the Stroop task (Botvinick, Braver, Barch, Carter, & Cohen, 2001). The authors suggested that the demand for control may be evaluated by monitoring for conflicts in information processing. It has also been proposed that “the ACC is a cortical region where a cognitive/motor command, coming from a different cortical region (e.g., prefrontal cortex), is being modulated and funnelled to the motor system, and this modulation takes place within distinct, motor output-specific subregions of the ACC, thus emphasizing the motor character of this region” (Paus, Petrides, Evans, & Meyer, 1993; Swick & Jovanovic, 2002; Turken & Swick, 1999). Another fMRI study on Stroop tasks proposed that the precuneus is a part of a system which is responsible for executive aspects of attentional selection for the task-relevant information. In particular, it sets a top-down bias for selecting certain types of information, e.g., color (Banich et al., 2000). The precuneus is associated with high-level behavioral correlates, as well as being a part of an attentional system (Cavanna & Trimble, 2006). Despite these studies, the neural mechanisms correlated with the Stroop effect are still debatable (Botvinick & Cohen, 2014). While fMRI studies can pinpoint brain areas important for an effect, a sequence of activations of brain areas from sensory stimulation to a motor response remains unknown.

EEG and MEG brain-imaging techniques have a good time resolution. Thus they can determine time aspects of sensorimotor processing in the human brain, and reveal a sequence of activation of brain areas from sensory recognition to motor execution. An ERP study of the temporal course of the Stroop color-word interference effect (Liotti, Woldorff, Perez, & Mayberg, 2000) found for the Incongruent vs. Congruent ERP difference wave a medial dorsal negativity between 350-500 ms post stimulus and a prolonged positivity developed between 500-800 ms post-stimulus over the left superior temporo-parietal scalp. “A possible interpretation of these results is that Stroop color-word interference first activates anterior cingulate cortex (350-500 ms post-stimulus), followed by activation of the left temporo-parietal cortex, possibly related to the need of additional processing of word meaning”. The ERP technique allowed the separation of neural correlates of cognitive control and conflict processing in the Stroop task (West, 2003). Cognitive control was associated with the occipital-parietal region and conflict processing was associated with the N450 component associated with the anterior cingulate cortex (ACC). These findings demonstrated that different neural systems support cognitive control and conflict processing (West, 2003). However, a poor spatial resolution does not allow us to fully understand coupling steps of sensorimotor processing in the human brain.

A combined fMRI - EEG study investigated the time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex (ACC) during top-down attentional control (Silton et al., 2010). Results indicated that left dorsolateral prefrontal cortex (LDLPFC) activity influenced Stroop performance via its relationship with later dorsal ACC activity. “The degree to which dACC influenced Stroop performance depended on the level of earlier LDLPFC activity. When LDLPFC activity levels were high, there was little impact of dACC activity on Stroop performance, suggesting that, when LDLPFC provides sufficient attentional control, dACC plays a smaller role in affecting overt performance” (Silton et al., 2010). These results are consistent with the temporal course hypothesis posited by the cascade-of-control model (Banich et al., 2000; M. P. Milham, Banich, Claus, & Cohen, 2003; Michael P. Milham et al., 2002; Michael P. Milham & Banich, 2005), which aims to explain performance on the Stroop task.

The development of microelectrode recording techniques unveiled neural correlates of sensorimotor processing with precise time and space localization. A microelectrode study revealed a neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey (Shadlen & Newsome, 2001). In this study, authors recorded the activity of single neurons in the posterior parietal cortex (area LIP) of two rhesus monkeys while they discriminated the direction of motion in random-dot visual stimuli. The neural activity predicted the monkey’s direction judgment, whether or not the judgment was correct. “Thus the activity of single neurons in area LIP reflects both the direction of an impending gaze shift and the quality of the sensory information that instructs such a response. The time course of the neural response suggests that LIP accumulates sensory signals relevant to the selection of a target for an eye movement” (Shadlen & Newsome, 2001). This experiment revealed that some areas in the monkey brain show an increased activity right after a sensory stimulation and shortly before a motor response, which demonstrates a complex nature of sensorimotor processing in primate brains. However, the microelectrode recording techniques are not applicable to healthy human subjects.

1.2 Sensorimotor processing under natural conditions

Current tendency to investigate sensorimotor processing in the human brain shifts experimental recordings from the laboratory chamber to natural conditions with full body movements (Klaus Gramann, Jung, Ferris, Lin, & Makeig, 2014; Mcdowell et al., 2013; Snider, Plank, Lee, & Poizner, 2013). A shift in the field of neuroscience to a greater appreciation of natural behavior is placing new demands on recording techniques. Participants have to act and move in order to experience the proprioceptive and vestibular sensations under natural conditions (Klaus Gramann et al., 2014; Mcdowell et al., 2013; Snider et al., 2013), parallel to a simultaneous recording of participants' motor actions and external events influencing cognition. Many traditional noninvasive imaging techniques that estimate brain activity, like fMRI (Scott A. Huettel; Allen W. Song; Gregory McCarthy et al., 2004) or magnetoencephalography (MEG) (Hansen, Kringelbach, & Salmelin, 2010), have mobility constraints that limit the ability for use in natural conditions. In contrast, mobile brain/body imaging ("MoBI") has the potential to provide unique insight into cognition in normal everyday behaviors (Ojeda, Bigdely-Shamlo, & Makeig, 2014). Near-infrared spectroscopy (NIRS) (Workman & Weyer, 2007) allows mobility and can provide blood oxygenation level dependent signals regarding brain function, but its temporal resolution is not close to that of EEG. The high temporal resolution of EEG makes it the prime candidate for a measurement technique of brain dynamics in modern real-world paradigms investigating cognition under natural contexts and those involving sensorimotor coupling and actions (Aspinall, Mavros, Coyne, & Roe, 2013; De Sanctis, Butler, Malcolm, & Foxe, 2014; K Gramann, Gwin, Bigdely-Shamlo, Ferris, & Makeig, 2010; Makeig, Gramann, Jung, Sejnowski, & Poizner, 2009).

A desire for making EEG recordings under natural conditions, and potential uses in the gaming and wellness industries, have triggered the development of affordable and mobile EEG systems. These systems tend to be low-budget, easy to set up, and convenient for wearing over an extended period (Hairston et al., 2014; Hairston & Lawhern, 2015). Multiple aspects of EEG systems can influence the quality of the signal. For example, conventional EEG sensors are usually based on a conductive gel, which leads to time-consuming setup and breakdown (gel removal from the hair and the electrodes afterward). Newer systems can use dry electrodes or saline-based electrodes to remove the impact of gel in the hair, but the quality of the EEG signal is usually reduced in some manner. EEG amplifiers can have varying parameters like noise performance, power consumption, signal bandwidth, and cost (Badillo, Ponomaryov, Ramos, & Igartua, 2003; Hairston, Proie, Conroy, & Nothwang, 2015; Harrison & Charles, 2003; Lin, Lee, Tseng, Wu, & Jiang, 2006). More recent developments include wireless data transition directly from a cap (M. De Vos, Kroesen, Emkes, & Debener, 2014), ultra-low power digitization, and a minimal

use of cables (Warchall et al., 2016). All of these aspects have a potential to alter EEG data. How much do these changes affect the recorded signals? Are the results obtained with new mobile EEG systems and research-grade EEG systems equivalent?

Recent studies have attempted to compare recordings obtained from different EEG systems. In one study (Gargiulo et al., 2010), investigators compared a new EEG system with dry electrodes to a standard and clinically available EEG system with wet electrodes using parallel and serial recording methods. The comparison included two experimental paradigms in which frequency domain and correlation coefficients of channel data, and ERPs were compared. Another study examined the correlation of signals from dry foam-based EEG sensors and wet EEG sensors, as well as the impedance at the sensor-skin contact during long-term EEG measurements (Liao et al., 2012). (Yeung et al., 2015) completed a comparison of foam-based and spring-loaded dry EEG electrodes using three experimental paradigms by linear correlation analyses. Additional studies have looked at the application-based performance using P300 brain-computer interfaces with dry and gel-based electrodes (M. De Vos et al., 2014; Guger, Krausz, Allison, & Edlinger, 2012). The general conclusion from these studies was that EEG data could be successfully collected using non-research grade EEG systems when taking into account the number and placement of electrodes.

With the increasing number of comparisons of EEG systems, we need a quantitative standard for comparing EEG systems (Oliveira, Schlink, Hairston, König, & Ferris, 2016b). A systematic benchmark based on a variety of paradigms testing different brain-response effects would allow for the assessment of how results vary across EEG systems (low-budget, dry electrode, research-grade). This would allow for comparison of not only systems evaluated in a particular paper but also any new EEG system (Senevirathna et al., 2016) in the future to the database of performance results already collected. Furthermore, to rightly estimate how large variability due to systems is, we have to compare it to other factors of variance, like subjects or repeated sessions for each subject-system combination.

1.3 Sensorimotor processing frameworks

A classical view conceptualizes sensorimotor processing as a series of functional modules activated in a sequence from sensory input to motor execution. For a usual lab setup with stimulus presentation and speeded response, such a view implies that all functional modules participate in sensorimotor processing only once and that functional modules related to sensory processing have no direct influence on action execution. Such a view also implies that action-related modules do not influence perceptual modules. This classic view explains many experiments on monkeys with

straightforward stimulus-response paradigms in perceptual discrimination and value-based decision tasks (Shadlen & Newsome, 2001; Sugrue, Corrado, & Newsome, 2005) and “describes how to form decisions using priors, evidence, and value to achieve certain desirable goals” (Gold & Shadlen, 2007).

However, other frameworks propose that sensory and motor processing are intimately related (Engel, Maye, Kurthen, & König, 2013; König, Wilming, Kaspar, Nagel, & Onat, 2013). Embodied cognition (Foglia & Wilson, 2013) supports the idea that not only perception guides motor actions, but motor actions also influence perception, as well as the idea that brain concepts are grounded in perception and action (Gallese & Lakoff, 2005). This framework is supported by experimental findings where action intentions influence visual processing (Bekkering & Neggers, 2002), as well as by the motor-visual attentional (Craighero, Fadiga, Rizzolatti, & Umiltà, 1999) and visuo-motor priming (Vogt, Taylor, & Hopkins, 2003) effects. The common-coding approach implies that “there are certain products of perception on the one hand and certain antecedents of action on the other that share a common representational domain” (Prinz, 1997). This proposes similarity between afferent and efferent codes instead of the separate-coding-and-translation approach (Prinz, 1992), so that afferent and efferent patterns share common codes. This view lead to the claim that “this approach is more powerful than, traditional approach to the relationship between perception and action, which invokes separate rather than common coding” (Prinz, 1997) in explaining effects like “negative asynchrony” (Aschersleben & Prinz, 1995; P. G. Vos, Mates, & van Kruysbergen, 1995) and “stimulus-response compatibility” (Fitts & Seeger, 1953; Lippa, 1996). Furthermore, sensorimotor contingency theory emphasizes the key role of action for perception. “Recognition of an object occurs, not when neural excitation due to the object arrives in some cortical area, but when we are exercising our mastery of the way the object behaves under exploratory manipulation” (O’Regan & Noe, 2001). Sensorimotor contingency serves as the theoretical foundation of sensory substitution, when visual sensation is perceived through touch (Bach-y-Rita, 1983), or learning a new sense with a sensory augmentation device (Kärcher, Fenzlaff, Hartmann, Nagel, & König, 2012; Kaspar, König, Schwandt, & König, 2014; Nagel et al., 2005). All these different frameworks support the idea that sensorimotor processing can be a more holistic process.

1.4 Sensorimotor processing and working memory

Human actions and sensorimotor processing occur at different time scales. Saccadic eye movement is an example of a sensory-motor primitive of a physical act. The next level is a primitive task, like uttering a sentence or dialing a telephone number (Dana H. Ballard, Hayhoe, Pook, & Rao, 1997; Newell, 1994). Sequential actions require a buffer and a workspace to store and operate over different features of sensory information (D H Ballard, Hayhoe, Li, & Whitehead, 1992). Short-term memory temporarily holds a limited amount of information in an accessible state, and working memory is used to plan and carry out behavior. “Working memory includes short-term memory and other processing mechanisms that help to make use of short-term memory” (Cowan, 2008). Short-term memory has a capacity and temporal decay limitations. Chunk capacity limits were estimated in different studies to be about 3-7 units (Cowan, 2001, 2008; Miller, 1956). This capacity limitation of the working memory may not be a shortage but a consequence of the cognitive architecture of the human brain that allows efficient searching, grouping, and processing of information.

A possible mechanism to deal with capacity limitations of the short-term memory is to combine multiple items into a composed unit (Miller, 1956). A study of the perceptual structure showed that stronger chess players were able to encode the position into larger perceptual chunks (Chase & Simon, 1973). In another study of reading a twisted text, subjects fixated at words only once in a normal case, and at individual letters during a reading of the reversed text (Kowler & Anton, 1987). Small saccades were necessary in the absence of familiar letter patterns. Learning of composed units allows performance of complex sensorimotor tasks.

Some authors argue that the environment itself can be treated as an extension of internal memory and the mind (Clark & Chalmers, 1998; O’Regan, 1992). Human performance in the block-copying task can be characterized as a sequence of deictic instructions based on a small number of primitive operations (identification, location, and deictic access). The use of “just-in-time” representation avoids complex memory representations and the carrying cost of the information. “The limited number of variables need only be a handicap if the entire task is to be completed from memory; in that case, the short-term memory system is overburdened. In the more natural case of performing the task with ongoing access to the visual world, the task is completed perfectly” (Dana H. Ballard et al., 1997).

Eye movements reflect a sequence of subtasks (location, identification) in the process of solving a larger cognitive goal. The complexity of visual computation can be substantially reduced by decomposing a given task into simpler identification and

location tasks, thereby making it computationally more tractable (Dana H. Ballard et al., 1997).

“Saccadic eye movements have often been thought to reflect cognitive events. Subjects use fixation as a deictic pointing device to serialize the task and allow incremental access to the immediately task-relevant information” (Dana H. Ballard et al., 1997). Observations of individual eye movements made in a block-copying-task experiment showed that subjects chose to serialize the task by adding more eye fixations than might be expected. These fixations allow subjects to postpone the acquiring of task-relevant information just before its use (Dana H. Ballard, Hayhoe, & Pelz, 1995; Dana H. Ballard et al., 1997; Pelz, 1995). Rather minimal information is retained from the immediately prior fixation. The information retained from prior saccades is determined by what is currently relevant for the task (Dana H. Ballard et al., 1997). These observations show that sensorimotor processing tends to less cognitively demanding strategies when possible, which, however, may require more actions and take longer. That could be due to the high “price” for carrying costs of short-term or working memory, for attentional load, for learning, or some other reason. So that low memory strategies are preferred.

Selectivity is important for sensorimotor actions. The attention has a selective nature and may play a role of a pointer. “The conception of working memory as the set of currently active pointers also leads to a very simple interpretation of the tradeoffs between working memory load and eye movements, in which fixation can be seen as a choice of an external rather than an internal pointer” (Dana H. Ballard et al., 1997). Therefore, limitation in working memory could be no constraint but a consequence of the cognitive architecture of the brain.

Aimed limb movements have a strong correlation with saccadic eye movements. Eyes fixate the target location before or shortly after a hand movement is initiated. There are three major phases of an aimed limb movement: movement preparation, initial impulse, and error correction (Abrams, Meyer, & Kornblum, 1990; Carlton, 1981; Elliott, Helsen, & Chua, 2001; Meyer, Abrams, Kornblum, Wright, & Smith, 1988). The movement-preparation phase begins after a decision to produce a movement but before the movement starts. In the initial-impulse phase, a rapid ballistic movement starts and the limb traverses most of the distance between that starting and final positions. In the error-correction phase, apparent discrepancies between the current position of the limb and the goal are being minimized. “Spatial planning for coordinated saccades and hand movements are dissociated at the level of processing at which online visual information is integrated with information in short-term memory” (Issen & Knill, 2012).

1.5 Sensorimotor processing in cognitive architectures

Valuable insights of sensorimotor processing principles were learned from biological systems (Peter William Battaglia, 2008; O'Regan & Noe, 2001; Prinz, 1990). However, a deeper understanding of sensorimotor processing in the human brain may be possible through integrating these principles (e.g., embodiment, sensorimotor contingency, common coding, etc.) into artificial cognitive architectures and robotics technologies. And, in the other way around, development of machine learning approaches, e.g., recurrent, convolutional neural networks, and artificial cognitive architectures, e.g., HTM (Hawkins & George, 2006), Soar (Laird, 2012), ACT-R (Anderson, 2007), CLARION (Sun, 2003), could help to reveal insights of sensorimotor processing principles in the human brain. Both, biologically and engineering inspired approaches are pursuing to build a model of efficient translation of high-dimensional sensory inputs into goal-directed complex motor outputs.

Simple cognitive architectures use a direct translation of sensory information to motor outputs. Such an approach encounters a problem of deriving "efficient representations of the environment from high-dimensional sensory inputs" (Mnih et al., 2015) and a problem of representation and execution of complex motor actions. This approach works well until a certain level of complexity of sensory input, motor output, or a task. Computer vision systems utilize different methods for object representation, feature extraction, database organization, and model indexing (Dickinson, 1999).

One approach to representing objects is to utilize conceptual spaces where points denote objects, and regions denote concepts. A conceptual space is a geometric structure that consists of many quality dimensions, which denote basic features by which concepts and objects can be compared. Examples of such dimensions are color, pitch, temperature, weight, and the three ordinary spatial dimensions (Zenker & Gärdenfors, 2015). Convex regions of a conceptual space can represent concepts, and one can calculate the center of gravity of a region to find prototypical members.

Another approach to handling high-dimensionality of sensory information is to utilize internal models (Ito, 2008) which maintain efficient representations of the environment. They can also predict sensory consequences from issued motor commands and can calculate necessary feedforward motor commands from desired outcomes (Kawato, 1999). Internal models allow for storage structured knowledge suitable for reasoning and conduct feature based calculations. Some authors argue that the environment itself can be treated as an extension of internal memory and the mind (Clark & Chalmers, 1998; O'Regan, 1992). Therefore, complexity and high-dimensionality of sensory information can be tackled by compact representations (Yankov & Keogh, 2006) of internal models.

Internal models in artificial cognitive architectures which deal with 2D or 3D environments can be based on computer-game engines, which simulate rich physics and make robust and fast inferences in complex natural scenes (P W Battaglia, Hamrick, & Tenenbaum, 2013). Such internal models allow us to process physical interactions, learn causality and dynamical properties of objects, and to interact with the environment. Cognitive architectures based on internal models require translation from a high-dimensional sensory representation to discrete states and objects. Current approaches to learning 3D structures from images encompass a combination of deep neural networks, probabilistic inference and render engines such as OpenGL (Wright Jr, Haemel, Sellers, & Lipchak, 2010).

Objects have a hierarchical structure, where some parts may play decisive roles in certain contexts. Different approaches to building a hierarchical representation have been proposed. The Learning a Hierarchy of Parts approach learns a hierarchy of spatially flexible compositions, i.e., parts, in an unsupervised, statistics-driven manner in lower, category-independent layers, while requiring minimal supervision to learn higher, categorical layers. It aims to enable recognition and detection of a large number of object categories (Fidler & Leonardis, 2007).

Only recognition of sensory information is not enough for a goal-directed behavior. Distinct representation of objects in the environment and their tokenization would endue objects and parts of the environment with affordances (Greeno, 1994; Turvey, 1992). Affordances encode relationships between actions, objects, and effects (Montesano, Lopes, Bernardino, & Santos-Victor, 2008). Affordances can be represented as graphical models, points in multidimensional spaces, etc. for which standard machine learning techniques can be applied. Digitalization into discrete states allows hierarchical representation of possible transitions between states, as well as semantic, and causal relationships. For many board and computer games, like “Chess”, “Go”, “Unblock Me”, “Brain it On!”, “Vovu”, etc. affordances are possible moves and actions in gameplay, which can be learned through observation and interaction with the gameplay.

Computer games reflect human cognitive capacities and therefore can serve as valuable benchmarks for cognitive architectures, artificial intelligence and machine learning approaches (Mnih et al., 2013). Current artificial intelligence algorithms play many games better than humans (Chess, Go, Atari games) (Silver et al., 2016). However, these are narrow AI algorithms which exploit computational power and convenience of well-defined rules of a game and therefore do not fully reflect computational principles of the human brain.

Finding a solution in a novel environment is a skill correlated with intelligent behavior. However, a novelty of a situation or a task that, e.g., people can efficiently

manage is quite finite. For example, a math test requires application of already known rules to novel examples. 2D or 3D computer games require prior knowledge of intuitive physics (Weitnauer & Ritter, 2012), consequences of motion in 2D or 3D environment, navigational and orientational skills, etc. Therefore, finding a solution in a novel environment can be seen as a search process of new connections between existing states.

Observation of dynamical and sequential changes in an environment may be used to build graphical models of possible transitions or train neural networks (Zhang, Leitner, Milford, Upcroft, & Corke, 2015), and in this way facilitate learning of affordances. However, learning by actions is required in order to integrate motor outputs into these transitions (Mnih et al., 2015). Such graphical or neural network models allow for searching for an action that translated a system from the current state to a desirable state. Therefore, self-actions are required to build a goal directed structure of possible transitions and related motor actions. Moreover, the same set of stimuli may result in different actions based on contextual and intentional states (Silver et al., 2016).

One of the well-known ways for agents to learn how to act optimally in controlled Markovian domains is Q-learning. “By trying all actions in all states repeatedly, it learns which are best overall, judged by long-term discounted reward” (Watkins & Dayan, 1992). However, in architectures which work in complex or real world environments, numbers of possible states and actions may be too big to try out all combinations. Contextual clustering and hierarchical organization may help algorithms to end up with a moderate number of goal-relevant states.

Human-like artificial cognitive architecture should make use of the hierarchical organization of states (Ravasz & Barabási, 2003; Sridharan, Wyatt, & Dearden, 2010). Hierarchical Task-Network planning (Erol, Hendler, & Nau, 1994) uses actions and states of the world. Each state of the world is represented by the set of atoms true in that state. Actions correspond to state transitions. Such organization allows a representation of a solution for any goal in only one step between the current and the desired states, which can be later split up into a chain of intermediate states. Therefore, the recursive breaking down of a goal into intermediate states will eventually bring the processing to the level of execution of atomic actions. For example, one may have a goal to attend a conference, which is a one step goal at a higher level of abstraction. Getting a flight ticket can be an intermediate goal**.** In order to buy a ticket, one should find a ticket; and in order to find a ticket, one should know the dates. Therefore, an atomic action to execute would be to click a computer-mouse button to open a file which contains dates of the conference. The approach of the hierarchical organization of states can carry out generalization and concept understanding mechanisms in the form of learned shortcuts between states at higher levels of abstraction.

Cognitive architectures have a natural limitation to adjust plans and actions according to a single goal at a time. Attentional mechanisms conduct information filtering and selection. “Attention is the behavioral and cognitive process of selectively concentrating on a discrete aspect of information, whether deemed subjective or objective, while ignoring other perceivable information” (https://en.wikipedia.org/w/index.php?title=Attention&oldid=792646136). Top-down attention fulfills active goals and intentions. Bottom-up attention helps in selecting what information will be cognitively processed further. New, moving, or intense information draws attention stronger. Attentional mechanisms are insufficiently developed in artificial cognitive architectures.

There are multiple cognitive architectures with a decent history of development. One of them is Soar (Laird, 2012). Soar’s architecture is based on goals, problem spaces, states, and operators. It has four main types of memory system: procedural memory, semantic memory, episodic memory, and perceptual memory. Soar responds to impasses by a generation of substates and recursive problem-space search. An agent in Soar can play, for example, Infinite Mario game (Mohan & Laird, 2011). The cognitive architecture was applied in many domains, e.g., game-based environment simulations (Quake III, Unreal 3 Atari 2600 emulator, etc.) simulated pilots, robotics. Another cognitive architecture with over 30 years of history is ACT-R. In this architecture, cognition emerges through the interaction of a number of independent modules (Anderson, 2007). ACT-R has applications in cognitive psychology, human-computer interaction, education, and computer-generated forces. One more well-known cognitive architecture is CLARION which consists of a number of distinct subsystems: action-centered subsystem, non-action-centered subsystem, motivational subsystem, and meta-cognitive subsystem (Sun, 2003). Each subsystem has implicit and explicit representations. Different cognitive architectures share many computational structures, however, they have different emphasis, like on general intelligence, cognitive modeling, or simulation of creativity.

1.6 About this thesis

EEG Correlates Of Sensorimotor Processing: Independent Components Involved In Sensory And Motor Processing (Chapter 2)

Sensorimotor processing is one of the key functionalities of the human brain, and we know of cortical areas which are specialized for sensory recognition or motor execution. But unlike early afferent and late efferent brain activities, the coupling between these steps remain concealed. In the previous sections, I described sensorimotor frameworks which are attempting to explain sensorimotor coupling in the human brain. A classical view conceptualizes sensorimotor processing as a series of functional modules activated in a sequence from sensory input to motor execution. Other frameworks propose that sensory and motor processing are intimately related.

In our experimental paradigm, we used stimulus presentation in complex speeded categorization-response tasks (variations of a Stroop task), which require recognition, decision, and motor response, to test the hypothesis that some functional modules are participating in both sensory as well as motor processing. We operationalize functional modules as independent components (ICs) yielded by an independent component analysis (ICA) of EEG data. We consistently found across subjects ICs with event-related ITC responses related to both sensory stimulation and motor response onsets. Thus, the current study reveals some EEG correlates of tightly coupled sensorimotor processing in the human brain. Our results are compatible with such frameworks like embodied cognition, common coding, and sensorimotor contingency.

Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data? (Chapter 3)

Contemporary research of sensorimotor processing in the human brain shifts experimental paradigms from an enabling environment of laboratory chamber to natural conditions with full body movements. The shift in the field of neuroscience to a greater appreciation of natural behavior is placing new demands on recording techniques. Participants have to act and move in order to experience the proprioceptive and vestibular sensations under natural conditions parallel to a simultaneous recording of participants' motor actions and external events influencing cognition.

The high temporal resolution and mobility of some new EEG systems make it the prime candidate for a measurement technique of brain dynamics in modern real-world paradigms investigating cognition under natural contexts and those involving sensorimotor coupling and actions.

It is often assumed that the specific choice of EEG system has limited impact on the data and does not add variance to the results. However, many low cost and mobile EEG systems are now available, and there is some doubt as to the how EEG data vary across these newer systems. We sought to determine how variance across systems compares to variance across subjects or repeated sessions. We tested four EEG systems: two standard research-grade systems, one system designed for mobile use with dry electrodes, and an affordable mobile system with a lower channel count. We recorded four subjects three times with each of the four EEG systems. This setup allowed us to assess the influence of all three factors on the variance of data. The two standard research-grade EEG systems had no significantly different means from each other across all paradigms. However, the two other EEG systems demonstrated different mean values from one or both of the two standard research-grade EEG systems in at least half of the paradigms. For systems with a lower number of electrodes, their spatial distributions are not suitable for all experiments. Users should be aware of the shortcomings and use such system accordingly. In this study, in addition to providing specific estimates of the variability across EEG systems, subjects, and repeated sessions, we also proposed a benchmark to evaluate new mobile EEG systems by means of ERP responses.

The Price of Saccades (Chapter 4)

Visual working memory has a limited capacity of 3-7 items (Luck & Vogel, 1997; Cowan, 2001; Todd & Marois, 2004) and temporal decay limitations. Humans tend to not reach these limits and choose to serialize a task by adding more eye fixations. These fixations allow participants to postpone the acquiring of task-relevant information just before its use.

In this study, we tested how a price for a certain type of saccades influenced eye movements in a block-copying task (Dana H. Ballard et al., 1997). We introduced a 700 ms delay for uncovering of a model of blocks when subjects wanted to get relevant information. Our results demonstrated a decrease in the penalize type of eye movements. However, that shifted an eye-movement behavior to a less optimal strategy in comparison to a case where subjects would tolerate the penalty and maintain the unpenalized eye-movement behavior. This new eye-movements strategy resulted in a longer performance time. This finding indicates the price of a higher cognitive load of short-term and working memories. Sensorimotor processing

tends to less cognitive demanding strategies when possible, which, however may take longer.

Sensorimotor processing in cognitive architectures (Chapters 5, 6, 7)

The most comprehensive understanding of how a mechanism works (e.g. sensorimotor processing) comes during a process of creation of the instance. Cognitive architectures summarize knowledge from different branches of cognitive science into a formalized theory of the human brain which can be implemented and executed as a computer program. Well-known examples of cognitive architectures are ACT-R (Anderson, 2007), Soar (Laird, 2012), CLARION (Sun, 2003), HTM (Hawkins & George, 2006).

The motivation for these studies (Chapters 5, 6, 7) was a deeper understanding and implementation of spatial affordances in tasks which require intuitive physics knowledge and spatial reasoning. These types of a task are currently problematic for robotics, machine learning, and artificial intelligence approaches (Weitnauer & Ritter, 2012). However, humans can intuitively perform such tasks. “When affordances are perceptible, they offer a direct link between perception and action. Complex actions can be understood in terms of groups of affordances that are sequential in time or nested in space” (Gaver, 1991). I implemented narrow cognitive architectures for navigation in 2D environments and spatial reasoning.

Summary

In summary, these studies investigated different aspects of sensorimotor processing in the human brain and in cognitive architectures. The first study (Chapter 2) revealed EEG correlates of sensorimotor processing in the human brain. The second study (Chapter 3) determined how variance across systems compares to variance across subjects or repeated sessions. This study allowed assessing new mobile EEG systems suitable for recordings under natural conditions with full body movements investigating sensorimotor coupling and actions. In the third study (Chapter 4), we tested how a penalty for a certain type of saccades influenced eye movements in a block-copying task. In the fourth, fifth, and sixth studies (Chapters 5, 6, 7), I implemented several narrow cognitive architectures for navigation in 2D environments and spatial reasoning.

Chapter 2. EEG Correlates Of Sensorimotor Processing: Independent Components Involved In Sensory And Motor Processing

This study has been published in “Nature Scientific Reports”.

Melnik, A., Hairston, W. D., Ferris, D. P., & König, P. (2017). EEG correlates of sensorimotor processing: independent components involved in sensory and motor processing. Scientific Reports, 7.
http://dx.doi.org/10.1038/s41598-017-04757-8

AUTHOR CONTRIBUTIONS
A.M. and P.K. conceived and designed the experiments. A.M. performed the experiments. A.M. analysed the data. A.M. and P.K. wrote the manuscript. A.M., P.K., W.D.H., and D.P.F. revised the manuscript for important intellectual content. All authors reviewed the manuscript.

2.1 Abstract

Sensorimotor processing is a critical function of the human brain with multiple cortical areas specialised for sensory recognition or motor execution. Although there has been considerable research into sensorimotor control in humans, the steps between sensory recognition and motor execution are not fully understood. To provide insight into brain areas responsible for sensorimotor computation, we used complex categorization-response tasks (variations of a Stroop task requiring recognition, decision-making, and motor responses) to test the hypothesis that some functional modules are participating in both sensory as well as motor processing. We operationalize functional modules as independent components (ICs) yielded by an independent component analysis (ICA) of EEG data and measured event-related responses by means of inter-trial coherence (ITC). Our results consistently found ICs with event-related ITC responses related to both sensory stimulation and motor response onsets (on average 5.8 ICs per session). These findings reveal EEG correlates of tightly coupled sensorimotor processing in the human brain, and support frameworks like embodied cognition, common coding, and sensorimotor contingency that do not sequentially separate sensory and motor brain processes.

2.2 Introduction

In everyday life, humans respond to a variety of stimuli to produce a range of motor commands in diverse contexts. Object manipulation, object avoidance, social interaction, and navigation are all examples of daily sensorimotor processing. Complex speeded-response tasks, which often take about 500 ms in response latency (Fize et al., 2000), provide an opportunity to trace activation of brain areas involved in sensorimotor processing using brain imaging techniques such as electroencephalography (EEG) or magnetoencephalography (MEG). These two brain imaging techniques, in particular, are helpful for research on human sensorimotor processing as they are noninvasive yet provide fine temporal resolution. Past studies have identified cortical areas characterised as being specialised for sensory recognition or motor execution (Bear et al., 2007), but we know less about the brain processes involved in moving from sensor recognition to motor execution. From a theoretical perspective of sensorimotor control, it would be very beneficial to identify areas (or modules) specifically involved in turning sensory information into motor actions during sensorimotor processing.

A classical view conceptualises sensorimotor processing as a series of functional modules activated in a sequence from sensory input to motor execution. For a usual lab setup with stimulus presentation and response, such a view implies that all functional modules participate in sensorimotor processing only once. Presumably, functional modules related to sensory processing do not have a direct influence on action execution. Such a classical view also implies that action-related modules do not influence perception. This classical view explains many experiments on monkeys with straightforward stimulus-response paradigms in perceptual discrimination and value-based decision tasks (Shadlen & Newsome, 2001; Sugrue et al., 2005), and provides a framework that "describes how to form decisions using priors, evidence, and value to achieve certain desirable goals" (Gold & Shadlen, 2007).

Other views of human brain function propose that sensory and motor processing are intimately related (Engel et al., 2013; König et al., 2013). Embodied cognition (Foglia & Wilson, 2013) supports the ideas that perception guides motor actions, motor actions influence perception, and brain concepts are fundamentally grounded in both perception and action (Gallese & Lakoff, 2005). This framework is supported by experimental findings of action intentions influencing visual processing (Bekkering & Neggers, 2002), motor-visual attentional priming (Craighero et al., 1999), and visuo-motor priming (Vogt et al., 2003). The common-coding approach implies that "there are certain products of perception on the one hand and certain antecedents of action on the other that share a common representational domain" (Prinz, 1997). This view proposes similarity between afferent and efferent codes instead of a separate-coding-and-translation approach (Prinz, 1992). Some have claimed that the approach is more powerful than the traditional approach to the perception and action relationship, which invokes separate coding rather than common coding (Prinz, 1997) in explaining effects like negative asynchrony (Aschersleben & Prinz, 1995; P. G. Vos et al., 1995) and stimulus-response compatibility (Fitts & Seeger, 1953; Lippa, 1996). Lastly, sensorimotor contingency theory emphasises the key role of action for perception. "Recognition of an object occurs, not when neural excitation due to the object arrives in some cortical area, but when we are exercising our mastery of the way the object behaves under exploratory manipulation" (O'Regan & Noe, 2001). Sensorimotor contingency serves as the theoretical foundation of sensory substitution, when visual sensation is perceived through touch (Bach-y-Rita & W. Kercel, 2003) or learning a new sense with a sensory augmentation device (Kärcher et al., 2012; Kaspar et al., 2014; Nagel et al., 2005). All three of these different frameworks support the idea that sensorimotor processing can be a holistic process rather than a series of separate modules.

In the present study, we examined the neural substrates of sensorimotor processing in humans using a complex cognitive task and speeded-categorization response. We investigated whether we could consistently find functional modules participating in both sensory processing and motor processing**.** We operationalized functional

modules as independent components (ICs) yielded by an independent component analysis (ICA) of EEG data, resulting in decomposition of many electrode channels recorded in parallel into presumed independent sources of brain activity (Bell & Sejnowski, 1995; Delorme & Makeig, 2004; Makeig, J. Bell., Jung, & Sejnowski, 1996). The classical view in its pure form leads to the hypothesis that activity of functional modules is strictly related to either sensory stimulation or motor response, but not to both. The alternative frameworks propose that some functional modules will relate to both sensory stimulation and motor response. These modules would likely participate for the whole sensorimotor process, from sensory recognition to motor response. In this study, we tested these competing hypotheses to provide insight into the theoretical constructs of sequential processing and the alternative frameworks (embodied cognition, common coding, and sensorimotor contingency).

The current study revealed demonstrable EEG correlates of sensorimotor processing in the human brain. Our experimental paradigm used stimulus presentation and speeded categorization-response tasks which required recognition, decision, and motor response. A stimulus consisted of three sources of information about colour which appeared simultaneously: the written name of a colour, the hue of the font, and the pronounced word. The tasks were either to detect a specific colour or to compare different sources of information about colour and then to respond by pressing the appropriate key. We measured event-related responses by means of inter-trial coherence (ITC) (Makeig, Debener, Onton, & Delorme, 2004; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). We consistently found independent components with event-related ITC responses related to both sensory stimulation and motor response onsets. We refer to such independent components as Sensorimotor ICs. In order to check whether the number of channels in an EEG system influences a number of Sensorimotor ICs, we recorded subjects with both 64- and 127- channel EEG systems, and additionally analysed a subset of 32 channels (from the 64-channel EEG recordings). We found that the number of channels does not influence the number of Sensorimotor ICs. We conclude that EEG data support frameworks of sensorimotor processing that directly couple sensory recognition and motor responses rather than sequentially separate them.

2.3 Results

Results general introduction

We recorded healthy students in the EEG experiment while they were performing speeded categorization-response tasks (variations of a Stroop task). Subjects completed motor responses by pressing one of two buttons on a keyboard. EEG data were filtered with a 3–45 Hz bandpass finite impulse response filter. After elimination of noisy periods, we processed EEG data by ICA using EEGLAB's 'runica' function (Bell & Sejnowski, 1995; Delorme & Makeig, 2004; Makeig et al., 1996). Further data analysis and results were based on the resulting ICs and their activity. In order to measure sensory and motor related activity of ICs, we calculated the ITC of all trials in a recording session twice per independent component. The first ITC calculation was done with stimulus alignment of trials, and the second ITC calculation was done with button-press alignment of trials. The maximum value of ITC in a sensory time-frequency window of interest was assigned as the Sensory ITC value of the IC (see Methods and Fig. 3 for details). The maximum value of ITC in a motor time-frequency window of interest was assigned as the Motor ITC value of the IC. Thus, each IC had sensory-related and motor-related ITC values, which served as scalar measures of sensory and motor event-related responses of an IC, respectively. We used these values as indicators of EEG correlates of sensorimotor processing in the human brain. In order to investigate localization of ICs, an equivalent dipole (ED) was calculated for each IC. See the Methods section for more details about subjects, stimuli, tasks, data acquisition, and data processing.

Artefact-related ICs rejected from data analysis

We identified artefact ICs, which were presumably not of neuronal origin, using two measures. The first measure of artefact-related ICs was the residual variance (RV) of the equivalent dipole. If the residual variance of the equivalent dipole exceeded 15%, then the corresponding IC was marked as residual-variance-bad (Hammon, Makeig, Poizner, Todorov, & de Sa, 2008; Onton, Delorme, & Makeig, 2005; Wyczesany & Ligeza, 2014) and excluded from data analysis. Fig. 1B and Fig. 1D show such examples of excluded ICs. Fig. 2 also shows ICs with residual variance greater than 15% on the right of the vertical dotted line. In contrast, Fig. 1C shows an examples of IC with residual variance less than 15% that was valid for further data analysis.

The second measure of artefact-related ICs was topographical sparseness (TS). When a topography of an IC is widely distributed (Fig. 1C), then topographical

sparseness is low. When a topography of an IC is concentrated over a small area (e.g., over one electrode, as shown in Fig. 1A), then topographical sparseness is high. We assume that ICs with high topographical sparseness do not represent a source in the brain, but may arise due to a change in contact properties between an electrode and skin during the experiment, bad channels, or other electrode related artefacts. Our topographical sparseness measure was calculated for each IC using the following formula: TS = max(abs(PW-mean(PW)))/std(PW), where PW is a vector of projection weights (PW) of an IC (column of the inverse weight matrix, EEG.icawinv). We obtained the EEG.icawinv matrix using EEGLAB's 'runica' function (Bell & Sejnowski, 1995; Delorme & Makeig, 2004; Makeig et al., 1996). The max(x) function returns the largest element in x. The abs(x) function returns absolute values of elements in x. The mean(x) function returns the mean value of elements in x. The std(x) function returns the standard deviation of elements in x.

This value of topographical sparseness captures the sparseness of projection weights of an IC and is dependent on the number of EEG channels. For example, TS = 5.5 for a vector of 32 EEG channels with only one non-zero element (e.g., [1,0,0,...]) but TS = 7.9 for a vector of 64 EEG channels with only one non-zero element and TS = 11.2 for a vector of 127 EEG channels with only one non-zero element. Achieving a constant TS = 5.5, 32 EEG channels would need one non-zero element, 64 EEG channels would need two non-zero elements, and 127 EEG channels would need four non-zero elements. The size of an area covered by one electrode in an EEG system with 32 channels is approximately equal to a size of an area covered by two electrodes in an EEG system with 64 channels, and 4 electrodes in an EEG system with 127 electrodes.

ICs with the topographical sparseness value lower than 5 have a size of a contributing area exceeding one-electrode coverage in 32 channels EEG system. Therefore, we excluded from data analysis all ICs with the topographical sparseness value higher than 5. Fig. 1A and Fig. 1B show such examples of excluded ICs. Fig. 2 also shows excluded ICs over the horizontal dotted line that had topographical sparseness values higher than 5.

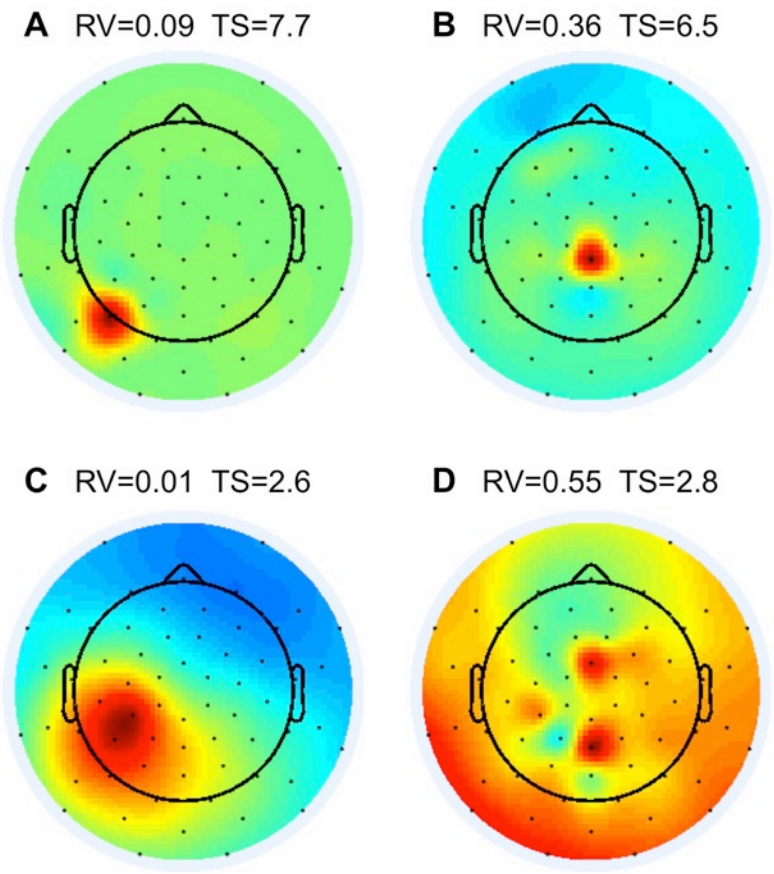


Figure 1. Examples of (A, B & D) artefact ICs which were excluded and (C) a good IC kept for further analysis. The black dots are positions of channels in the 64-channel EEG system. Colour-coding indicates projection weights of an IC (column of the inverse weight matrix, EEG.icawinv in EEGLAB). Panels A and B demonstrate ICs with topographical sparseness (TS) > 5, that were excluded from further analysis. Panels B and D demonstrate ICs with residual variance (RV) of the equivalent dipoles exceeding 15%. These ICs were also excluded from further analysis. Panel C demonstrates an IC which was kept for further analysis (TS > 5 and RV < 15%).

We obtained a total of 1412 accepted ICs (the lower left rectangle in the scatter plot in Fig. 2) and 2757 rejected ICs. The ICA method automatically produces the same number of ICs as the number of channels in EEG data (e.g., 127 ICs from 127-channel data). A systematic analysis of ICA with different numbers of EEG electrodes found that robust "electrocortical sources can be well captured using an electrode montage with as few as 35 channels" (Lau, Gwin, & Ferris, 2012). Therefore many ICs in a 127-channel collection may contain information not relevant for the experimental paradigm. From 32-channel datasets (red dots in Fig. 3), 419 ICs were accepted and 253 ICs were rejected. This was, on average, 19.95 ICs accepted per session and 12.05 ICs rejected per session. From 64-channel datasets (blue dots in Fig. 3), 529 ICs were accepted and 815 ICs were rejected. This was, on

average, 25.19 ICs accepted per session and 38.81 ICs rejected per session. From 127-channel datasets (black dots in Fig. 3), 464 ICs were accepted and 1689 ICs were rejected. This was, on average, 27.29 ICs accepted per session and 99.35 ICs rejected per session. More than half of the rejected ICs are from 127-channel datasets and less than 10% are from 32-channel datasets. This suggests that after a certain number of channels, increasing the number of channels does not result in the detection of additional independent sources of brain activity in stationary subject conditions using the experimental paradigm in this study. More unconstrained human behaviours in varying environmental conditions with substantive head motion would likely benefit from a higher number of EEG channels (Lau et al., 2012).

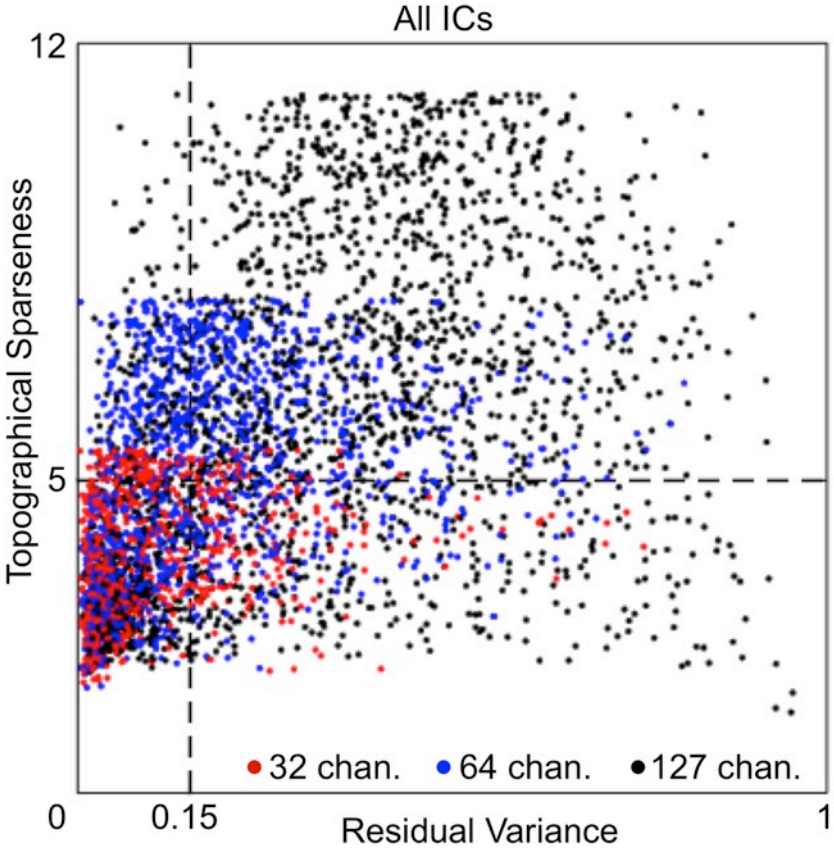


Figure 2. Scatter plot of Topographical Sparseness and Residual Variance values of all ICs in the study. Each dot represents an IC. X-axis = residual variance of an equivalent dipole of the IC. Vertical dotted line shows the residual-variance threshold = 0.15 (15%). Y-axis = topographical sparseness (TS) of an IC. Horizontal dotted line shows the TS threshold = 5. The number of all ICs from all recorded sessions in the study amounts to 4169 ICs, which consist of the following three groups: 672 ICs from 32-channel datasets (red dots), 1344 ICs from 64-channel datasets (blue dots), and 2153 ICs from 127-channel datasets (black dots). ICs in the lower left rectangle in the scatter plot were selected for further data analysis.

Measuring sensory and motor event-related responses of ICs

To address the question of whether activity of some ICs relates to both sensory stimulation and motor response, we calculated event-related responses two times for each IC in each recording session (Fig. 3). The first time was for sensory stimulus alignment of trials and the second time was for motor response alignment of trials (button press). Varying reaction time among trials in a session allowed us to differentiate stimulus and button-press event-related responses by extracting epochs either by stimulus onsets or by motor response onsets (button-press events).

In order to estimate to what extent each IC represents sensory and/or motor processing, we calculated ITC of trials epoched separately by sensory and motor onsets. Fig. 3A shows an example of an IC with clear event-related responses for both sensory and motor alignment of trials. Aligning the approximately 5000 trials of a recording session to the stimulus onset revealed a sensory event-related ITC response. The maximal value in the time-frequency window of interest of ITC (see details in Methods) served as a scalar measure of the sensory event-related response of the IC. In the example in Fig. 3A, this Sensory ITC value was equal to 0.4. Likewise, aligning the same trials on motor onset revealed a motor event-related ITC response. The maximal value in the time-frequency window of interest of ITC served as a scalar measure of the motor event-related response of the IC. In the example in Fig. 3A, the Motor ITC value was equal to 0.38. Fig. 3B shows an example of an IC with motor event-related response only. Fig. 3C shows an example of an IC with sensory event-related response only. Fig. 3D shows an example of an IC with neither sensory nor motor event-related responses. Thus, each IC in the study had two ITC values (sensory and motor) which identified the extent that the IC represented sensory or motor processing.

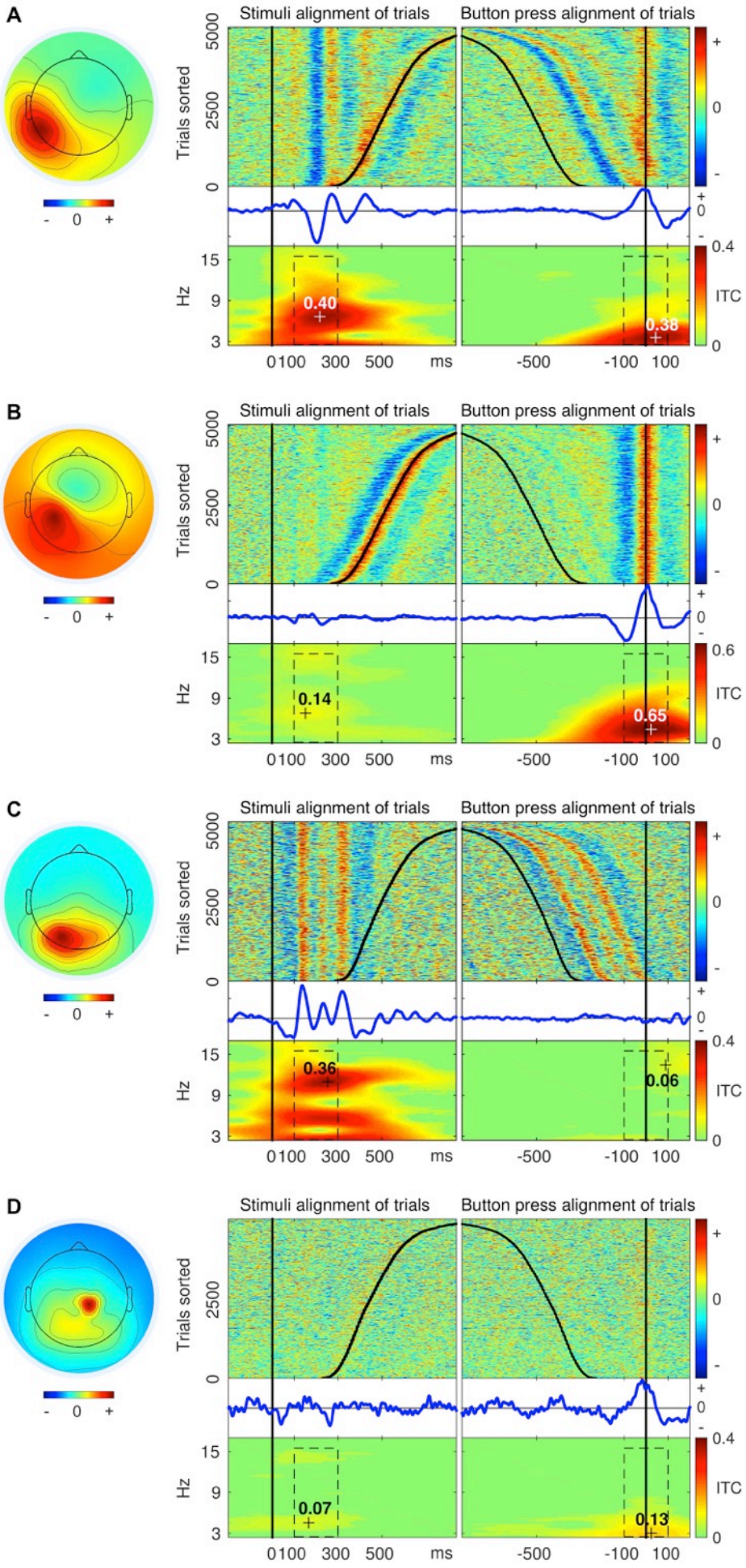

A
B
C
D
Stimuli alignment of trials
Button press alignment of trials
Trials sorted
5000
2500
0
Hz
15
9
3
0100 300 500 ms
-500 -100 100
ITC
0.40
0.38
0.14
0.65
0.36
0.06
0.07
0.13

Figure 3. Examples of (A) Sensorimotor, (B) Motor, (C) Sensory, and (D) Unspecified ICs. The exact procedure of grouping of all ICs into the four groups is described below. The round component scalp maps on the left side represent topography of ICs from 64-channel EEG recordings. The component scalp map values returned by ICA were proportional to μV (scaling is distributed between the component maps and activity time courses) (Delorme & Makeig, 2004; Makeig & Onton, 2011). Panels on the right side of each scalp map depict activity data of the IC. (1) The "top-left" plot in each panel shows colour-coded amplitude of the IC activity in a recording session. Each line in the plot represents a trial. Trials were sorted according to latency of reaction time. The black vertical line (OX = 0 ms) shows the onsets of the stimuli and the black curve in the positive direction shows the moments of a button-press event. (2) The "top-right" plot in each panel depicts the same trials as the "top-left" plot, but this time trials were aligned by onset of a button-press event (OX = 0 ms). The straight vertical line shows the moments of the button-press events and the curve in the negative direction shows the onsets of the stimuli. (3) The "middle-left" and (4) "middle-right" plots in each panel show Event Related Potentials (blue curves) derived from trials depicted in the plots above ("top-left" and "top-right"). (5) The "bottom-left" plot in each panel shows inter-trial coherence (ITC) of trials from the "top-left" plot. The black dashed rectangle represents the time-frequency window of 100 ms to 300 ms by 3 Hz to 15 Hz. The maximum value of ITC in this time-frequency window represents Sensory ITC value of the IC. (6) The "bottom-right" plots in each panel show ITC of trials from the "top-right" plot. The black dashed rectangle represents the time-frequency window of -100 ms to 100 ms by 3 Hz to 15 Hz. The maximum value of ITC in this time-frequency window represents a Motor ITC value of the IC.

Four groups of ICs: Motor, Sensory, Sensorimotor, and Unspecified

By establishing a threshold for Sensory and Motor ITC values, we separated all ICs in the study into four groups (Fig. 4A). Those ICs with Sensory and Motor ITC values exceeding the threshold (red dots in Fig. 4A; also Fig. 3A) were identified as Sensorimotor ICs. Those ICs with only Motor ITC value exceeding the threshold (blue dots in Fig. 4A; also Fig. 3B) were identified as Motor ICs. Those ICs with only Sensory ITC value exceeding the threshold (green dots in Fig. 4A; also Fig. 3C) were identified as Sensory ICs. ICs with neither Sensory nor Motor ITC values exceeding the threshold (black dots in Fig. 4A; also Fig. 3D) were identified as Unspecified ICs.

For classification of ICs into the four groups and further localization of Sensorimotor ICs, we had to select an ITC threshold. At sensory and motor ITC thresholds of 0.2, we observed a maximal number of sensory ICs (green lines in Fig. 4B) and a near-

maximal number of motor ICs (blue lines in Fig. 4B). Therefore, we selected this value for classification of ICs into Motor, Sensory, Sensorimotor, or Unspecified (Fig. 4). With this choice of critical threshold values, we observed on average 5.8 Sensorimotor ICs in each subject-session combination, which was in between the number of Sensory and Motor ICs (i.e. a substantial portion of the total).

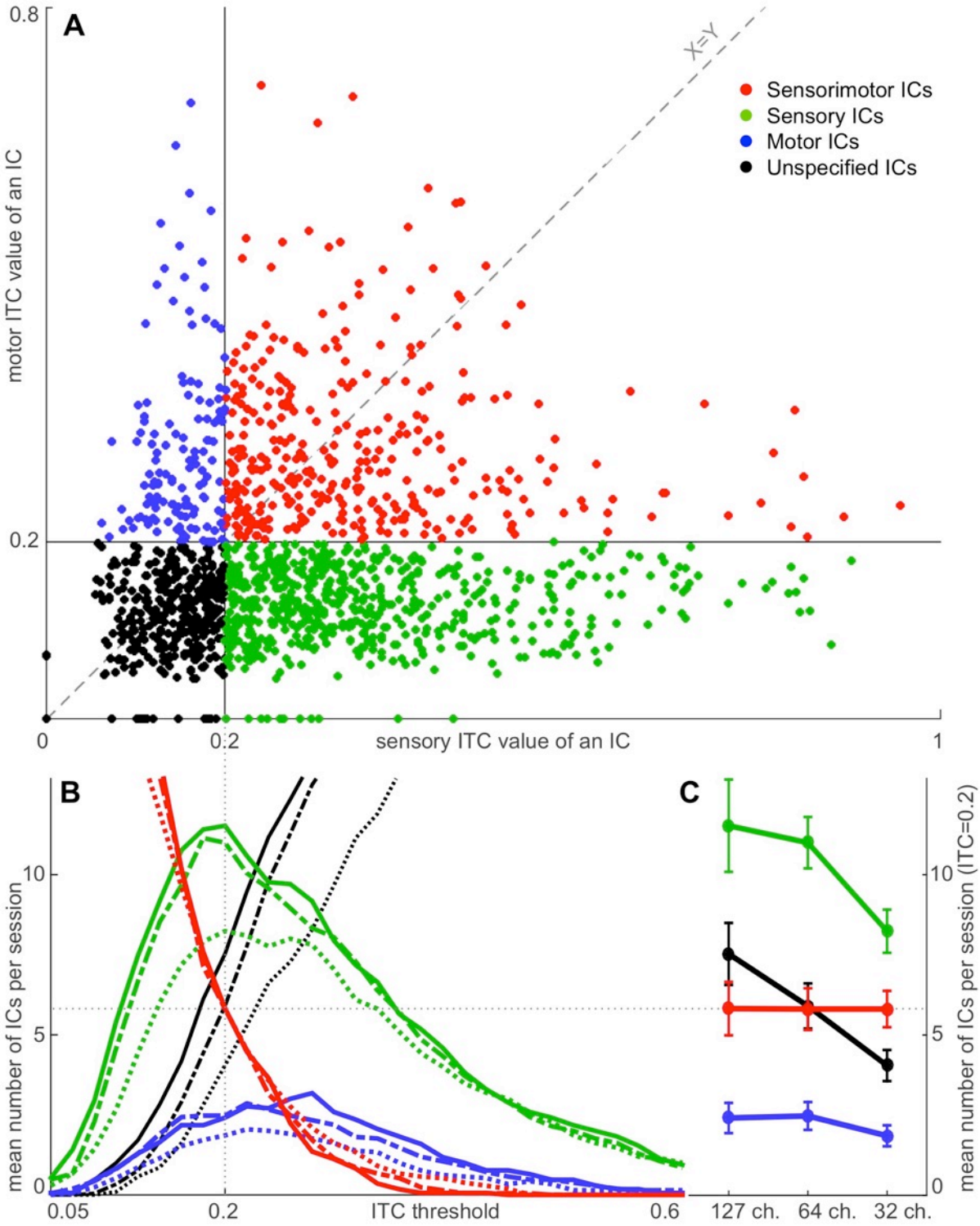


Figure 4. Distribution of ICs into the four groups: Motor, Sensory, Sensorimotor, and Unspecified. **Panel (A)** depicts 1412 accepted ICs in the study from 32-, 64- and 127-channel datasets. Each dot represents an IC with a Sensory ITC value as the X coordinate and a Motor ITC value as the Y coordinate. Blue dots represent Motor ICs. Green dots represent Sensory ICs. Red dots represent Sensorimotor ICs. Black dots represent Unspecified ICs. The vertical black line at X = 0.2 represents the sensory ITC threshold, and the horizontal black line at Y = 0.2 represents the motor ITC threshold to separate ICs into the four groups. The grey diagonal dashed line (X = Y) represents the ITC-threshold trajectory of the intersection of the threshold lines. The results of distribution of ICs into the four groups at different points of the trajectory are shown in **Panel (B)**. The X-axis is derived from the X = Y dashed line in **Panel (A)**. The Y-axis represents a mean number of ICs in a group per session at different points of the grey diagonal dashed ITC-threshold-trajectory line (X = Y) in **Panel (A)**. The grey dotted line at X = 0.2 indicates the selected ITC threshold for Sensory and Motor ITC values in **Panel (A)**. X-axis data points interval in the illustration is equal to 0.02. Red curves represent Sensorimotor ICs. Green curves represent Sensory ICs. Blue curves represent Motor ICs. Black curves represent Unspecified ICs. Dotted curves represent 32-channel datasets. Dashed curves represent 64-channel datasets. Solid curves represent 127-channel datasets. **Panel (C)** shows the projection of the intersection points of the 12 coloured lines in **Panel (B)** with the grey dotted line at X = 0.2. Y-axis and colour coding remain the same as in **Panel (B)**. Error bars represent standard error of the mean.

at ITC threshold = 0.2	32 channels	64 channels	127 channels
Sensorimotor ICs	5.8 ± 0.6	5.8 ± 0.6	5.8 ± 0.8
Sensory ICs	8.2 ± 0.7	11.0 ± 0.8	11.5 ± 1.4
Motor ICs	1.9 ± 0.3	2.5 ± 0.4	2.4 ± 0.5
Unspecified ICs	4.0 ± 0.5	6.0 ± 0.7	7.5 ± 1.0
All groups, good ICs	20.0 ± 0.8	25.2 ± 1.3	27.3 ± 2.5

Table 1. Average number of ICs per session (±s.e.m.) in 32-, 64-, and 127-channel datasets and in the four groups, given the sensory and motor ITC thresholds = 0.2

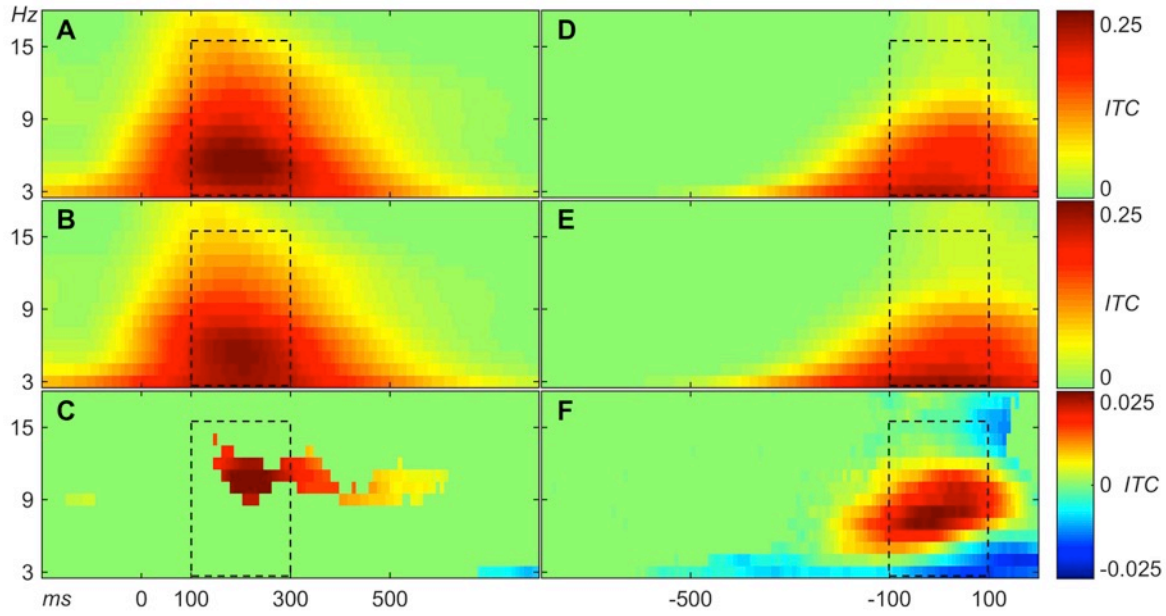


Figure 5. Average ITC plots of Sensory, Motor, and Sensorimotor ICs. A. Average ITC over 600 Sensory ICs. **B.** Average ITC over 343 Sensorimotor ICs. **C.** Average ITC of Sensory ICs minus average ITC of Sensorimotor ICs. Note that colour-coding of subtraction is ten times amplified relative to panels A and B. 0 ms in plots A, B, and C indicate stimulus onset **D.** Average ITC over 132 Motor ICs. **E.** Average ITC over 343 Sensorimotor ICs. **F.** Average ITC of Motor ICs minus average ITC of Sensorimotor ICs. Note that color-coding of subtraction is ten times amplified relative to panels D and E. 0 ms in plots D, E, and F indicate motor onset (button-press event). Panels C and F were masked according to two-sample t-test with p-value threshold < 0.01. The black dashed rectangles represent the same time-frequency windows of interest as in Figure 3.

Does number of channels in EEG recordings influence number of Sensorimotor ICs?

One potential concern was that there might have been a minimum number of ICs necessary to fully resolve the brain processes (Lau et al., 2012). Because the number of ICs produced was equal to the number of EEG channels, the number of Sensorimotor ICs observed may have been dictated by the number of EEG channels used. To address this, we examined data recorded from subjects with 64- and 127-channel EEG systems, using the same experimental design twice, and additionally extracted a subset of 32 channels from the 64-channel EEG recordings. The extraction was performed after filtering and downsampling but before the ICA procedure. The subsets of 32 channels had the same head coverage as 64-channel recordings, but with half of the electrode density, resulting in a total of twenty-one 32-

channel datasets, twenty-one 64-channel datasets, and seventeen 127-channels datasets. We processed all three types datasets (32, 64, and 127 channels) applying the same data-processing procedure in order to find out whether the number of channels influences the number of Sensorimotor ICs. We applied a series of ITC thresholds from 0 to 1 with a step of 0.02 to all accepted ICs in the study (1412 ICs) and plotted the results of separation of ICs into the four groups (Motor, Sensory, Sensorimotor, or Unspecified) in Fig. 4B. Colours of the lines in Fig. 4B are congruent with colours of the dots in Fig. 4A; solid lines in Fig. 4B represent 127-ch. datasets, dashed lines represent 64-ch. datasets, and dotted lines represent 32-ch. datasets. Notably the red solid, dashed, and dotted lines in Fig. 4B are overlaying, suggesting that the number of Sensorimotor ICs does not correlate with the number of channels in an EEG system. This supports the interpretation that these are genuine Sensorimotor ICs.

Localization of sensorimotor ICs

Where are the Sensorimotor ICs located? For each IC, we found an equivalent dipole (ED). In order to check whether locations are reproducible across subjects, we marked as reproducible only those Sensorimotor EDs which had neighbouring dipoles of at least 50% of subjects within a given radius R (coloured dots in Fig. 6). We marked Sensorimotor EDs which had neighbouring dipoles of less than 50% of subjects within a given radius R (grey dots in Fig. 6) as not reproducible across subjects. Moreover, a large part of non-reproducible Sensorimotor EDs were located outside of the cortex. To estimate a meaningful order of magnitude for the radius R, we calculated the volume of the brain accounted per Sensorimotor ED and extracted the radius of the sphere of the derived volume: $V_{sphere} = V_{brain}/N$; where $V_{sphere} = (4*\pi*R^3)/3$; V_{brain}= 1202 cm^3 - average volume of a human brain across genders(Allen, Damasio, & Grabowski, 2002); N=5.8 - average number of Sensorimotor ICs in a dataset in the study (Table 1). These calculations resulted in R=36.7 mm. The average volume of a human brain is the sum of the grey and white matters. To have a conservative selection, we took 75% of this radius (which a bit less than half of the volume), which resulted in R=27.5 mm. Based on this criterion, 151 Sensorimotor EDs were marked as not reproducible across subjects (grey dots in Fig. 6) and were excluded from further analysis. The previous rejections of artefact-related ICs (Fig. 1) were based on individual sessions. Here we use spatial information of repeated sessions, completing the selection process of ICs. Complementary, 192 Sensorimotor EDs were marked as reproducible across subjects (coloured dots in Fig. 6) and were analysis subsequently.

We used the 192 Sensorimotor EDs that were marked as reproducible across subjects in the k-means clustering. We set the number of clusters equal to 3

because the average number of Sensorimotor EDs reproducible across subjects per dataset was equal to 3.25 (192 Sensorimotor EDs divided by 59 datasets). Cluster 1 had its centroid at MNI = [-4 -2 45], Talairach = [-3 0 41] (http://sprout022.sprout.yale.edu/mni2tal/mni2tal.html) (Lacadie, Fulbright, Rajeevan, Constable, & Papademetris, 2008), which is left cerebrum, limbic lobe, cingulate gyrus, Brodmann area 24 (www.talairach.org/applet/) (J. L. Lancaster et al., 1997; Jack L. Lancaster et al., 2000); and encompasses 58 Sensorimotor EDs (0.98 per dataset). Cluster 2 had its centroid at MNI = [-50 -18 30], Talairach = [-48 -17 29], which is left cerebrum, frontal lobe, postcentral gyrus, white matter; and encompasses 53 Sensorimotor EDs (0.90 per dataset). Cluster 3 had its centroid at MNI = [-5 -51 38], Talairach = [-4 -48 36], which is left cerebrum, parietal lobe, precuneus, Brodmann area 31 (on the border with the posterior cingulate cortex); and encompasses 81 Sensorimotor EDs (1.37 per dataset). Two of the derived clusters were mostly symmetrical along the sagittal plane, and one cluster was located in the left hemisphere.

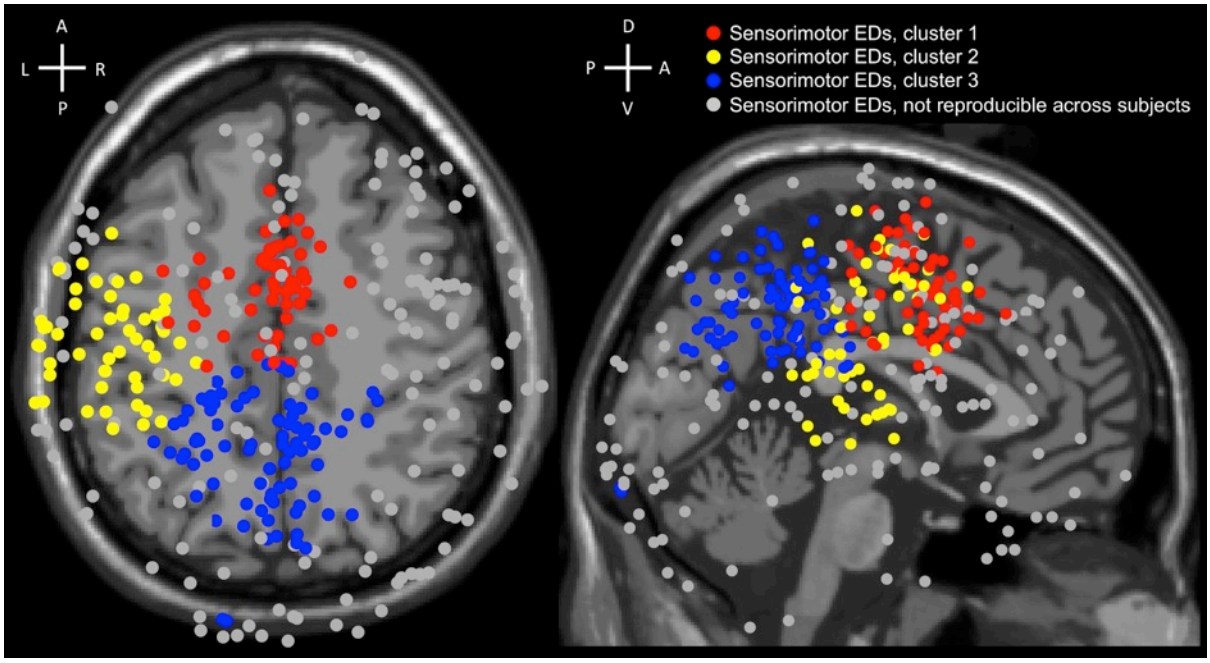
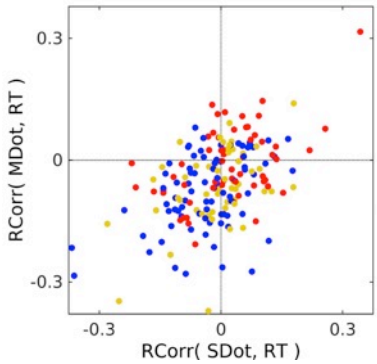


Figure 6. Three clusters (coloured in red, yellow, and blue) obtained by the k-means clustering algorithm from 192 Sensorimotor EDs reproducible across subjects from 32-, 64-, and 127-channel datasets. 151 Sensorimotor EDs, which were not reproducible across subjects, were coloured in grey and did not participate in the k-means clustering. Cluster 1 = Sensorimotor EDs coloured in red; Cluster 2 = Sensorimotor EDs coloured in yellow; Cluster 3 = Sensorimotor EDs coloured in blue. Labels at the crosses represent anterior (A), posterior (P), dorsal (D), ventral (V), left (L), and right (R) sides.

A



B

mean ± STD	RCorr(SDot,RT)	RCorr(MDot,RT)	Corr(SDot,MDot)
All 192 SM EDs	-0.01 ± 0.10	-0.05 ± 0.10	0.04 ± 0.06
SM EDs cluster 1	0.02 ± 0.11	-0.01 ± 0.09	0.03 ± 0.06
SM EDs cluster 2	-0.01 ± 0.09	-0.06 ± 0.09	0.04 ± 0.06
SM EDs cluster 3	-0.03 ± 0.10	-0.08 ± 0.09	0.05 ± 0.07

Figure 7. Rank correlation of sensory and motor related amplitudes in single trials with reaction time. (A) Scatter plot with 192 points representing all Sensorimotor (SM) EDs. Red points represent Sensorimotor EDs in cluster 1 (58 SM EDs). Yellow points represent Sensorimotor EDs in cluster 2 (53 SM EDs). Blue points represent Sensorimotor EDs in cluster 3 (81 SM EDs). The OX axis represents rank correlation (Spearman's rho) of reaction times (RT) in trials and dot products (SDot) of the sensory related ERP and single trials on the interval from 100 ms to 300 ms after the stimulus onset in a SM IC. The OY axis represents rank correlation of reaction times in trials and dot products (MDot) of the motor related ERP and single trials on the interval from -100 ms to 100 ms around the motor onset in a SM IC. **(B)** Table of mean values ± STD of correlation coefficients for all 192 SM EDs and for individual clusters depicted in Fig. 7A. RCorr(SDot,RT) represents rank correlation (Spearman's rho) of sensory related dot products in trials and reaction time in trials. RCorr(MDot,RT) represents rank correlation of motor related dot products in trials and reaction time in trials. Corr(SDot,MDot) represents linear correlation of sensory and motor related dot products in trials. Negative rank correlation coefficients reflect a correspondence of higher amplitude (dot product) with shorter reaction time in trials in an IC. We calculated the mean values without application of Fisher z-transformation, as the extreme values of correlation coefficients were between -0.4 to 0.4 with the highest density around zero.

2.4 Discussion

Our study revealed EEG correlates of sensorimotor processing in the human brain. Consistently across subjects, we found Sensorimotor ICs (Fig. 3A) with both event-related ITC responses tied to the stimulus onset and to the button-press event. We interpret these results as the Sensorimotor ICs represent functional neural modules that were active in both phases of sensorimotor processing. Thus, the findings support the second hypothesis that the human brain directly couples sensory recognition and motor actions rather than sequentially separating them. Our findings are compatible with alternative frameworks like embodied cognition (Foglia & Wilson, 2013), sensorimotor contingency (O'Regan & Noe, 2001), and common coding (Prinz, 1997). However, we also have found roughly as many Motor and Sensory ICs as Sensorimotor ICs. Thus, our findings do not exclude the classical view, which conceptualizes sensorimotor processing as a series of functional modules activated in a sequence from sensory input to motor execution.

The number of ICs in the Motor and Sensory groups increased with a higher number of channels, from 32 to 64 channels (Table 1). However, the increment of the number of channels from 64 to 127 did not lead to a noticeable increment of the number of ICs in these groups (Table 1). These results suggest that the number of independent sources in the brain that were involved in the sensorimotor task in our experiment and can be identified with ICA, was limited. Therefore, increasing the number of EEG channels beyond a certain number for stationary subject experiments does not lead to an increase of the number of Sensory, Motor, or Sensorimotor ICs. This is in agreement with previous work showing a functional maxima in the number of channels necessary (Lau et al., 2012).

ICA could have potentially failed to separate sensory and motor sources. Therefore, the presence of such sensorimotor components would not be due to integrated sensory and motor processing, but to failure of ICA to separate these two separate processes. Even then, a joint IC would suggest the colocalization of such processes, because ICA components are spatial filters. To address this potential issue, we recorded subjects with two EEG systems with different number of electrodes, and additionally we analysed a subset of 32 channels from 64-channels datasets. The number of Sensorimotor ICs did not correlate with the number of channels in an EEG system, supporting our expectation that these are genuine Sensorimotor ICs.

The inverse problem of source localization and possible artefacts within ICA, as well as our assumption that ICs can be represented by a single equivalent-dipole model, are potential limitations which constrain interpretation of the current results. We also

do not know the precise properties of such ICs within the frameworks of embodied cognition (Foglia & Wilson, 2013), sensorimotor contingency (O'Regan & Noe, 2001), and common coding (Prinz, 1997). If sensorimotor functional modules (represented by Sensorimotor ICs in the study) form a complex and distributed structure then the generation of more complex higher-order electric fields, e.g., a quadrupole, is possible. A dipole fit to such ICs would have a high residual variance and lead to a rejection by the dipole fit filtering procedure. However, in the current study we report the existence of sensorimotor functional modules represented by Sensorimotor ICs; thus, we deliberately chose the conservative approach to minimize the chances of false positive results and be sure that such Sensorimotor ICs are real.

The topographical sparseness threshold used to remove ICs from further analysis was set to 5 (Fig. 2). We did an additional control to check whether the threshold TS=5 might cause a selection bias on components selected for further analysis. Those results indicated that the threshold choice did not cause a selection bias. We found that TS threshold selection below 5 did not influence the proportion of Sensorimotor ICs (Sensory and Motor ITC values > 0.2) in the 32-, 64-, and 127-channel datasets. Having a TS threshold higher than 5 was not appropriate because we wanted to remove from further data analysis ICs which had a one-channel-dominant-projection-weight topography. In 32-channel datasets, such a topography corresponds to TS=5.5. Therefore, the current selection TS=5 seems reasonable and did not cause a selection bias on Sensorimotor ICs selected for further analysis.

The ITC threshold used to divide all ICs into the four groups (Sensorimotor, Sensory, Motor, and Unspecified - Fig. 3 and Fig. 4) was set to 0.2. Fig. 4 shows how the distribution of ICs into Sensory, Motor, and Sensorimotor groups changes when the ITC threshold moves. There was no gap of bi-modal distribution, and we did not find indications that the results depended on the exact ITC threshold. The selected ITC threshold = 0.2 seems to be a reasonable choice, as it is around the peaks of Sensory (green) and Motor (blue) lines (Fig. 4B). Furthermore, also for deviant choices of the threshold value, the number of Sensorimotor ICs was minimally dependent on the number of channels (Fig. 4B, solid, dashed, and dotted red line). Thus, we did not find any indication that Sensorimotor ICs were a result of squeezing sensory and motor modules into a single IC due to a limited number of channels. Instead, additional degrees of freedom were taken by the Sensory ICs and the Unspecified ICs. The ITC threshold level = 0.2 gave a good separation of ICs into the groups and shifting the threshold slightly up or down did not change the ratio of Sensorimotor ICs in datasets with a different number of channels (32, 64 and 127 ch.).

Location of the centroid of Cluster 1 coincides with the anterior cingulate cortex (ACC). The anterior cingulate cortex is associated with conflict monitoring in the

engagement of cognitive control (Botvinick et al., 2004; Kerns, 2004), i.e., in the Stroop task (Stroop, 1935). Indeed, the paradigms in the current study are variations of the Stroop task. It has also been proposed that “the anterior cingulate cortex is a cortical region where a cognitive/motor command, coming from a different cortical region (e.g., prefrontal cortex), is being modulated and funnelled to the motor system, and this modulation takes place within distinct, motor output-specific subregions of the anterior cingulate cortex , thus emphasizing the motor character of this region” (Paus et al., 1993; Swick & Jovanovic, 2002; Turken & Swick, 1999). Errors in the localization accuracy of EDs are possible, limiting our conclusions to the anterior cingulate cortex likely being involved with sensorimotor processing for our task.

The centroid of Cluster 2 lies in white matter in the vicinity (±10 mm) of Brodmann areas 1, 2, 3, 4, 6, and 40 (Lacadie et al., 2008). Activity of the primary motor and somatosensory cortices could explain motor event-related ITC response of Sensorimotor ICs in Cluster 2. Indeed, the index and the ring fingers of the right hand were used in the experiment to conduct motor responses. However, Sensorimotor ICs have sensory components as well, and the stimuli were centrally presented, which does not correlate with the leftward bias. Thus, it is not completely obvious why using the right hand should lead to a bias towards left hemispheric Sensorimotor ICs.

Location of the centroid of Cluster 3 coincides with the precuneus, in the area bordering the posterior cingulate cortex. Both the precuneus and the posterior cingulate cortex are associated with high-level behavioural correlates, as well as being a part of an attentional system (Cavanna & Trimble, 2006; Leech & Sharp, 2014). An fMRI study of Stroop tasks proposed that the precuneus is a part of a system which is responsible for executive aspects of attentional selection, in particular, sets a top-down bias for selecting certain types of information, e.g., colour (Banich et al., 2000). Another event-related fMRI study where subjects performed a variation of the Wisconsin Card Sorting Test concluded that an increase in neural activity in the precuneus correlates closely with the execution of attentional shifts between object features (Nagahama et al., 1999). The complex speeded-response tasks in the current study require attentional resources to compare and match stimuli features, as well as a top-down bias for selecting relevant information. Therefore, event-related ITC responses of Sensorimotor EDs from Cluster 3 during sensory, as well as motor processing, allow the assumption that the attentional system is involved in sensorimotor processing.

Rank correlation (Spearman's rho) showed that some Sensorimotor ICs have a noticeable correlation between reaction time and the amplitude of sensory or motor response in a trial. However, Sensorimotor ICs did not show any prominent shift towards the positive or negative direction. Mean values of all Sensorimotor EDs as

well as their sub clusters 1, 2, and 3 lie within one standard deviation from zero. Cluster 3 had the highest negative rank correlation coefficients -0.08 ± 0.09 STD (Fig. 7) for motor responses. Negative rank correlation coefficients in Fig. 7 indicated a correspondence of higher amplitude (dot product) with shorter reaction time in trials in an IC. Overall, there was a positive correlation of sensory and motor dot products in ICs, which means that rank correlation sign had a higher probability to be the same for sensory and motor rank correlations with reaction time in an IC. It has been shown before that some ICs may have both sensory and motor component (Makeig, Delorme, et al., 2004). However, as far as we know it has not been systematically studied as a primary variable.

There are other noninvasive imaging techniques that can quantify brain activity, like functional magnetic resonance imaging (fMRI) (Scott A. Huettel; Allen W. Song; Gregory McCarthy et al., 2004). An fMRI study would give a higher spatial resolution and more precise localization and structure of sensorimotor functional modules in the brain. However, for the analyses highlighted here, resolution in the temporal domain is a critical feature. In BOLD fMRI, the response is a smooth continuous function and a faster sampling than once per second yields only linear interpolation, since a hemodynamic response lasts over 10 seconds (Scott A. Huettel; Allen W. Song; Gregory McCarthy et al., 2004). Thus, given that the average reaction time in the study was 515 ms, it would be difficult or impossible to differentiate sensorimotor processes from sensory processes and motor processes by means of fMRI measurements in the current setup. Recordings from an electrode array in the brain tissue may provide both, good spatial and good temporal resolution, but they require either animal or clinical studies. Finding the exact structure and localization of sensorimotor functional modules are important questions for future research.

Several fMRI studies reported increased activity of the anterior cingulate cortex, the posterior cingulate cortex, and the precuneus during the execution of Stroop tasks (Banich et al., 2000; Michael P. Milham, Banich, & Barad, 2003; Swick & Jovanovic, 2002; Turken & Swick, 1999). These areas are compatible with a rich spectrum of functional significance, including attention, self-consciousness, conflict monitoring, and cognitive control (Bear et al., 2007). These findings support our results, since these brain areas coincide with Cluster 1 and Cluster 3 centroids.

Different types of sensorimotor tasks (i.e., whole body movement) also activate brain areas which coincide with the detected clusters of Sensorimotor ICs in the current study. A study of vestibular sensory processing with motor tasks revealed that impending loss of balance is associated with electrocortical dynamics in the anterior cingulate, posterior cingulate, anterior parietal, bilateral sensorimotor, and superior dorsolateral-prefrontal cortex (Sipp, Gwin, Makeig, & Ferris, 2013). Interestingly, that study also found a left hemisphere bias in sensorimotor IC clusters. Another EEG study showed that the premotor and parietal areas are involved in visually guided

gait in humans. Mu, beta, and lower gamma rhythms in premotor and parietal cortices are suppressed during conditions that require an adaptation of steps in response to visual input. The study concluded that “activity in the parietal cortex likely reflects direct visuomotor transformations required by the task. Increased activity in the premotor cortex may indicate motor planning involved in adapting the steps to the visual input” (Wagner, Solis-Escalante, Scherer, Neuper, & Müller-Putz, 2014). Thus, different types of tasks, as well as processing in different sensory modalities coupled with motor tasks, correlate with electrocortical dynamics in the brain areas related to the detected clusters of Sensorimotor ICs in the current study.

The present results are congruent with multiple frameworks emphasizing a coupling of sensory and motor brain processes. The embodied cognition framework, espousing that perception guides motor actions and motor actions influence perception is one such conceptual theory. Another is the sensorimotor-contingency theory (O’Regan & Noe, 2001), which emphasizes the key role of action for perception. Our results show bias in Sensorimotor ED locations toward the left hemisphere (Fig. 6). One reason for that might be that the right hand was used in the experiment to conduct motor responses. However, Sensorimotor ICs have sensory components as well, and the stimuli were centrally presented, which does not correlate with the leftward bias. This is congruent with the “stimulus-response compatibility” effect which is explained by the common-coding theory put forward by Wolfgang Prinz (Prinz, 1992, 1997). This theory is built on the idea that sensory and motor codes share some common features and that perceptual and motor representations are linked. Thus, according to the common-coding theory, there are some neural substrates involved in the decoding of sensory information coding of motor actions. Sensorimotor ICs in the current study have strong event-related responses after a stimulus onset as well as at the moment of a button-press event. However, additional processing at higher cortical areas may be required to clarify decoded information and modulate to an appropriate action. This may lead to a certain variance in reaction time and a delay between sensory and motor event-related responses of Sensorimotor ICs. Other evidence for this common-coding theory comes from observations in human and rhesus monkeys where the same neurons responded after a stimulus onset and timed to the motor response (Cogan et al., 2014; Hickok, Okada, & Serences, 2009; Roitman & Shadlen, 2002). Combined with the current results, it is possible to interpret the findings as supporting functional modules directly coupling sensory and motor processes.

In this study, we consistently found across subjects a substantial number of Sensorimotor ICs, in which event-related ITC responses are related to both sensory stimulation and motor response onsets. Sensorimotor ICs are possible EEG correlates of sensorimotor processing in the human brain. We hypothesised that these Sensorimotor ICs represent functional modules in the brain, combining sensory recognition and motor responses. This is compatible with such frameworks

as embodied cognition (Foglia & Wilson, 2013), sensorimotor contingency (O'Regan & Noe, 2001), and common coding (Prinz, 1992, 1997). Future studies could expand understanding of functional meaning of sensory and motor processing in such Sensorimotor ICs using different tasks and trying to more precisely identify their neuronal sources.

2.5 Methods

Participants & EEG recordings

Twenty-one paid healthy volunteers (eleven males), aged 19-29 years (mean = 24, sd = 3.1), participated in the EEG study. All participants had normal (self-reported) or corrected-to-normal vision. We obtained written informed consent from all subjects prior to the experiment and the protocol had been approved by the review board of the University of Osnabrück. Experimental procedures conformed to the Declaration of Helsinski and national guidelines. This project was a part of a larger study, in which participants were recorded two times in the same EEG experiment but with different EEG system. In the following, it was important for us that these EEG systems had a different number of channels. The first session was recorded with a 64-channel ActiCAP EEG system and the second session was recorded with a 127-channel ASA-Lab EEG system on a different day. These two EEG systems provide statistically indistinguishable EEG data (Melnik et al., 2017). Seventeen participants were recorded in both sessions and four participants were recorded only in the first session.

Stimuli and tasks

The stimuli were a combination of a coloured text (visual angle: 2x1°) on a black background presented centrally on a computer screen and a simultaneously occurring auditory word through a headset (Fig. 8). Thus, each stimulus contains three sources of information about a colour, which may or may not match. The experiment was conducted in German; four different colours were used: rot (German red), grün (German green), blau (German blue), and gelb (German yellow), yielding 64 possible combinations. The setup is reminiscent of the Stroop effect paradigm(Stroop, 1935), but is a distant variation of it due to containing an auditory component as well as different tasks and type of responses.

Each recording consisted of 5940 trials (one stimulus and one response per trial) divided into 6 blocks with breaks in between. One categorization rule per block was used. The categorization rule in Block 1: to press Arrow Left (True) when colour=red

and text=red, Arrow Right (False) in other cases; in Block 2: to press Arrow Left (True) when colour=red and sound=red, Arrow Right (False) in other cases; in Block 3: to press Arrow Left (True) when text=red and sound=red, Arrow Right (False) in other cases; in Block 4: to press Arrow Left (True) when colour=text, Arrow Right (False) in other cases; in Block 5: to press Arrow Left (True) when colour=sound, Arrow Right (False) in other cases; in Block 6: to press Arrow Left (True) when text=sound, Arrow Right (False) in other cases. Participants categorized each stimulus in accordance with a given rule by pressing on a keyboard Arrow Left with the index finger of the right hand or Arrow Right with the ring finger of the right hand. 50% of trials in each block belonged to one of the two categories. The two categories appeared in a random sequence. We pooled all six blocks to get more data for the ICA and ITC analyses.

The duration of the auditory part of a stimulus were as follows: rot (German red) = 340 ms, grün (German green) = 340 ms, blau (German blue) = 340 ms, gelb (German yellow) = 330 ms. The visual part of the stimulus were presented on the monitor until a button-press event occurred. We measured reaction time from the common onset of the auditory and the visual parts of a stimulus and until a button-press event. In the case of a correct response, the coloured word disappeared from the screen for 600 ms, and after that, a new stimulus appeared. In the first half of trials of each block, the disappearance of the visual part of the stimulus occurred immediately after a button-press event (in the case of a correct response). In the second half of trials of each block, the disappearance of the visual part of the stimulus occurred after a short delay of 150 ms after a button-press event (in the case of a correct response). In the case of an incorrect response, the coloured word stayed on the screen unchanged for 5000 ms to provide feedback to the participant.

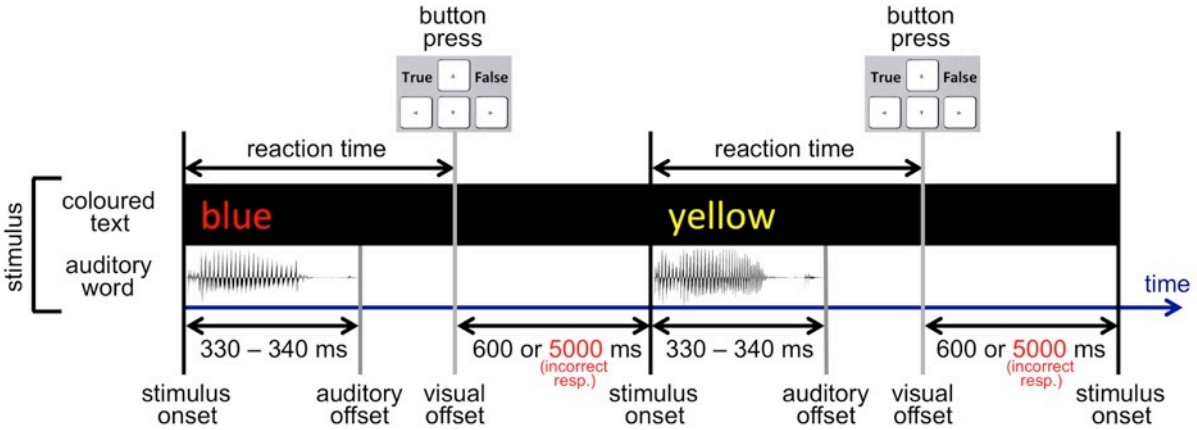


Figure 8. Experimental setup. A stimulus is a coloured text appearing on a monitor simultaneously with an auditory word through a headset. In a reaction-time task, subjects had to categorize each stimulus by pressing Arrow Left (True) or Arrow Right (False) on a keyboard according to a given categorization rule in each block. Please note, that the figure represents the first half of trials of each block, in the second half of trials of each block, the disappearance of the visual part of the

stimulus (visual offset) occurred after a short delay of 150 ms after a button-press event (in the case of a correct response).

Participants were motivated to give a correct response as quickly as possible. They always had the possibility to press a pause button and check their current mean reaction time and error-response statistics, as well as to compare their reaction time to that of other participants, which increased their motivation to give a correct response as quickly as possible. Subjects could also earn a bonus of up to 30% of the basic payment; the shorter the reaction time, the bigger the payment. However, in order to ensure accuracy, if the number of incorrect responses exceeded 5%, they lost their bonus completely, as well as being moved down in the reaction-time score table. Therefore, the usual strategy was to shorten reaction time down to an efficient level, where the number of incorrect responses did not exceed 5%. The mean number of error responses per session was 4.1% ± 1% (STD). All participants got a short practice before the real experiment to learn the strategy and reach a plateau of performance.

The following trials were discarded from EEG-data processing and reaction time analysis: each ten first trials after the beginning of each new block; trials with error responses, as well as two trials after them; one trial before and three trials after each pause (subjects were able to pause the experiment anytime). The mean number of discarded trials per recording was 14.7% ± 3.5% (SD). All not discarded trials were marked as valid.

We measured reaction time as the time from the common onset of the auditory and the visual stimulus components until a button-press event occurred (Fig. 8). The grand mean reaction time of valid trials (not discarded) across all thirty-eight recording sessions was 515 ms ± 213 ms (STD) (dotted line in Fig. 9). The mean reaction time for "Arrow Left (True)" categorizations was 486 ms ± 184 ms (STD) and for "Arrow Right (False)" categorizations was 543 ms ± 236 ms (STD). The mean reaction time for "Arrow Left (True)" and for "Arrow Right (False)" categorizations in the six blocks was the following (Fig. 9): in Block 1, 405 ms ± 115 ms (STD) and 413 ms ± 138 ms (STD) respectively; in Block 2, 402 ms ± 137 ms (STD) and 442 ms ± 161 ms (STD) respectively; in Block 3, 397 ms ± 112 ms (STD) and 442 ms ± 142 ms (STD) respectively; in Block 4, 596 ms ± 195 ms (STD) and 652 ms ± 226 ms (STD) respectively; in Block 5, 604 ms ± 209 ms (STD) and 704 ± 293 ms (STD) respectively; in Block 6, 511 ms ± 168 ms (STD) and 598 ms ± 225 ms (STD) respectively.

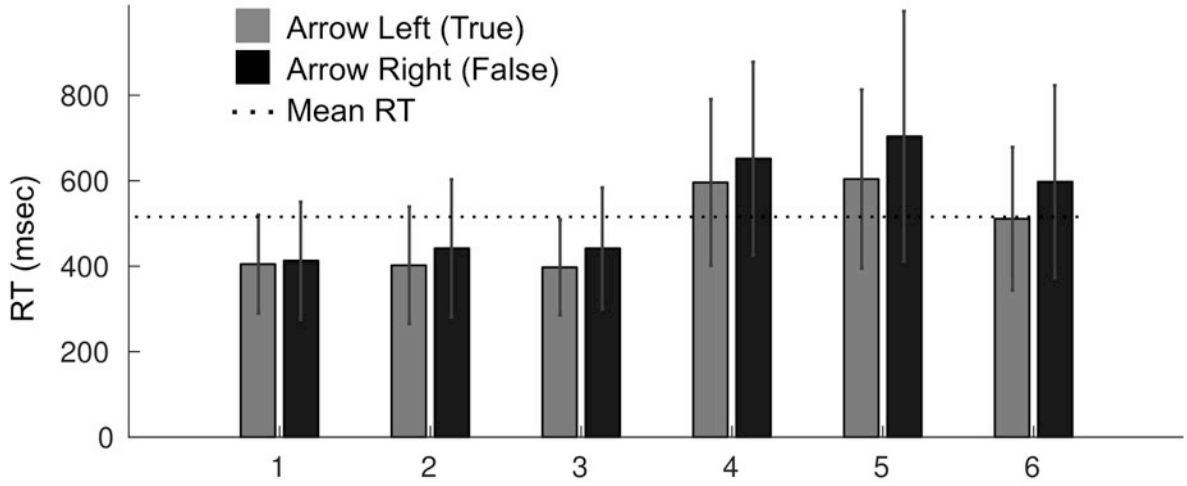


Figure 9. Reaction time for button presses Arrow Left (True) and Arrow Right (False) in the six blocks. The dotted line depicts the grand mean RT (515 ms) in the study. The X-axis depicts block ID. The Y-axis depicts the mean reaction time (RT) of all respective trials across all recording sessions. Error bars represent standard deviation.

EEG-data processing

EEG-data processing and analysis was performed using EEGLAB (Delorme & Makeig, 2004) in Matlab. Each recording was filtered with a 3–45 Hz bandpass finite impulse response filter. Sampling rate was reduced from 1000 to 500 Hz for 64-channel ActiCAP EEG system and from 1024 to 512 Hz for 127-channel ASA-Lab EEG system. Noisy periods of EEG data were manually marked and cut out. The EEG data after elimination were processed in each recording by ICA using EEGLAB's 'runica' function (Bell & Sejnowski, 1995; Delorme & Makeig, 2004; Makeig et al., 1996). Further data analysis and results were based on the resulting ICs and their activity. Please note, "as scale and polarity information is distributed in the ICA decomposition (not lost!) between the projection weights (column of the inverse weight matrix, EEG.icawinv) and rows of the component activations matrix (EEG.icaact), the absolute amplitude and polarity of component activations are meaningless and the activations have no unit of measure (through they are proportional to microvolt)" (https://sccn.ucsd.edu/wiki/Chapter_10:_Working_with_ICA_components).

Calculation of Sensory and Motor ITC values of ICs

We calculated the ITC of all trials in a session two times per IC (Fig. 3): the first time, for stimulus alignment of trials and the second time, for button-press alignment of

trials. ITC calculation was done in EEGLAB using the function 'pop_newtimef' with two-tailed permutation significance probability test ('alpha' = 0.01). Non-significant features of the output were zeroed out and plotted in green. We measured the maximum ITC value in relevant time-frequency windows of interest. For stimulus-onset-aligned trials, the time-frequency window of interest was from 100 ms to 300 ms and from 3 Hz to 15 Hz. The maximum value of ITC in the time-frequency window of interest was assigned as the Sensory ITC value of the IC (Fig. 3). For motor-response-onset-aligned trials, the time-frequency window of interest was from -100 ms to 100 ms and from 3 Hz to 15 Hz. The maximum value of ITC in the time-frequency window of interest was assigned as the Motor ITC value of the IC (Fig. 3). Thus, each IC had sensory-related and motor-related ITC values, which served as scalar measures of sensory and motor event-related responses of an IC, respectively. This kind of dual-analysis of stimulus-locked and response-locked trials is also common in monkey neurophysiology saccade experiments with intracranial recordings in order to differentiate neurons that are sensory- or motor-potential related(Shadlen & Newsome, 2001).

Equivalent dipoles localization

In order to investigate localization of ICs, an equivalent dipole (ED) was calculated for each IC using the 'Fine fit' procedure for a single dipole model in the "DIPFIT v2.3" toolbox (EEGLAB). The head model 'standard_vol.mat' and the MRI file 'standard_mri.mat' (MNI Colin27) (Holmes et al., 2015) provided in the toolbox were used as the basis. Since we do not have MRI data from subjects, we chose this relatively simple method of dipole localization. Another option would be to calculate dipole density (Delorme, Palmer, Onton, Oostenveld, & Makeig, 2012) using dipole density function of EEGLAB. Prior to each EEG recording we digitized the 3D positions of all electrodes and three major fiducials (nasion, left and right preauricular points) using the optical ANT Neuro xensor™ system (ANT Neuro, Enschede, Netherlands; www.ant-neuro.com/products/xensor). We warped these positions of electrodes to fit the MNI Colin27 template afterwards. Five reference points were used to obtain the best fitting of a recorded electrode positioning model and the head model: three fiducial points (nasion, left and right preauricular), a position of the most posterior electrode, and a position of the uppermost electrode. After the best fitting between the head model and a recorded electrode positioning model was achieved (by moving, rotating, and resizing of the electrodes model along axes), electrodes were projected onto the head model. If residual variance of an ED exceeded 15% (Hammon et al., 2008; Onton et al., 2005; Wyczesany & Ligeza, 2014), then the corresponding IC was marked as residual-variance-bad and excluded from further data analysis (Fig. 2).

2.6 Acknowledgments

This work was supported by the Cognition and Neuroergonomics Collaborative Technology Alliance (W911NF-10-2-0022). The authors would like to thank Scott Makeig, Tim Kietzmann, Benedikt Ehinger, and Nicolas Kuske for useful comments on an earlier version of the manuscript and Allison Moreno-Drexler for proofreading.

Chapter 3. Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data?

This study has been published in “Frontiers in Human Neuroscience”.

Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., & König, P. (2017). Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data?. Frontiers in Human Neuroscience, 11.
https://doi.org/10.3389/fnhum.2017.00150

AUTHOR CONTRIBUTIONS
AM, SMK, WDH, DPF and PK conceived and designed the experiments; AM, PL and KI performed the experiments; AM and PK analyzed the data; AM, WDH, DPF and PK wrote and reviewed the article.

3.1 Abstract

Lab-based electroencephalography (EEG) techniques have matured over decades of research and can produce high-quality scientific data. It is often assumed that the specific choice of EEG system has limited impact on the data and does not add variance to the results. However, many low cost and mobile EEG systems are now available, and there is some doubt as to the how EEG data vary across these newer systems. We sought to determine how variance across systems compares to variance across subjects or repeated sessions. We tested four EEG systems: two standard research-grade systems, one system designed for mobile use with dry electrodes, and an affordable mobile system with a lower channel count. We recorded four subjects three times with each of the four EEG systems. This setup allowed us to assess the influence of all three factors on the variance of data. Subjects performed a battery of six short standard EEG paradigms based on event-related potentials (ERPs) and steady-state visually evoked potential (SSVEP). Results demonstrated that subjects account for 32% of the variance, systems for 9% of the variance, and repeated sessions for each subject-system combination for 1% of the variance. In most lab-based EEG research, the number of subjects per study typically ranges from 10 to 20, and error of uncertainty in estimates of the mean (like ERP) will improve by the square root of the number of subjects. As a result, the variance due to EEG system (9%) is of the same order of magnitude as variance due to subjects (32% / sqrt(16) = 8%) with a pool of 16 subjects. The two standard research-grade EEG systems had no significantly different means from each other across all paradigms. However, the two other EEG systems demonstrated different mean values from one or both of the two standard research-grade EEG systems in at least half of the paradigms. In addition to providing specific estimates of the variability across EEG systems, subjects, and repeated sessions, we also propose a benchmark to evaluate new mobile EEG systems by means of ERP responses.

3.2 Introduction

Electroencephalography (EEG) techniques have matured over the last 80-plus years since the study of the electrical activity of the human brain by Hans Berger (Berger, 1935). The hardware and processing methods for EEG data have advanced to being very high quality. Recording scalp potentials in fixed temporal relation to events, e.g. stimulus presentation or reports of recognition by button presses, allows for the capture of electrocortical activity related to sensory, motor, or cognitive processes. Such scalp potentials recorded and averaged over many trials, called event-related potential (ERP) (Luck, 2005; Sur & Sinha, 2009), help to reveal important insights about how the human brain works. For example, processing of auditory stimuli was studied with the N100 ERP component (Davis, 1939) and the N170 ERP component relates to the neural processing of faces (Rossion & Caharel, 2011). Currently, EEG is one of the most widely used techniques in noninvasive brain research to study correlates of perceptual, cognitive, and motor activity associated with processing of information.

A shift in the field of neuroscience to a greater appreciation of natural behavior is placing new demands on recording techniques. Participants have to act and move in order to experience the proprioceptive and vestibular sensations under natural conditions (Klaus Gramann et al., 2014; Mcdowell et al., 2013; Snider et al., 2013), parallel to a simultaneous recording of participants' motor actions and external events influencing cognition. Many traditional noninvasive imaging techniques to estimate brain activity, like functional magnetic resonance imaging (fMRI) (Scott A. Huettel; Allen W. Song; Gregory McCarthy et al., 2004) or magnetoencephalography (MEG) (Hansen et al., 2010), have mobility constraints that limit the ability for use in natural conditions. In contrast, mobile brain/body imaging ("MoBI") has the potential to provide unique insight into cognition in normal everyday behaviors (Ojeda et al., 2014). Near-infrared spectroscopy (NIRS) (Workman & Weyer, 2007) allows mobility and can provide blood oxygenation level dependent signals regarding brain function, but its temporal resolution is not close to that of EEG. The high temporal resolution of EEG makes it the prime candidate for a measurement technique of brain dynamics in modern real-world paradigms investigating cognition under natural contexts and those involving sensorimotor coupling and actions (Aspinall et al., 2013; De Sanctis et al., 2014; K Gramann et al., 2010; Makeig et al., 2009).

A desire for making EEG recordings under natural conditions, and potential uses in the gaming and wellness industries, have triggered the development of affordable and mobile EEG systems. These systems tend to be low-budget, easy to set up, and convenient for wearing over an extended period of time (Hairston et al., 2014; Hairston & Lawhern, 2015). Multiple aspects of EEG systems can influence the quality of the signal. For example, conventional EEG sensors are usually based on a

conductive gel, which leads to time-consuming setup and gel removal from the hair and the electrodes afterward. Newer systems can use dry electrodes or saline-based electrodes**.** EEG amplifiers can have varying parameters like noise performance, power consumption, signal bandwidth, and cost (Badillo et al., 2003; Hairston et al., 2015; Harrison & Charles, 2003; Lin et al., 2006). More recent developments include wireless data transition directly from a cap (M. De Vos et al., 2014), ultra-low power digitization, and a minimal use of cables (Warchall et al., 2016). All of these aspects have a potential to alter EEG data. How much do these changes affect the recorded signals? Are the results obtained with new mobile EEG systems and research-grade EEG systems equivalent?

Recent studies have attempted to compare recordings obtained from different EEG systems. In one study (Gargiulo et al., 2010), investigators compared a new EEG system with dry electrodes to a standard and clinically available EEG system with wet electrodes using parallel and serial recording methods. The comparison included two experimental paradigms, in which frequency domain and correlation coefficients of channel data, and ERPs were compared. Another study examined the correlation of signals from dry foam-based EEG sensors and wet EEG sensors, as well as the impedance at the sensor-skin contact during long-term EEG measurements (Liao et al., 2012). (Yeung et al., 2015) completed a comparison of foam-based and spring-loaded dry EEG electrodes using three experimental paradigms by linear correlation analyses. Additional studies have looked at the application-based performance using P300 brain-computer interfaces with dry and gel-based electrodes (M. De Vos et al., 2014; Guger et al., 2012). The general conclusion from these studies was that EEG data could be successfully collected using non-research grade EEG systems when taking into account the number and placement of electrodes.

With the increasing number of comparisons of EEG systems, we need a quantitative standard for comparing EEG systems (Oliveira et al., 2016b). A systematic benchmark based on a variety of paradigms testing different brain-response effects would allow assessing how results vary across EEG systems (low-budget, dry electrode, research-grade). That will allow comparing not only systems evaluated in a particular paper but also any new EEG system (Senevirathna et al., 2016) in the future to the database of performance results currently being collected. Furthermore, to rightly estimate how large variability due to systems is, we have to compare it to other factors of variance, like subjects or repeated sessions for each subject-system combination.

The primary purpose of this study was to quantify the variance across EEG systems, subjects, and sessions for some standard EEG measures. We tested four EEG systems: two standard research-grade systems, one system designed for mobile use with dry electrodes, and one mobile low-budget system with a lower channel count. We used six well-established ERP paradigms: (1) auditory evoked potentials (AEPs);

(2) steady-state visually evoked potential (SSVEP); (3) motor potentials (MPs); (4) visual mismatch negativity (vMMN); (5) face-sensitive N170 component; and (6) target-distractor visual decision-making (vDM). We recorded four subjects three times with each of the four EEG systems. This setup allowed us to assess the influence of all three factors (System, Subject, Session) on the variance of EEG data.

3.3 Materials and methods

GENERAL METHODS

Subjects

Four paid healthy volunteers (three males, mean age: 24 years, range 23-25 years), participated in the study. All subjects had normal or corrected-to-normal vision (self-reported) and normal hearing (self-reported). Subjects sat in a darkened EEG recording chamber in front of a computer monitor from a distance of 90 cm and with a headset on. We obtained written informed consent from all subjects before the experiment and the protocol had been approved by the University Osnabrück ethics committee.

Design

We recorded each subject three times with each of four EEG systems, which gave 12 EEG sessions per subject or 48 EEG sessions in the study (Fig. 1). Median time interval between recording sessions was 2 day. In each recording session, subjects performed a battery of six standard EEG paradigms (Fig. 2) based on event-related potentials (ERPs) and steady-state visually evoked potential (SSVEP): (Fig. 2A) auditory evoked potentials (AEPs); (Fig. 2B) SSVEP, evoked by an alternating contrast checkerboard; (Fig. 2C) motor potentials (MPs) elicited by voluntary tapping; (Fig. 2D) visual mismatch negativity (vMMN) (ERP waveforms subtraction ‘deviant minus standard’); (Fig. 2E) face-sensitive N170 component (ERP waveforms subtraction 'faces minus cars'); (Fig. 2F) visual decision-making (vDM) N240 component (ERP waveforms subtraction ‘targets minus distractors’). The current selection of paradigms covers different modalities (visual, auditory, as wells as motor) and ranges from a low level to a higher level cognitive load. This setup allowed us to assess the influence of the factors: System, Subject, and Session on EEG data in different standard EEG paradigms.

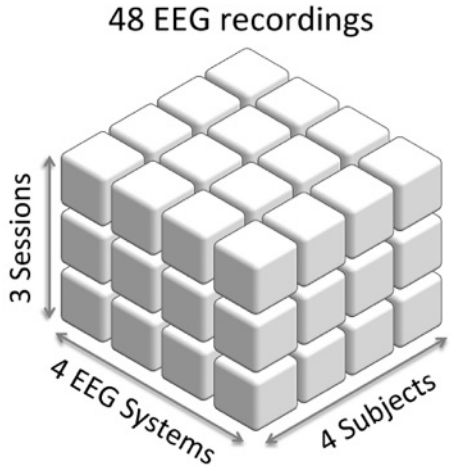


Figure 1. Experimental setup of EEG recordings. Each cube represents an EEG recording. We recorded four subjects with four EEG systems and three sessions for each combination of a subject and an EEG system. This results into 48 EEG recordings in the study.

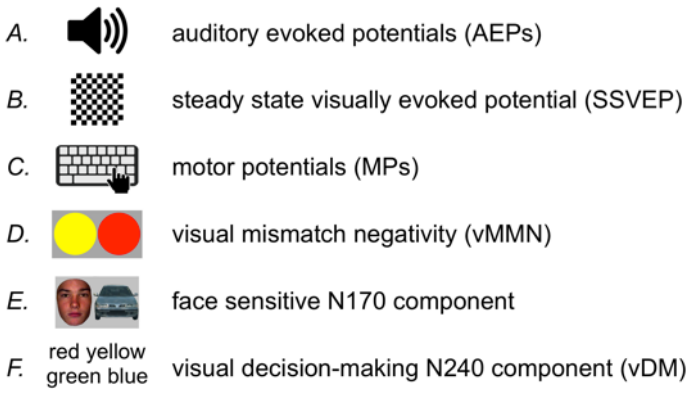


Figure 2. In each recording session, subjects performed a battery of six standard EEG paradigms. Pictograms A-F schematically represent the paradigms.

The order of the paradigms was identical for all recordings and had the following sequence: vMMN, SSVEPs, N170, MPs, vDM, AEP. We had additional paradigms which related not to ERPs but to time-frequency analysis and eye movements, which are however beyond the scope of the present paper. The additional paradigms took approximately 20 minutes before and 20 minutes after the paradigms processed in the current study. Each of the paradigms in the current study was divided into blocks and subjects could decide themselves when to start a new block and how long pause to take between blocks. Thus, duration of a recording session varied between 2 and 3 hours, depending on how long pauses between blocks and paradigms took a subject.

Paradigm 1: auditory evoked potentials

In the auditory paradigm (Fig. 2A), we presented short stimuli by a headset every 600 ms. The duration of the tone was 200 ms and interstimulus interval was 400 ms. Stimuli in the first half of the paradigm were pure tones at the frequency 1 kHz, and stimuli in the second half of the paradigm were white-noise audio signals. The paradigm required no response, but only to passively hear tones. Therefore we

asked subjects to close their eyes. We assumed that it should help subjects to better concentrate on the task, relax their eyes, and avoid eye-blinks. 1000 stimuli in the paradigm were equally distributed across four blocks. The peak-to-peak amplitude of the ERP components N1 and P2 defined the dependent variable in this paradigm.

Paradigm 2: steady-state visually evoked potential (SSVEP)

In this paradigm (Fig. 2B) we presented an alternating contrast checkerboard on a monitor (visual angle: 3.5x3.5°) at the frequency of 12 Hz. 2400 alternations were equally distributed across four blocks. The paradigm required no response, but only to passively observe the flickering checkerboard. A 4x4-pixel cross in the center of the checkerboard served as an anchor for the fixation point. Minimum to maximum peak difference of the evoked potential defined the dependent variable in this paradigm.

Paradigm 3: motor potentials (MPs) elicited by voluntary tapping

In this paradigm (Fig. 2C) we asked subjects to rhythmically press the key Arrow Down on a keyboard with the index finger of the right hand with a constant pace of about once per second. At the start of each block, subjects heard a guiding pace through a headset, which consisted of 10 beeps played once per second. After that, subjects could start to press the key Arrow Down on the keyboard. We asked subjects to close their eyes. We assumed that it should help subjects to better concentrate on the task, relax their eyes, and avoid eye-blinks. The paradigm consisted of three blocks. In each block, subjects had to conduct 240 button presses. Thus we obtained 720 button presses in this paradigm. The peak-to-peak amplitude of the ERP components Motor Potential (MP) and Reafferent Potential (RAP) defined the dependent variable in this paradigm.

Paradigm 4: visual mismatch negativity (vMMN)

In this paradigm (Fig. 2D) we presented sequentially standard (red) and deviant (yellow) circles on a monitor (visual angle: 2x2°) with a deviant-to-standard stimuli ratio of 1 to 4. We spaced deviant stimuli according to Poisson distribution (Ord, 1967), while we set the minimal number of standard stimuli in between to 1 and the maximal number of standard stimuli in between to 8. We asked subjects to look and concentrate their attention at the blinking circles. The duration of a stimulus was equal to 280 ms. An interval between stimuli onsets was randomly selected for each new stimulus from intervals of 630 ms to 830 ms. 1200 stimuli in the paradigm were equally distributed across four blocks. The difference-amplitude peak of subtracted ERPs (deviant stimuli ERP minus standard stimuli ERP) at around 200 ms after a stimulus onset defined the dependent variable in this paradigm.

Paradigm 5: face-sensitive N170 component

In the paradigm (Fig. 2E), we presented pictures of faces, cars and noise in a random sequence on a monitor (visual angle: 3.1x3.5°). Two sets of 43 colored photographs of full front faces (21 males) and cars were adopted from a different study (Rossion & Caharel, 2011). “Faces were presented without glasses, facial hair or make-up, and with neutral expression. All face pictures were trimmed to remove their variable backgrounds, clothing and hairline. Car pictures were also edited to remove background. Noise stimuli were made by scrambling the faces and the cars using a Fourier phase randomization procedure” (Rossion & Caharel, 2011). We asked subjects to look and concentrate their attention on the appearing pictures. The duration of a stimulus was equal to 280 ms. An interval between stimuli onsets was randomly selected for each new stimulus from intervals of 630 ms to 830 ms. 1032 stimuli in the paradigm, with a ratio of face, car, and noise pictures equal to 1:1:1, were equally distributed across four blocks. The difference-amplitude peak of subtracted ERPs (face stimuli ERP minus car stimuli ERP) at around 170 ms after a stimulus onset defined the dependent variable in this paradigm.

Paradigm 6: visual decision-making N240 component (vDM)

In the paradigm (Fig. 2F), we presented text labels with the German name of a color on a monitor (visual angle: 1.5x0.6°). We used four text labels: rot (red), grün (green), blau (blue), and gelb (yellow), with a ratio of 1:1:1:1. A sequence of the text labels was random, but no two identical color names in a row. The color of the font was always black and presented on a gray screen. We asked subjects to look and concentrate their attention at the appearing text labels and to press the key Arrow Down on the keyboard with the index finger of the right hand, only when the text label “rot” (red) appears on the monitor. Other text labels did not require any action. The duration of a stimulus was equal to 280 ms. The interval between stimuli onsets was 1 second long. In the case of an error response, subjects got feedback: a white flash of the full screen (duration of 150 ms) and extended interval to the next stimulus onset (two seconds). 960 stimuli in the paradigm were equally distributed across four blocks. The difference-amplitude peak of subtracted ERPs (target stimuli ERP minus distractor stimuli ERP), at around 240 ms after a stimulus onset, defined the dependent variable in this paradigm.

PHYSIOLOGICAL METHODS

EEG systems and data acquisition

We recorded electroencephalographic (EEG) data using four EEG systems (Table 1 & Fig. 3B): (1) asalab™ (ANT Neuro, Enschede, Netherlands; www.ant-neuro.com); (2) actiCAP (Brain Products GmbH; www.brainproducts.com); (3) g.tec's g.Nautilus; and (4) Emotiv EPOC. We conducted EEG recordings in batches: first, we conducted a batch of recordings for all subjects with g.Nautilus, then with actiCAP, then with asalab™, and the last batch of recordings with Emotiv.

asalab™ had 127 data channels and gel-based Ag/AgCl electrodes (waveguard™ caps). Electrodes were positioned according to the 10-5 international system (Oostenveld & Praamstra, 2001). The recording reference was Cz and the ground electrode laid on the left clavicle area. We kept scalp impedances below 10 kΩ. The sampling rate of continuously recorded EEG data was 1024 Hz. Neither band-pass nor notch filter was applied. The system transmits EEG data via a shielded electrical cord.

actiCAP had 64 data channels and gel-based Ag/AgCl active electrodes. Electrodes were positioned according to the equidistant spherical montage. The recording reference was Cz and the ground electrode laid near Fz. We kept scalp impedances below 10 kΩ. The sampling rate of continuously recorded EEG data was 1000 Hz. Neither band-pass nor notch filter was applied. The system transmits EEG data via an electrical cord.

g.Nautilus had 32 data channels with g.SAHARA dry electrode technology and wireless data transmission. A receiver situated in proximity to the EEG system. Electrodes were positioned according to the 10-20 international system. The recording reference was on the right ear and the ground electrode AFz. The EEG system had dry active electrodes and high impedance amplifiers. Internal impedance check was performed automatically via software. The sampling rate of continuously recorded EEG data was 500 Hz. Neither band-pass nor notch filter was applied. g.Nautilus wirelessly transmits data via the 2.4 GHz band (Bluetooth). We shifted all timestamps (event triggers) in recordings with g.Nautilus for 11 ms earlier, since we encountered the constant delay in comparison to asalab™ and actiCAP research-grade EEG systems.

Emotiv EPOC had 14 data channels, saline-based electrodes, and wireless data transmission. A receiver situated in proximity to the EEG system. Electrodes were positioned according to the 10-20 international system. The recording references in the CMS/DRL noise cancellation configuration were P3/P4 electrodes. Sensors were adjusted until connectivity reached the "green" level, indicating that the impedance level required by the software was reached (the Emotiv EPOC software development

kit was used). The sampling rate of continuously recorded EEG data was 128 Hz. Emotiv EPOC has a built-in band-pass filter of 0.2 – 43Hz and notch filters at 50 Hz and 60 Hz and wirelessly transmits data via the 2.4 GHz band (Bluetooth).

We noticed an issue of timestamps of event triggers with the Emotiv EPOC EEG system. Timestamps stored in the EEG data were highly susceptible to jitter and delay (Hairston et al., 2014). To overcome this issue we synchronized biased timestamps of the event triggers in EEG data with true timestamps of the same event trigger logged by a presentation script during the recordings. Precise synchronization at different recording devises is an important aspect because residual jitter might contribute to a reduction of ERP amplitude, specifically of high frequency.

Raw EEG data from all systems were band-pass filtered from 1 Hz to 43 Hz (Matlab function firfilt.m) to remove drift and to match the Emotiv EPOC's built-in filter. Then we applied the standardized preprocessing PREP pipeline (Bigdely-Shamlo, Mullen, Kothe, Su, & Robbins, 2015), which removed line-noise, robust referenced the signal, and interpolated bad channels. "The robust average reference procedure tries to estimate the true average of the EEG channels after removing contamination by bad channels. The robust referencing approach produces the same results as average referencing if there are no bad channels". However, "researchers should proceed with caution when there are not enough channels to cover the head for accurate channel interpolation" (Bigdely-Shamlo et al., 2015).

Electrode positions digitizer: xensor™

Before each EEG recording, we digitized the 3D locations of all electrodes and three major fiducials (nasion, left and right preauricular points) using the optical ANT Neuro xensor™ system (ANT Neuro, Enschede, Netherlands; www.ant-neuro.com/products/xensor). Thus, we collected an individual electrodes digitization for each EEG recording.

System	Manufacturer	Electrode type and material	Reference location	Ground location	Number of data channels	Sampling rate (Hz)	Wireless transmission
asalab™	ANT Neuro	Gel-based Ag/AgCl	Cz	Left clavicle area	127	1024	None
actiCAP	Brain Products	Active gel-based Ag/AgCl	Cz	Near Fz	64	1000	None
g.Nautilus	g.tec	Active dry gold-alloy coated (g.SAHARA)	Right earlobe	AFz	32	500	2.4 GHz band
EPOC	Emotiv	Saline-infused felt	P3 (or mastoid)	P4 (or mastoid)	14	128 (2048 internal)	Bluetooth Smart 2.4 GHz

Table 1. General properties of the EEG systems.

INTERPOLATION OF ELECTRODE POSITIONS

Positions of electrodes did not overlap for all the different EEG systems. For example, asalab™ had over a hundred electrodes, and Emotiv EPOC had only fourteen. Additionally, actiCAP had a spherical montage but Emotiv EPOC and g.Nautilus used the 10-20 international system, and asalab™ used the 10-5 international system. The different electrode locations did not allow us to directly compare activity from electrodes in the same places across the different EEG systems. Moreover, subjects had different head sizes, but not all systems had multiple cap sizes. Emotiv EPOC and g.Nautilus had only one cap size (these models had wires covered in a resilient plastic). This could have resulted in possibly different brain areas under the electrodes from different EEG systems. Additionally, for the same subject and the same EEG system, there could have been shifts of a few centimeters across data collection sessions due to a bias in the placement procedure of the cap.

To overcome these issues, we interpolated channel positions (and their activities) over a mesh-head model, taken from EEGLAB (colin27headmesh.mat) that consisted of 1082 mesh points. After the interpolation (EEGLAB function *headplot.m*), we calculated the activity of 1082 mesh channels. Mesh points are depicted as white dots on the head model in Figure 3. Black labels represent the 10-20 international system and depict positions of 35 electrodes. The procedure for each recording session consisted of two steps: (1) The first step is to project positions of electrodes digitized with xensor™ onto a mesh-head model. This step was necessary because the mesh-head model and subjects' head forms were not exactly the same. (2) The second step was to interpolate electrode positions and

activity of related channels over the positions of 1082 mesh points (mesh channels). After that, the 1082 mesh channels had the same positions on the mesh-head model in all EEG recordings in the study. Therefore, we could compare selected mesh channels between all EEG recordings in the study.

For each paradigm we found a region of interest and selected a cluster of interpolated mesh channels (see definition of clusters for the paradigms below in the results). Due to the varying density of electrodes across EEG systems, some clusters in a part of recordings may lack of electrodes within the clusters. Table 2 demonstrates average distances from clusters to nearest electrodes.

	AEPs	SSVEP	MPs	vMMN	N170	vDM
asalab (127 chan.)	0.0	0.0	0.0	0.0	0.0	0.0
actiCAP (64 chan.)	0.4	0.0	0.3	0.0	0.1	0.1
g.Nautilus (32 chan.)	1.3	0.0	0.7	0.3	0.7	0.0
Emotiv EPOC (14 chan.)	1.4	2.4	0.5	1.0	1.3	0.8

Table 2. Average distances between a cluster of selected mesh channels in a paradigm and the nearest electrode to the cluster in a recording session. The values are given in centimeters and averaged over 12 recording sessions. Electrode positions were digitized in each recording session.

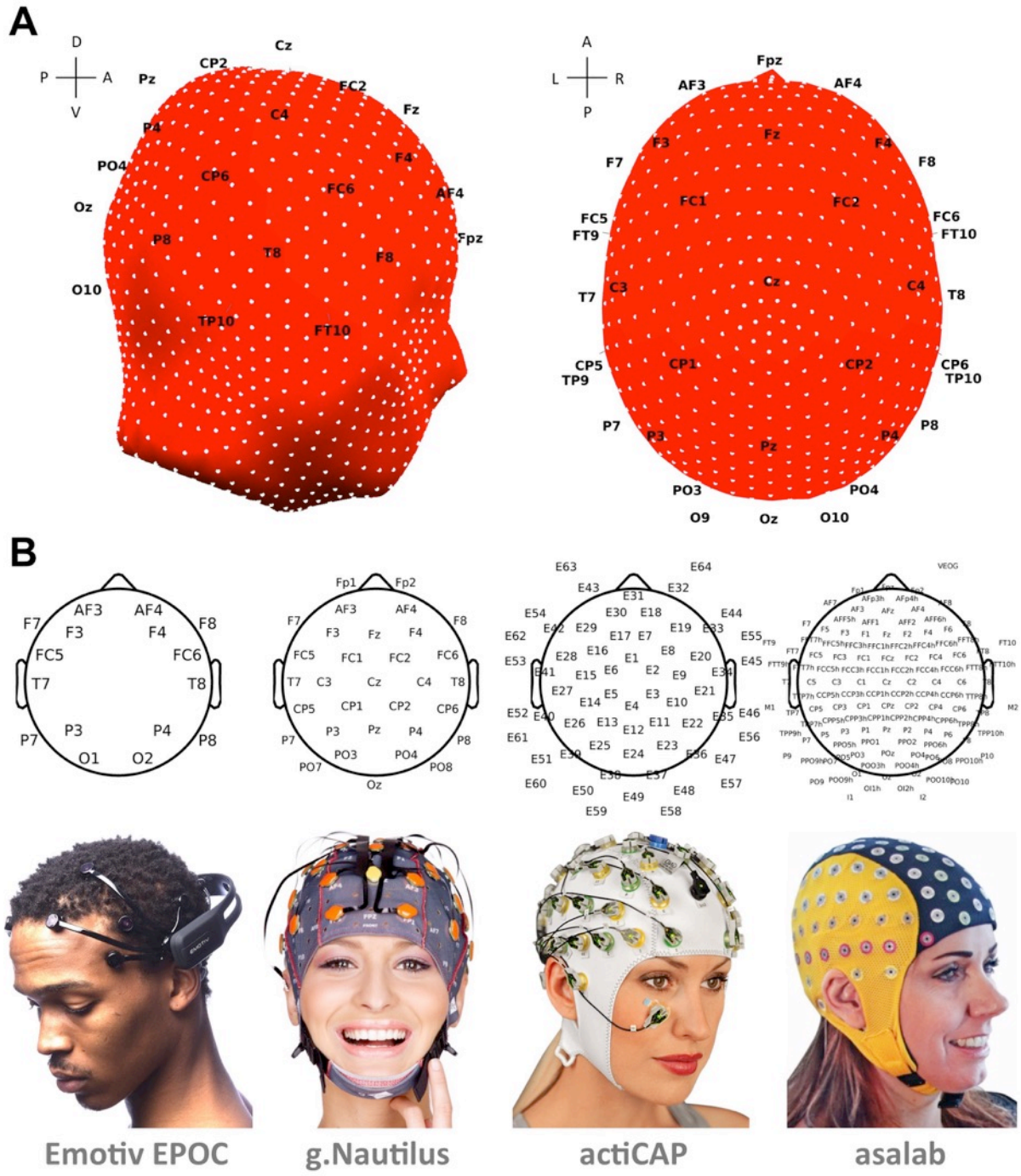


Figure 3. (A) Mesh-head model with the 1082 mesh channels shown as white dots. Black labels represent the 10-20 international system and depict positions of 35 electrodes. Labels at the crosses represent anterior (A), posterior (P), dorsal (D), ventral (V), left (L), and right (R) sides. **(B)** Allocation of electrodes. Beneath the topoplots are photos of the corresponding EEG systems used in the study. [Photos with permission from Emotiv, g.tec, Brain Products, and ANT Neuro]

STATISTICAL ANALYSIS

ANOVA

Repeated measurements of a paradigm recorded with different subjects, systems, or in different sessions encompass some variance related to these factors. We wanted to know to what extent each of these factors influenced the variance. The analysis of variance (ANOVA) approach offers itself to statistically estimate which factor influenced EEG data the most. Therefore, we used a 3-Way ANOVA of factors: System, Subject and Session and their two-factor interactions (Matlab function anovan.m, Sum of Squares Type III). ANOVA is from the family of linear models, and allows to model which of the different factors (Subject, System, Session) explains how much of the variance. The Sum of Squares (SS) value served as a measure of variance explained by the factor. ANOVA is a rather general approach which should be used before a more specific modeling is done.

We wanted to know whether repeated recordings introduced some variability. Therefore, we split a paradigm in each EEG recording into eight equal parts and analyzed each part individually. The specific number eight was a trade-off between a larger number of parts we needed and the length of an individual part, which is a function of the amount of data available. We estimated that eight parts were a reasonable compromise. The division into eight parts was not used to increase the number of available data, but only to have an estimate of the variance within one paradigm in a recording. Thus, we gained eight measurements of a dependent variable per paradigm and EEG recording. As we had 48 EEG recordings, we yielded 384 such measurements per paradigm for the statistical analysis ANOVA.

We did not want to be dependent on one specific paradigm. Therefore, we selected a whole range of paradigms frequently used in the literature and which covered a range of cognitive processes. We treated each paradigm itself like an own experiment and analyzed the sources of variance independently from all the others paradigms. Therefore, there is no interaction between ANOVAs for different paradigms.

In Paradigm 1 we measured for each part (119 trials per part) a peak-to-peak amplitude N1 – P2 (Fig. 4 & Fig. 5). The red numbers in Fig. 5 indicate peak-to-peak amplitudes N1 – P2 for each part. The search interval for the peaks was from 100 ms to 200 ms relative to a stimulus onset.

In Paradigm 2 we measured for each part (270 trials per part) a peak-to-peak amplitude (minimum-to-maximum) of a SSVEP (Fig. 6). The search interval for the peaks was from 120 ms to 220 ms relative to a stimulus onset.

In Paradigm 3 we measured for each part (85 trials per part) a peak-to-peak amplitude MP – RAP (Fig. 7). The search interval for the peaks was -40 ms to 140 ms relative to a stimulus onset.

In Paradigm 4 we derived two ERPs from each part: deviant stimuli ERP (25 trials per part) and standard stimuli ERP (100 trials per part). We measured a negative peak amplitude of a resulting subtraction of deviant stimuli ERP minus standard stimuli ERP for each part (Fig. 8). The search interval for the negative peak was 150-220 ms relative to a stimulus onset. We used the negative peak amplitude as the dependent variable.

In Paradigm 5 we derived two ERPs from each part: face stimuli ERP (40 trials per part) and car stimuli ERP (40 trials per part). We measured a negative peak amplitude of a resulting subtraction of face stimuli ERP minus car stimuli ERP for each part (Fig. 9). The search interval for the negative peak was 130-190 ms relative to a stimulus onset. We used the negative peak amplitude as the dependent variable.

In Paradigm 6 we derived two ERPs from each part: target stimuli ERP (25 trials per part) and distractor stimuli ERP (80 trials per part). We measured a negative peak amplitude of a resulting subtraction of target stimuli ERP minus distractor stimuli ERP for each part (Fig. 10). The search interval for the negative peak was 200-280 ms relative to a stimulus onset. We used the negative peak amplitude as the dependent variable.

Despite the differences in sampling rates of the four EEG systems, interpolation to a common sampling frequency was not required, since we used search intervals to measure amplitude peaks. However, to build grand totals in Figures 4, 6, 7, 8, 9, and 10, we did interpolation to a common sampling frequency of 1000 Hz for all systems.

Post-hoc test

To estimate which levels of a factor were significantly different from each other and therefore explain the variance within the factor, we conducted the post-hoc multiple comparison test (Matlab function *multcompare.m*).

3.4 Results

PARADIGMS

Paradigm 1: auditory evoked potentials

Late-latency auditory evoked potentials (AEPs) “beginning with P1 (which is sometimes classified as middle-latency) at about 80 ms through to N2 at about 250 ms, all are cortical in origin and maximal in amplitude at the central top of the scalp” (Kraus & Nicol, 2009). The 3D map in Fig. 4B demonstrates the scalp distribution of the grand total ERP at 176 ms after the stimulus onset (P2 component) in the current study. Seven white dots on the top of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster is [0 37 85] in Montreal Neurological Institute (MNI) coordinates, the radius of the cluster is 20 mm. All further results in this paradigm were based on the average activity of the selected mesh channels of the cluster. In order to nullify possible influences of baseline position on P2 amplitude value, we selected the peak-to-peak amplitude of the components N1 and P2 as a dependent variable in the paradigm.

A comparison of ERPs taken from scientific literature and the current study is shown in Fig. 4A, which demonstrates similarity of these ERPs. Early auditory components of the ERP in the current study were smoothed, however, the most prominent components N1 and P2 have nearly identical forms. Grand average ERPs per level of a factor in Fig. 4C, 4D & 4E reveal that the peak-to-peak amplitude of the components N1 and P2 vary the most by the factor Subject, less by the factor System, and vary the least by the factor Session.

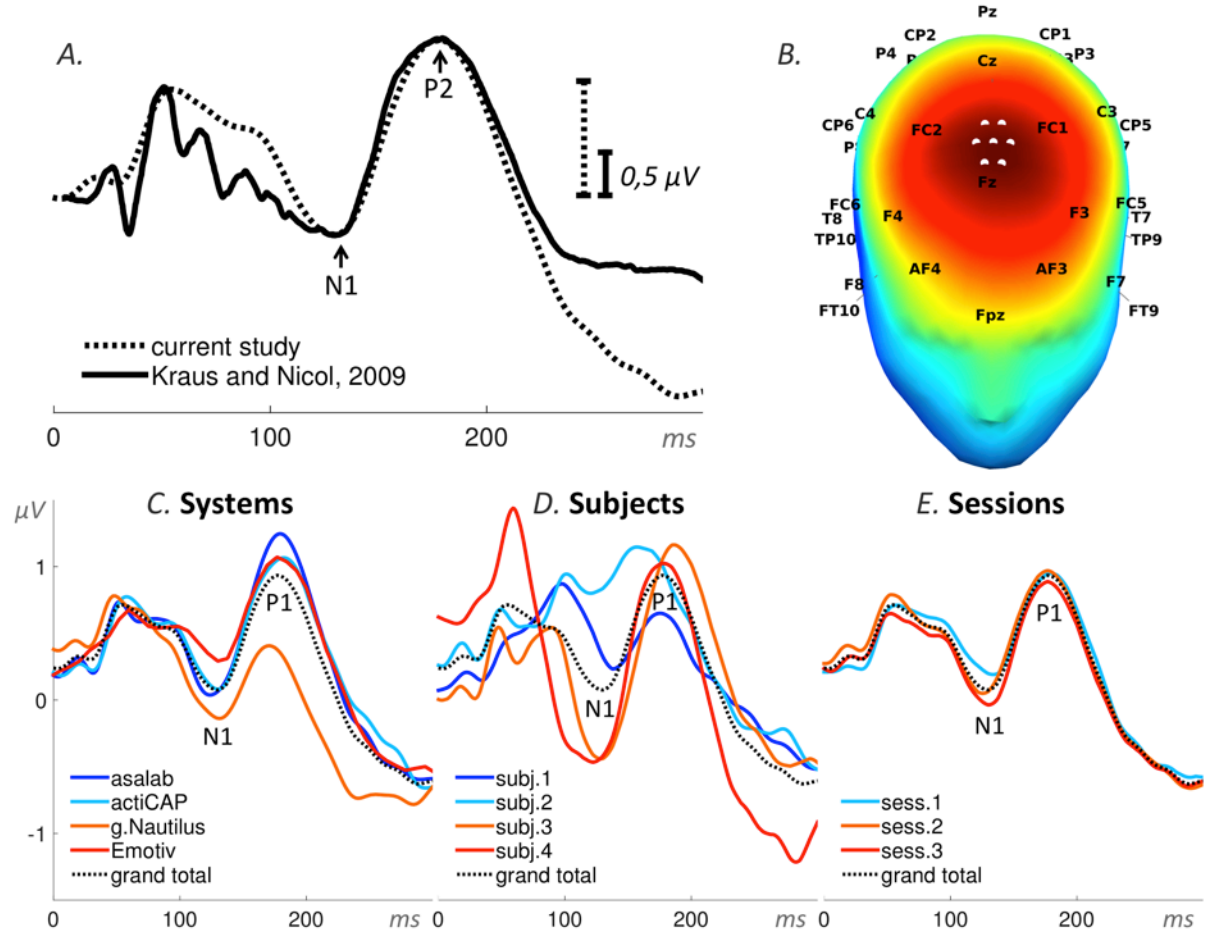


Figure 4. Auditory evoked potential.
A: superposition of an ERP from the literature (Kraus & Nicol, 2009) and the grand total ERP of all 48 sessions in the current study. We shifted the ERP from the literature by 16 ms to the right and reduced the voltage scale by 2.66 times, in order to find the best superposition of these two ERPs.
B: 3D map of the grand total ERP in the current study at 176 ms after the stimulus onset (P2 component in panel A). Seven white dots on the top of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster was [0 37 85] (MNI coordinates), the radius of the cluster was 20 mm. Black labels represent the 10-20 international system.
C, D & E: grand average ERPs of the factors specified in the title above each panel and derived from the cluster of selected mesh channels (shown in panel *B*). Each ERP (colored lines) in panels *C & D* is based on 12 sessions grouped by a level of a factor and in panel *E* on 16 sessions, e.g. the dark blue ERP in panel *C* represents the grand average ERP of 12 sessions recorded with the asalab™ EEG system and with four subjects, each recorded three times. The black dotted ERP in each panel is the grand total ERP of all 48 sessions in the study.

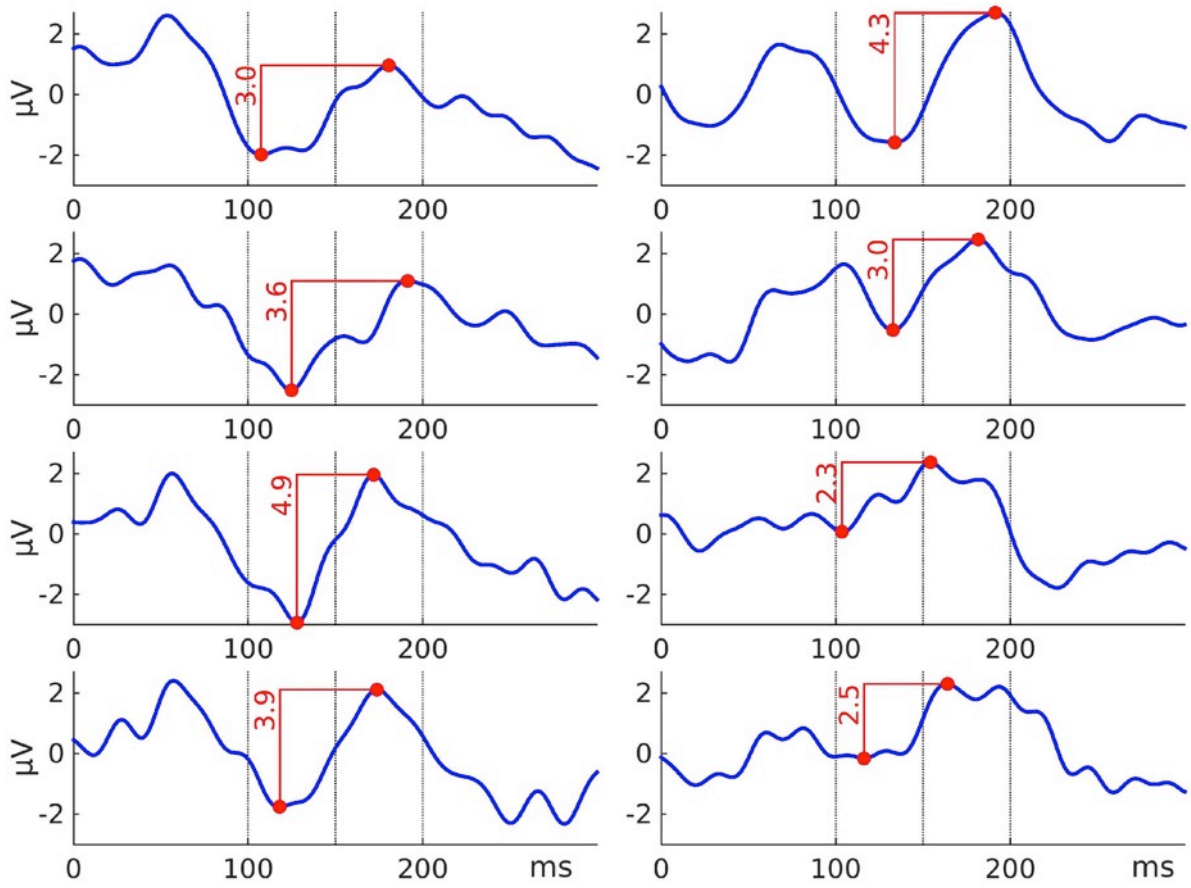


Figure 5. Eight ERPs yielded from eight sequential parts of a recording session (in the figure: asalab™, subject 4, session 3). Each ERP derived from 119 trials. Vertical dotted lines indicate intervals used for finding minimum (N1) and maximum (P2) peaks in the paradigm on intervals 100-150 ms and 150-200 ms after a stimulus onset, respectively. The red numbers indicate peak-to-peak amplitudes (N1 – P2) for each part. We processed 48 recordings and yielded 384 such peak-to-peak amplitude values of the paradigm.

The ANOVA results found statistical significance for these two factors: System (SS = 28.0, F = 23.7, p<.01) and Subject (SS = 150.5, F = 127.3, p<.01). The factor Session was not statistically significant in the ANOVA results (SS = 1.0, F = 1.3, p = 0.29). One two-factor interaction was statistically significant: System*Subject (SS = 15.9, F = 1.8, p<.01). Values of SS_{Error} and SS_{Total} were equal to 139.5 and 341.6 respectively. The ratio of SS_{System}, $SS_{Subject}$, and $SS_{Session}$ was 16%, 83%, and 1%, respectively (Fig. 11). ANOVA results in the paradigm revealed that the factor Subject was the biggest source of variance, relative to other factors.

See post-hoc-test summary for all paradigms in Fig. 13. In the factor System (Fig. 13A) with four levels (asalab™, actiCAP, g.Nautilus, Emotiv), asalab™'s mean value

was significantly different from g.Nautilus's and Emotiv's mean values; actiCAP's mean value was significantly different from g.Nautilus's mean value; g.Nautilus's mean value was significantly different from asalab™'s, actiCAP's, and Emotiv's mean values; Emotiv's mean value was significantly different from g.Nautilus's and asalab™'s mean values. In the factor Subject (Fig. 13B) with four levels (subject 1, subject 2, subject 3, subject 4), subject 1's mean value was significantly different from the mean values of subjects 3 and 4; subject 2's mean value was significantly different from the mean values of subjects 3 and 4; subject 3's mean value was significantly different from the mean values of subjects 1 and 2; subject 4's mean value was significantly different from the mean values of subjects 1 and 2. In the factor Session (Fig. 13C) with three levels (session 1, session 2, session 3), mean values of all three levels of the factor were not significantly different from each other. The post-hoc test revealed (Fig. 13) that two research-grade EEG systems (asalab™ and actiCAP) demonstrated similar results to each other. Emotiv and g.Nautilus demonstrated significantly different results from one or both of the research-grade EEG systems, respectively. Subjects split up in two statistically distinguishable groups: the first group - subjects 1 and 2 and the second group - subjects 3 and 4.

Paradigm 2: steady-state visually evoked potential (SSVEP)

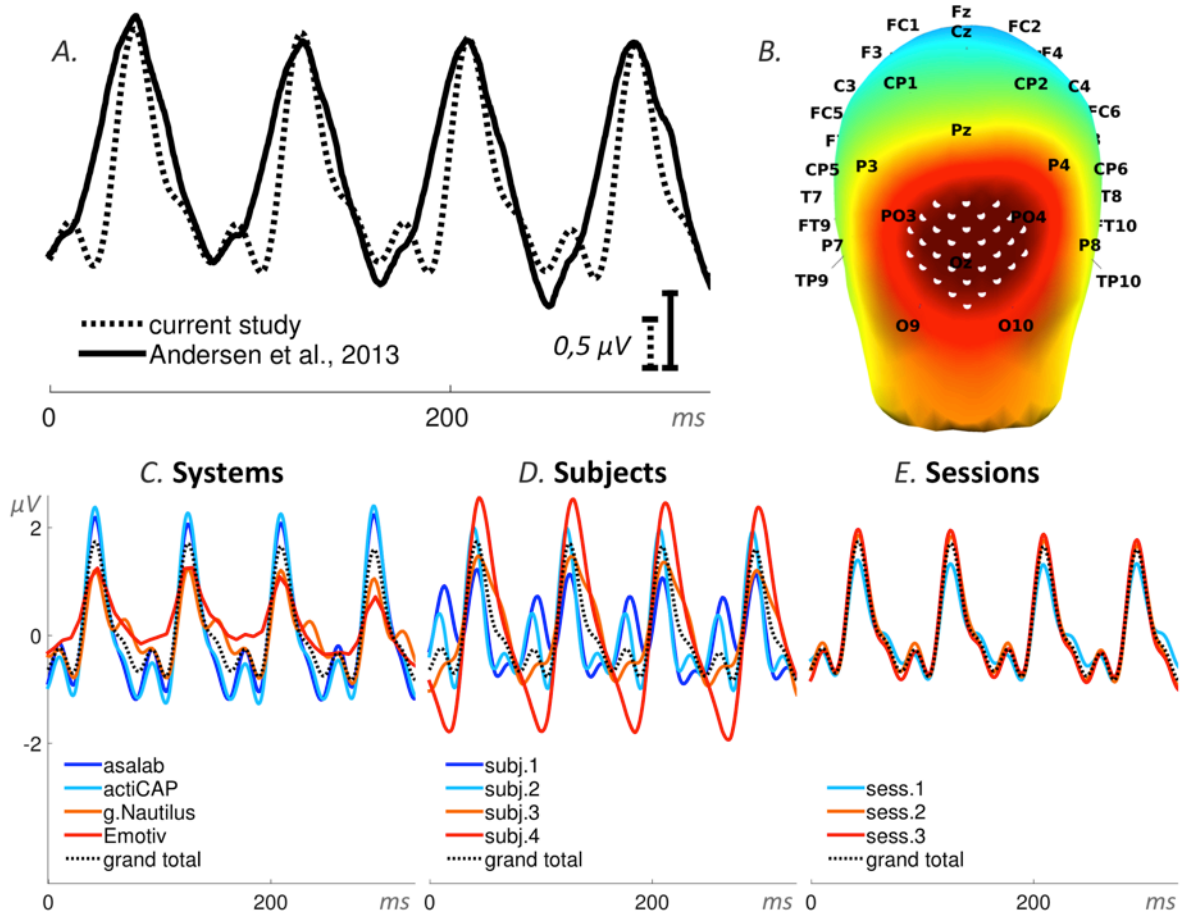


Figure 6. Steady-state visually evoked potential. **A:** superposition of an ERP from the literature (Andersen, Hillyard, & Müller, 2013) and the grand total ERP of all 48 sessions in the current study. We shifted the ERP from the literature by 40 ms to the left and increased the voltage scale by 1.54 times, in order to find the best superposition of these two ERPs. **B:** 3D map of the grand total ERP in the current study at 126 ms after the stimulus onset (second positive peak in panel A). Thirty-one white dots on the back of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster is [0 -118 14] (MNI coordinates), the radius of the cluster is 40 mm. Black labels represent the 10-20 international system. **C, D & E**: grand average ERPs of the factors specified in the title above each panel and derived from the cluster of selected mesh channels (shown in panel *B*). Each ERP (colored lines) in panels *C & D* is based on 12 sessions grouped by a level of a factor and in panel *E* on 16 sessions, e.g. the dark blue ERP in panel *C* represents the grand average ERP of 12 sess. recorded with the asalab™ EEG system and with four subjects, each recorded three times. The black dotted ERP in each panel is the grand total ERP of all 48 sess. in the study.

The ANOVA results found statistical significance for all three factors: System (SS = 205.5, F = 103.6, p<.01), Subject (SS = 289.2, F = 145.8, p<.01), and Session (SS = 6.8, F = 5.1, p<.01). All three two-factor interactions were statistically significant: System*Subject (SS = 97.7, F = 16.4, p<.01), System*Session (SS = 27.1, F = 6.8, p<.01), and Subject*Session (SS = 39.1, F = 9.8, p<.01). Values of SS_{Error} and SS_{Total} were equal to 234.1 and 899.6 respectively. The ratio of SS_{System}, $SS_{Subject}$, and $SS_{Session}$ was 41%, 58%, and 1%, respectively (Fig. 11). ANOVA results in the paradigm revealed that the factor Subject was the biggest source of variance, relative to other factors.

See post-hoc-test summary for all paradigms in Fig. 13. In the factor System (Fig. 13A) with four levels (asalab™, actiCAP, g.Nautilus, Emotiv), asalab™'s mean value was significantly different from Emotiv's mean value; actiCAP's mean value was significantly different from g.Nautilus's and Emotiv's mean values; g.Nautilus's mean value was significantly different from actiCAP's and Emotiv's mean values; Emotiv's mean value was significantly different from asalab™'s, actiCAP's, and g.Nautilus's mean values. In the factor Subject (Fig. 13B) with four levels (subject 1, subject 2, subject 3, subject 4), subject 1's mean value was significantly different from the mean values of subjects 2, 3, and 4; subject 2's mean value was significantly different from the mean values of subjects 1 and 4; subject 3's mean value was significantly different from the mean values of subjects 1 and 4; subject 4's mean value was significantly different from the mean values of subjects 1, 2 and 3. In the factor Session (Fig. 13C) with three levels (session 1, session 2, session 3), mean values of all three levels of the factor were not significantly different from each other. The post-hoc test revealed (Fig. 13) that two research-grade EEG systems (asalab™ and actiCAP) demonstrated similar results to each other. g.Nautilus and Emotiv demonstrated significantly different results from one or both of the research-grade EEG systems, respectively. Subjects split up in three statistically distinguishable groups: the first group - subject 1, the second group - subjects 2 and 3, and the third group - subject 4.

Paradigm 3: motor potentials (MPs) elicited by voluntary tapping

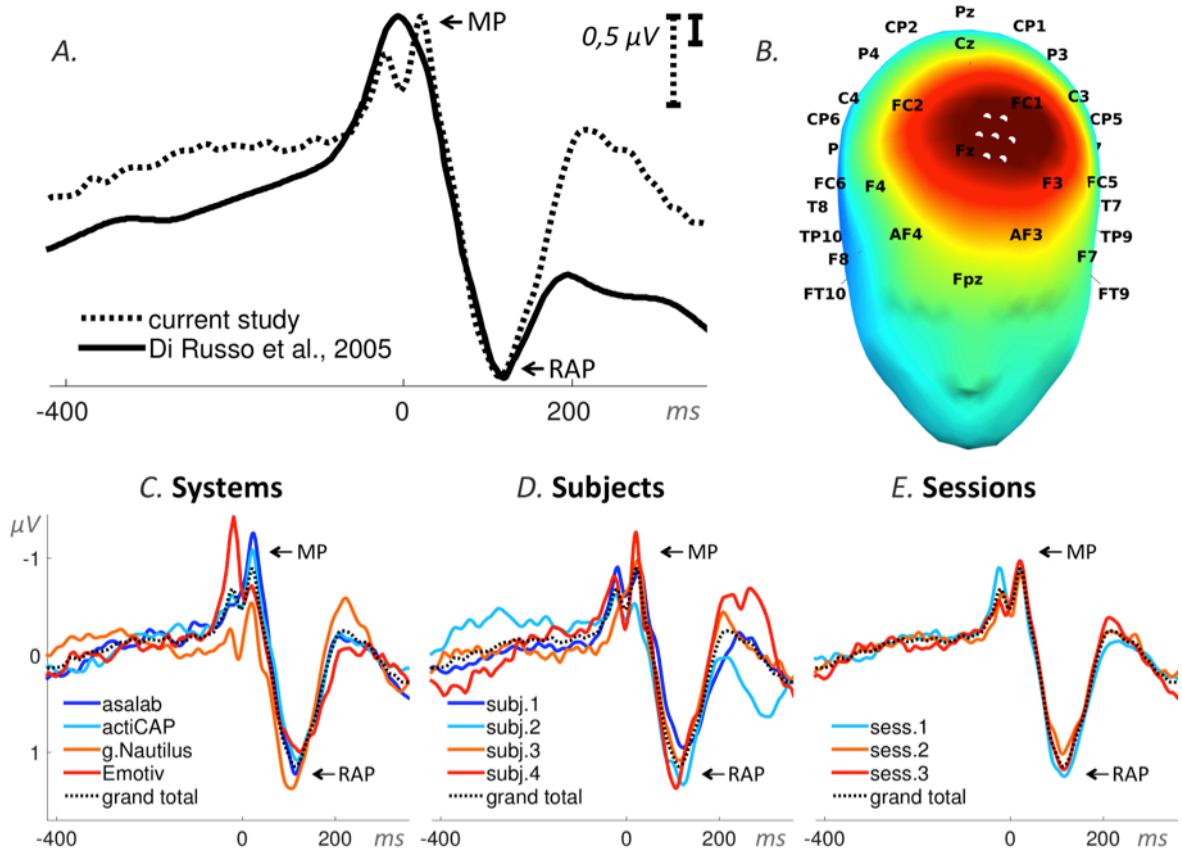


Figure 7. Motor potentials. **A:** superposition of an ERP from the literature (Di Russo, Pitzalis, Aprile, & Spinelli, 2005) and the grand total ERP of all 48 sessions in the current study. We shifted the ERP from the literature by 53 ms to the left and reduced the voltage scale by 3.3 times, in order to find the best superposition of these two ERPs. The peak-to-peak amplitude of the ERP components Motor Potential (MP) and Reafferent Potential (RAP) defined the dependent variable in this paradigm. **B:** 3D map of the grand total ERP in the current study at 115 ms after the stimulus onset (RAP component in panel *A*). Seven white dots on the front top of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster is [-16 48 76] (MNI coord.), the radius of the cluster is 20 mm. Black labels represent the 10-20 international system. **C, D & E**: grand average ERPs of the factors specified in the title above each panel and derived from the cluster of selected mesh channels (shown in panel *B*). Each ERP (colored lines) in panels *C & D* is based on 12 sessions grouped by a level of a factor and in panel *E* on 16 sessions, e.g. the dark blue ERP in panel *C* represents the grand average ERP of 12 sessions recorded with the asalab™ EEG system and with four subjects, each recorded three times. The black dotted ERP in each panel is the grand total ERP of all 48 sessions in the study.

The ANOVA results found statistical significance for all three factors: System (SS = 28.9, F = 15.1, p<.01), Subject (SS = 86.3, F = 45.0, p<.01), and Session (SS = 12.5, F = 9.8, p<.01). All three two-factor interactions were statistically significant: System*Subject (SS = 56.9, F = 9.9, p<.01), System*Session (SS = 29.6, F = 7.7, p<.01), and Subject*Session (SS = 11.3, F = 2.9, p<.01). Values of SS_{Error} and SS_{Total} were equal to 226.5 and 452.0 respectively. The ratio of SS_{System}, $SS_{Subject}$, and $SS_{Session}$ was 23%, 67%, and 10%, respectively (Fig. 11). ANOVA results in the paradigm revealed that the factor Subject was the biggest source of variance, relative to other factors.

See post-hoc-test summary for all paradigms in Fig. 13. In the factor System (Fig. 13A) with four levels (asalab™, actiCAP, g.Nautilus, Emotiv), asalab™'s mean value was not significantly different from mean values of the three other EEG systems; actiCAP's mean value was significantly different from Emotiv's mean value; g.Nautilus's mean value was significantly different from Emotiv's mean value; Emotiv's mean value was significantly different from actiCAP's and g.Nautilus's mean values. In the factor Subject (Fig. 13B) with four levels (subject 1, subject 2, subject 3, subject 4), subject 1's mean value was significantly different from the mean values of subjects 3 and 4; subject 2's mean value was significantly different from the mean values of subjects 3 and 4; subject 3's mean value was significantly different from the mean values of subjects 1, 2, and 4; subject 4's mean value was significantly different from the mean values of subjects 1, 2 and 3. In the factor Session (Fig. 13C) with three levels (session 1, session 2, session 3), session 1's mean value was significantly different from session 2's mean value; session 2's mean value was significantly different from session 1's mean value; session 3's mean value was not significantly different from mean values of the other two sessions. The post-hoc test revealed (Fig. 13) that two research-grade EEG systems (asalab™ and actiCAP) demonstrated similar results to each other. Emotiv demonstrated significantly different results from one of the research-grade EEG system. Subjects split up in three statistically distinguishable groups: the first group - subjects 1 and 2, the second group - subject 3, and the third group - subject 4.

Paradigm 4: visual mismatch negativity (vMMN)

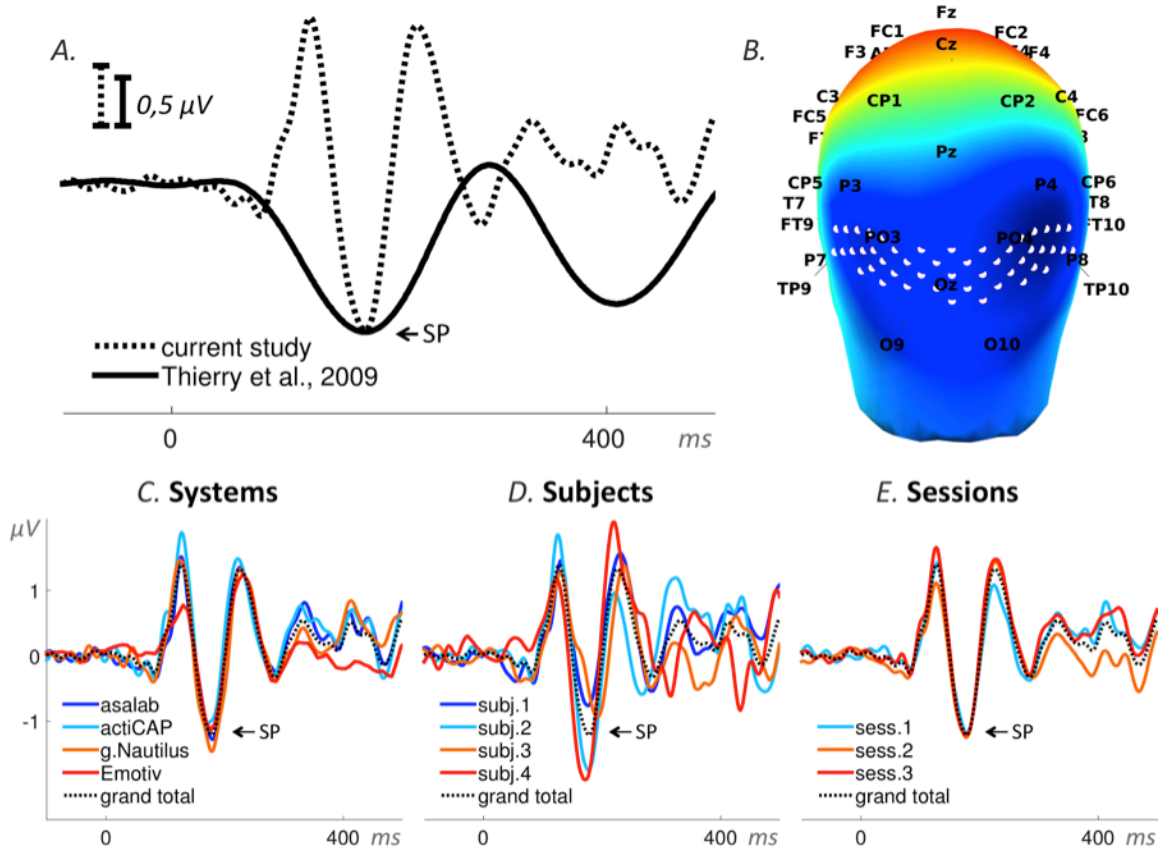


Figure 8. Visual mismatch negativity (ERP waveforms subtraction 'deviant minus standard'). **A:** superposition of an ERP from the literature (Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009) and the grand total ERP (subtraction 'deviant minus standard') of all 48 sess. in the current study. We shifted the ERP from the literature by 10 ms to the left and reduced the voltage scale by 1.25 times, in order to find the best superposition of these two ERPs. **B:** 3D map of the grand total ERP (subtraction 'deviant minus standard') in the current study at 180 ms after the stimulus onset (subtraction peak (SP) in panel *A*). Forty-eight white dots on the back of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster is [0 -118 14] (MNI). Black labels represent the 10-20 international system. **C, D & E**: grand average ERPs of the factors specified in the title above each panel and derived from the cluster of selected mesh channels (shown in panel *B*). Each ERP (colored lines) in panels *C & D* is based on 12 sess. grouped by a level of a factor and in panel *E* on 16 sess., e.g. the dark blue ERP in panel *C* represents the grand average ERP (subtraction 'deviant minus standard') of 12 sess. recorded with the asalab™ EEG system and with four subjects, each recorded three times. The black dotted ERP in each panel is the grand total ERP (subtraction 'deviant minus standard') of all 48 sess. in the study.

The ANOVA results found statistical significance for these two factors: System (SS = 26.1, F = 6.7, p<.01) and Subject (SS = 120.5, F = 31.1, p<.01). The factor Session was not statistically significant (SS = 1.1, F = 0.4, p = 0.66). One two-factor interaction was statistically significant: System*Subject (SS = 39.5, F = 3.4, p<.01). Values of SS_{Error} and SS_{Total} were equal to 457.6 and 660.6 respectively. The ratio of SS_{System}, $SS_{Subject}$, and $SS_{Session}$ was 18%, 81%, and 1%, respectively (Fig. 11). ANOVA results in the paradigm revealed that the factor Subject was the biggest source of variance, relative to other factors.

See post-hoc-test summary for all paradigms in Fig. 13. In the factor System (Fig. 13A) with four levels (asalab™, actiCAP, g.Nautilus, Emotiv), asalab™'s mean value was significantly different from g.Nautilus's mean value; actiCAP's mean value was significantly different from g.Nautilus's mean value; g.Nautilus's mean value was significantly different from asalab™'s, actiCAP's and Emotiv's mean values; Emotiv's mean value was significantly different from g.Nautilus's mean value. In the factor Subject (Fig. 13B) with four levels (subject 1, subject 2, subject 3, subject 4), mean values of all four levels of the factor were significantly different from each other. In the factor Session (Fig. 13C) with three levels (session 1, session 2, session 3), mean values of all three levels of the factor were not significantly different from each other. The post-hoc test revealed (Fig. 13) that two research-grade EEG systems (asalab™ and actiCAP) demonstrated similar results to each other. g.Nautilus demonstrated significantly different results from both of the research-grade EEG systems. All four subjects' had significantly different from each other mean values.

Paradigm 5: face-sensitive N170 component

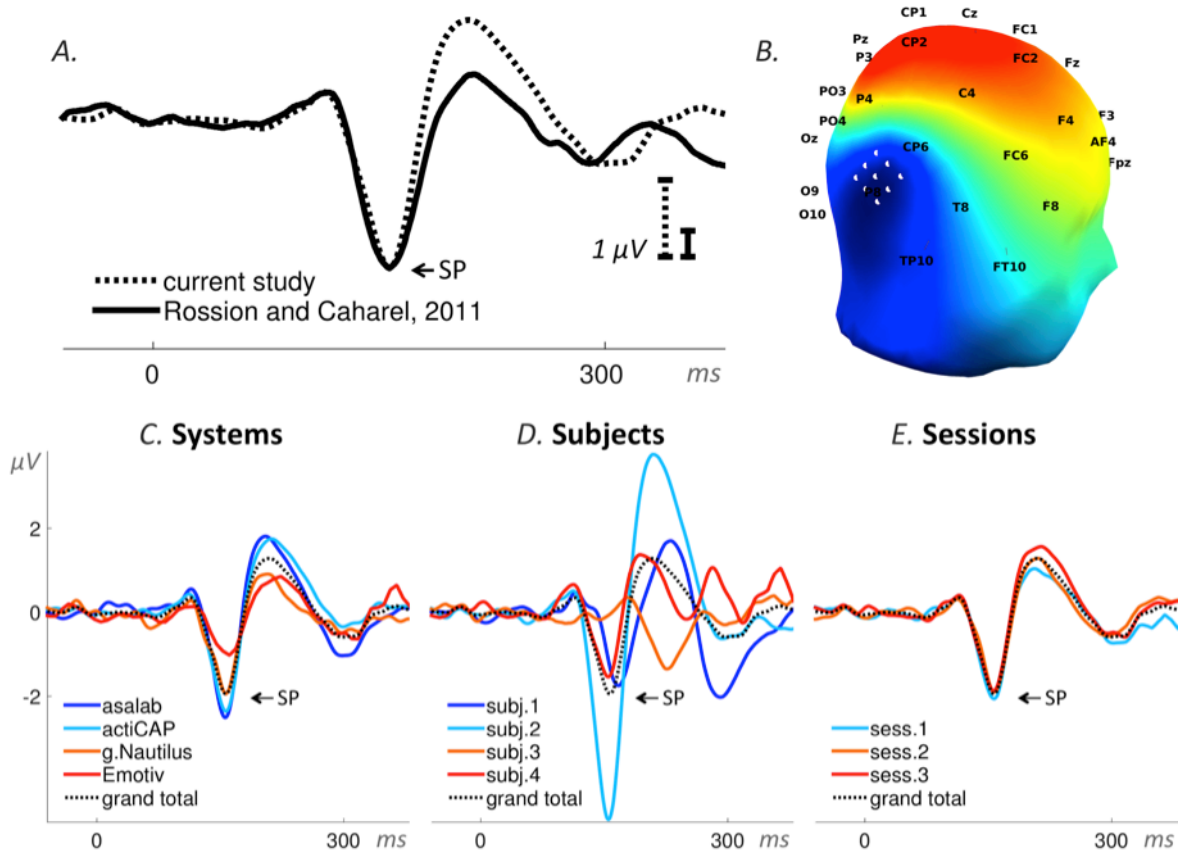


Figure 9. Face-sensitive N170 component (ERP waveforms subtraction 'faces minus cars') **A:** superposition of an ERP from the literature (Rossion & Caharel, 2011) and the grand total ERP (subtraction 'faces minus cars') of all 48 sessions in the current study. We shifted the ERP from the literature by 23 ms to the right and reduced the voltage scale by 2.95 times, in order to find the best superposition of these two ERPs. **B:** 3D map of the grand total ERP (subtraction: face minus car) in the current study at 160 ms after the stimulus onset (subtraction peak (SP) in panel *A*). Nine white dots on the right side of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster is [70 -83 6] (MNI coordinates), the radius of the cluster is 22 mm. Black labels represent the 10-20 international system. **C, D & E**: grand average ERPs of the factors specified in the title above each panel and derived from the cluster of selected mesh channels (shown in panel *B*). Each ERP (colored lines) in panels *C & D* is based on 12 sessions grouped by a level of a factor and in panel *E* on 16 sessions, e.g. the dark blue ERP in panel *C* represents the grand average ERP (subtraction 'faces minus cars') of 12 sessions recorded with the asalab™ EEG system and with four subjects, each recorded three times. The black dotted ERP in each panel is the grand total ERP (subtraction 'faces minus cars') of all 48 sessions in the study.

The ANOVA results found statistical significance for these two factors: System (SS = 120.9, F = 21.0, p<.01) and Subject (SS = 952.2, F = 165.0, p<.01). The factor Session was not statistically significant (SS = 6.7, F = 1.7, p = 0.18). Two two-factor interactions were statistically significant: System*Subject (SS = 132.1, F = 7.6, p<.01) and System*Session (SS = 49.0, F = 4.2, p<.01). Values of SS_{Error} and SS_{Total} were equal to 681.0 and 1955.6 respectively. The ratio of SS_{System}, $SS_{Subject}$, and $SS_{Session}$ was 11%, 88%, and 1%, respectively (Fig. 11). ANOVA results in the paradigm revealed that the factor Subject was the biggest source of variance, relative to other factors.

See post-hoc-test summary for all paradigms in Fig. 13. In the factor System (Fig. 13A) with four levels (asalab™, actiCAP, g.Nautilus, Emotiv), asalab™'s mean value was significantly different from Emotiv's mean value; actiCAP's mean value was significantly different from Emotiv's mean value; g.Nautilus's mean value was significantly different from Emotiv's mean value; Emotiv's mean value was significantly different from asalab™'s, actiCAP's, and g.Nautilus's mean values. In the factor Subject (Fig. 13B) with four levels (subject 1, subject 2, subject 3, subject 4), subject 1's mean value was significantly different from the mean values of subjects 2 and 3; subject 2's mean value was significantly different from the mean values of subjects 1, 3, and 4; subject 3's mean value was significantly different from the mean values of subjects 1, 2, and 4; subject 4's mean value was significantly different from the mean values of subjects 2 and 3. In the factor Session (Fig. 13C) with three levels (session 1, session 2, session 3), mean values of all three levels of the factor were not significantly different from each other. The post-hoc test revealed (Fig. 13) that two research-grade EEG systems (asalab™ and actiCAP) demonstrated similar results to each other. Emotiv demonstrated significantly different results from both of the research-grade EEG systems. Subjects split up in three statistically distinguishable groups: the first group - subjects 1 and 4, the second group - subject 2, and the third group - subject 3.

Paradigm 6: visual decision-making N240 component (vDM)

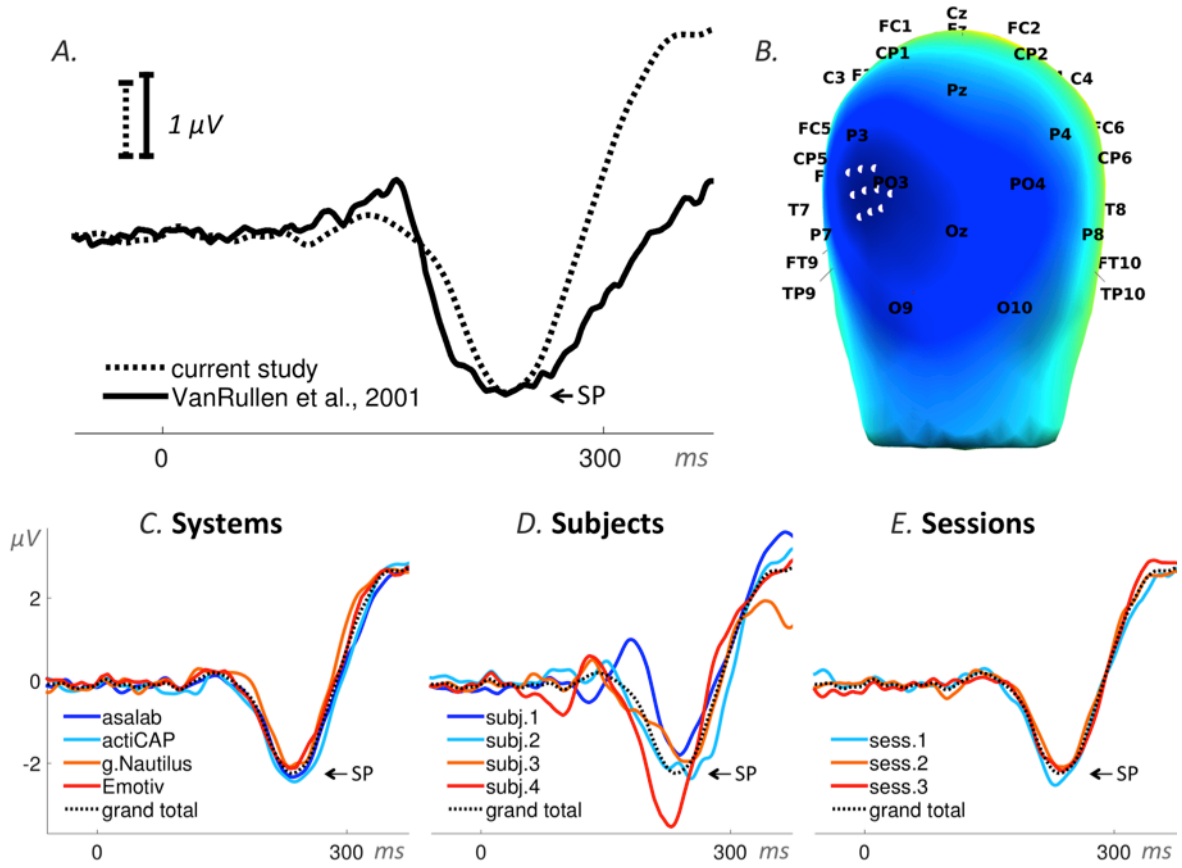


Figure 10. Visual decision-making N240 component (ERP waveforms subtraction 'targets minus distractors'). **A:** superposition of an ERP from the literature (VanRullen & Thorpe, 2001) and the grand total ERP (subtraction 'targets minus distractors') of all 48 sessions in the current study. We shifted the ERP from the literature by 23 ms to the right and increased the voltage scale by 1.11 times, in order to find the best superposition of these two ERPs. **B:** 3D map of the grand total ERP (subtraction 'targets minus distractors') in the current study at 235 ms after the stimulus onset (subtraction peak (SP) in panel *A*). Ten white dots on the left back side of the head depict the cluster of mesh channels selected for further analysis. The central point of the cluster is [-58 -90 28] (MNI coord.), the radius of the cluster is 20 mm. Black labels represent the 10-20 international system. **C, D & E**: grand average ERPs of the factors specified in the title above each panel and derived from the cluster of selected mesh channels (shown in panel *B*). Each ERP (colored lines) in panels *C & D* is based on 12 sessions grouped by a level of a factor and in panel *E* on 16 sessions, e.g. the dark blue ERP in panel *C* represents the grand average ERP (subtraction 'targets minus distractors') of 12 sessions recorded with the asalab™ EEG system and with four subjects, each recorded three times. The black dotted ERP in each panel is the grand total ERP (subtraction 'targets minus distractors') of all 48 sessions in the study.

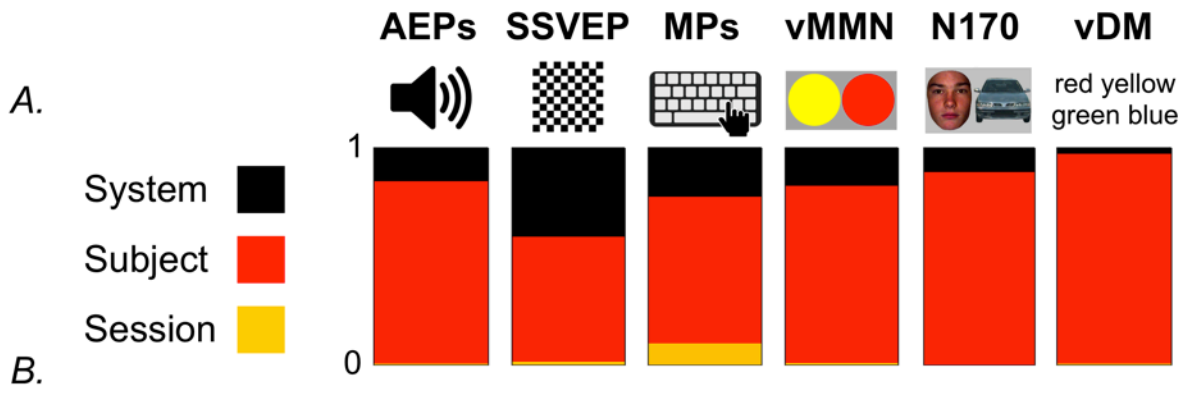
The ANOVA results found statistical significance for the following factor: Subject (SS = 258.0, F = 56.9, p<.01). The factors System (SS = 7.4, F = 1.6, p = 0.18) and Session (SS = 1.1, F = 0.4, p = 0.69) were not statistically significant . One two-factor interaction was statistically significant: Subject*Session (SS = 47.9, F = 5.3, p<.01). Values of SS_{Error} and SS_{Total} were equal to 534.9 and 892.4 respectively. The ratio of SS_{System}, $SS_{Subject}$, and $SS_{Session}$ was 3%, 97%, and 0%, respectively (Fig. 11). ANOVA results in the paradigm revealed that the factor Subject was the biggest source of variance, relative to other factors.

See post-hoc-test summary for all paradigms in Fig. 13. In the factor System (Fig. 13A) with four levels (asalab™, actiCAP, g.Nautilus, Emotiv), mean values of all four levels of the factor were not significantly different from each other. In the factor Subject (Fig. 13B) with four levels (subject 1, subject 2, subject 3, subject 4), subject 1's mean value was significantly different from the mean values of subjects 2, 3, and 4; subject 2's mean value was significantly different from the mean values of subjects 1 and 4; subject 3's mean value was significantly different from the mean values of subjects 1 and 4; subject 4's mean value was significantly different from the mean values of subjects 1, 2, and 3. In the factor Session (Fig. 13C) with three levels (session 1, session 2, session 3), mean values of all three levels of the factor were not significantly different from each other. The post-hoc test revealed (Fig. 13) that two research-grade EEG systems (asalab™ and actiCAP) demonstrated similar results to each other. Subjects split up in three statistically distinguishable groups: the first group - subjects 1, the second group - subjects 2 and 3, and the third group - subject 4.

SUMMARY

ANOVA

The ANOVA results are consistent across the six paradigms and suggest that the factor Subject was the largest source of variance in the study. However, ratio of variation over subjects (Fig. 11) was smallest in the "low-level" tasks like flickering checkerboard (Fig. 11 - SSVEP) and highest in the more cognitive-oriented tasks like decision-making task (Fig. 11 - vDM). The factor System was also a significant source of variance in all paradigms. The factor Session was a relatively small source of variance, and it had a significant *p*-value (p<0.01) only in two paradigms (motor potentials and SSVEP) out of six.



Source	d.f.	AEPs	SSVEP	MPs	vMMN	N170	vDM
System	**3**	**28.0**	**205.5**	**28.9**	**26.1**	**120.9**	**7.4**
Subject	**3**	**150.5**	**289.2**	**86.3**	**120.5**	**952.2**	**258.0**
Session	**2**	**1.0**	**6.8**	**12.5**	**1.1**	**6.7**	**1.1**
System*Subject	9	15.9	97.7	56.9	39.5	132.1	31.4
System*Session	6	4.1	27.1	29.6	11.2	49.0	11.6
Subject*Session	6	2.6	39.1	11.3	4.7	13.7	47.9
Error	354	139.5	234.1	226.5	457.6	681.0	534.9
Total	383	341.6	899.6	452.0	660.6	1,955.6	892.4

Figure 11. ANOVA Sum of Squares summary of six paradigms. **A.** Color bars represent ratios of Sum of Squares of factors: System, Subject, and Session in the paradigms AEPs, SSVEP, MPs, vMMN, N170, and vDM. **B.** ANOVA table summary of Sum of Squares due to each source for the six paradigms. "d.f.": degrees of freedom associated with each source.

Interpolation-related error

We introduced the interpolation step of the physically recorded channels to 1082 mesh channels in order to make the different systems comparable at a selected location. This does, however, raise a question regarding the validity of whether the interpolation step creates abnormal, unrealistic or other spurious data resulting in increased variance between the systems or poorer performance based solely on channel density. To address this, we compared an exemplar original recording by the full set of electrodes with a reduced and interpolated version thereof. Specifically, we used data from the system with the highest electrode count (asalab) and selected channels that were also present in the other systems, i.e. Emotiv EPOC, g.Nautilus, or similar electrodes (actiCAP had a spherical montage) to create an artificially sparse, but matching montage. Based on this sub-selection of electrodes, we interpolated the signals for the region of interest similar to had been performed for all previous analyses (Fig. 4B). We chose the AEP paradigm as a relevant example, as

not all systems have good coverage of the region of interest. Distances between the AEP cluster of mesh channels and the nearest electrodes were 0 cm for the original 127 channels and the 63- and 31-channels down sampling. This means that at least one physical electrode was present right in the region of interest and close to the mesh points. The distance for the 14-channels down sampling was 3.2 cm, which is more than two times greater than the average distance for Emotiv EPOC (1.4 cm, see Table 2).

Figure 12 compares the AEP of the original recording and reduced, interpolated data sets. The high similarity of the evoked potentials demonstrates that here the interpolation step does not introduce relevant additional variance (Fig. 12). Thus, in this case, as the topographic distribution of the evoked potential was relatively smooth, the interpolation could capture the evoked potential even though the distance of physical electrodes to the region of interest was relatively large. However, this does not imply that the number of electrodes is irrelevant. It only shows that the interpolation step does not by itself degrade the data or create artifactual results. The coverage of a system should be suited for a specific experiment.

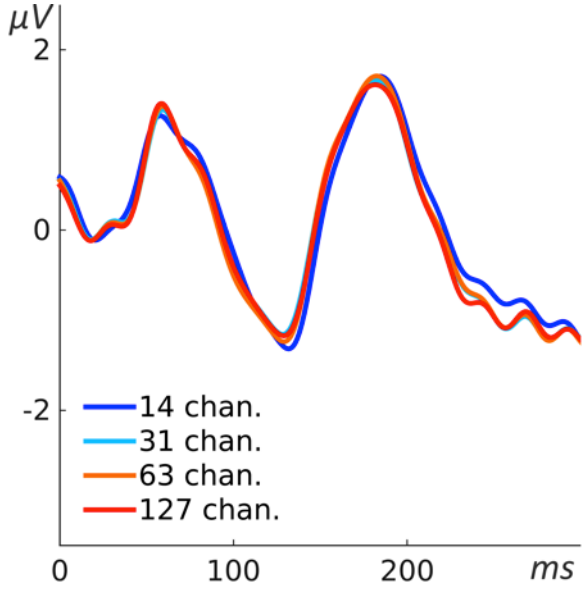


Figure 12. Interpolation-related error. Four ERPs gained from the auditory evoked potential (AEP) paradigm. The red line indicates the original data of one recording session of one subject with the 127 channel system (asalab). The additional lines indicate the evoked potential obtained by downsampling the number of channels to 63 (orange), 31 (light blue) and 14 (dark blue), respectively and interpolating to 1082 mesh-head channels.

Post-hoc test

The post-hoc tests revealed that the two research-grade EEG systems (asalab™ and actiCAP) had no significantly different values from each other in all paradigms (Fig. 13). However, g.Nautilus had significantly different values from one or both of the research-grade EEG systems in 3 out of 6 paradigms, and Emotiv EPOC also had significantly different values from one or both of the research-grade EEG systems in 4 out of 6 paradigms. In 4 out of 6 paradigms, Emotiv EPOC showed a stronger deviation of values, than g.Nautilus. The post-hoc test results also revealed subject differences. Subject 1 had the weakest effect responses (dependent variable) in 5 out of 6 paradigms. Subject 4 had the strongest effect responses (dependent variable) in 5 out of 6 paradigms. These results revealed subject specificity of EEG responses. The post-hoc test results demonstrated that the factor Session influenced EEG results at the very least since, in 5 out of 6 paradigms, all 3 sessions were not significantly different from one another and only in the motor potentials paradigm sessions 1 and 2 were significantly different from each other. We did not observe any systematic increase or decrease of mean amplitude values in sequences of sessions (Fig. 13C). Post-hoc results were congruent with ANOVA results and suggested that the factor Subject was the largest source of variance in the study; the factor System was also the prominent source of variance, however, mostly because of the mobile EEG systems (g.Nautilus and Emotiv EPOC); and the factor Session was the smallest source or variance in the study.

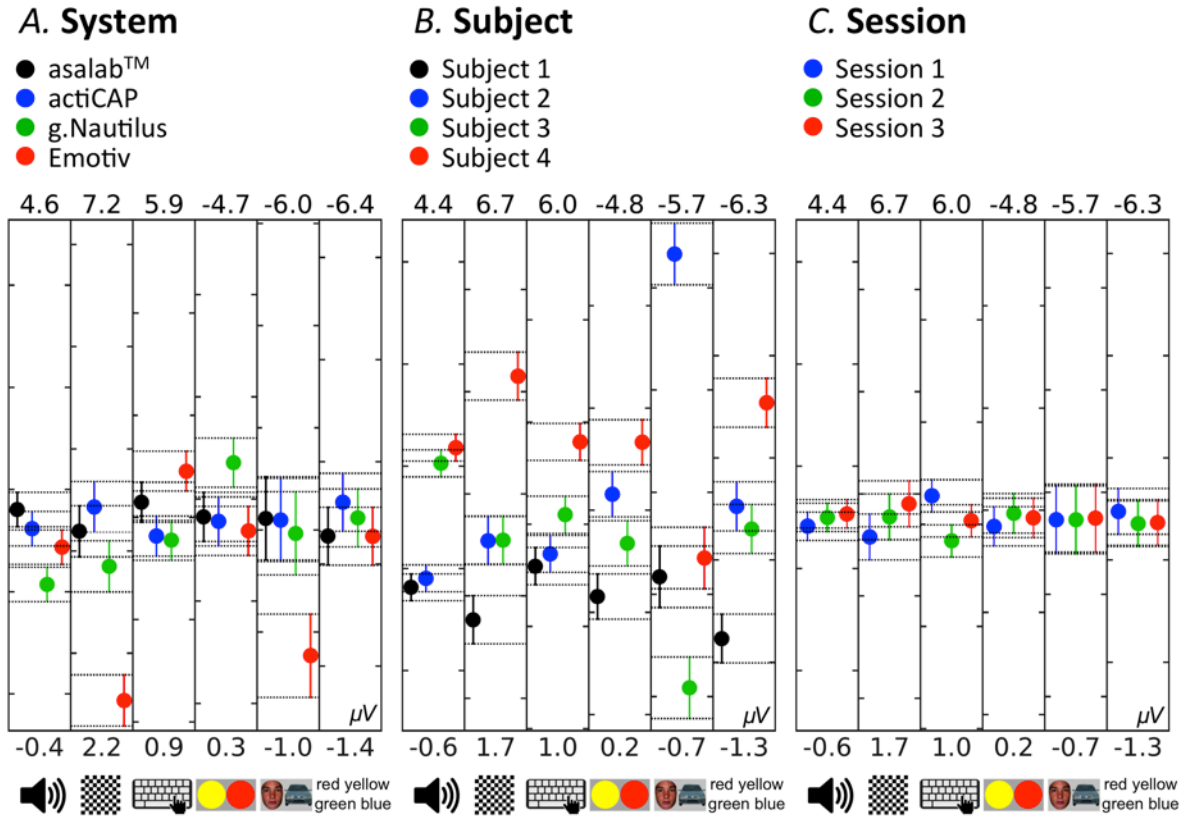


Figure 13. Post-hoc test: 'multiple comparison of means' in the six paradigms. *A*: comparison of levels of the factor System. *B*: comparison of levels of the factor Subject. *C*: comparison of levels of the factor Session. Colored dots in the plots show the mean value of the level of the factor, e.g. the black dot in the first plot from the left is equal to 1.8 μV and represents the mean value of 96 measurements of the dependent variable (12 EEG recordings x 8 measurements per recording) in the paradigm AEPs with the asalab™ EEG system. Error bars and related horizontal dotted lines indicate comparison intervals. If two dots have intersecting comparison intervals, then they are not significantly different, otherwise they are. Values above and below each plot show Y-axis limits of the plot. Additionally, all plots have 5 μV interval between the lower and upper Y-axis limits. The first tree plots in each panel represent paradigms which investigate a peak-to-peak amplitude of an ERP, where a higher positive value correlates with a larger manifestation of an effect, therefore these plots have the normal direction of Y-axis. The last three plots in each panel represent paradigms which investigate amplitude of a subtraction peak of two ERP waveforms, where a higher negative value correlates with a larger manifestation of an effect, therefore these plots have the reverse direction of Y-axis. Thus, the higher a dot in all of these plots is located, the larger the manifestation of the effect in a paradigm is. Pictograms on the bottom of each plot show the paradigm related to the plot.

3.5 Discussion

The present results demonstrate that subjects are the largest source of variance of ERPs in all paradigms. On average, they explain 32% of the total variance. However, systems are also a significant source of variance in all paradigms, on average 9%. In contrast, sessions are a relatively small source of variance, on average 1%. However, this 1% does not mean that ERPs from two sequential recordings with the same EEG system and the same subject were nearly identical. Rather, we did not find any systematic bias of ERP amplitudes, based on the ordinal number of a recording for the same subject and the same system. The two standard research-grade EEG systems had no significantly different means from each other through all paradigms. However, the g.Nautilus with g.SAHARA electrodes and Emotiv EPOC EEG systems used in the current study demonstrated different mean values from one or both of the two standard research-grade EEG systems in at least half of the paradigms.

The variance across systems is smaller than variance across subjects but is not negligible. It is relatively easy to collect more subjects, average over subjects, and reduce the variance. Typically, the number of subjects per study ranges from 10 to 20. Our ANOVA evaluated the variance associated with different independent variables, e.g., Subject, System, and Session. This variance is independent of the number of subjects. If we would pool data from 16 subjects, that would improve the error of uncertainty in estimates of the mean (like ERP) by the square root of the number of subjects (sqrt(16)), thus resulting from current 32% to 8%. It is unusual to average across different systems as most studies do not have access to multiple EEG systems. The uncertainty due to EEG systems (9%) would be of the same order of magnitude as the uncertainty due to subjects (32% / sqrt(16) = 8%), if we would pool data from 16 subjects. Increasing the number of subjects even higher, e.g. to 100, would have progressively diminishing returns, because we are still left with the variance of the systems due to different laboratories using different EEG systems.

The variance across systems may appear due to variability in amplifiers, contact materials, number of electrodes, etc. However, it was not in the focus of the current study to disentangle these interactions and whether a system could be improved by tweaking an individual component. Instead, we assumed that manufacturers designed complete optimized packages, with different parts tailored suitably. Hence, we tested the complete systems as they were supplied by the manufacturer.

We used the most prominent ERPs in each paradigm to calculate the variance due to EEG systems. If we would focus on more subtle ERPs, like middle-latency auditory evoked potentials, the variance related to systems could be higher. Modern EEG paradigms, which allow free movements under natural contexts, investigate prominent sources of EEG activity (Gwin, Gramann, Makeig, & Ferris, 2011), which are usually 'cognitive' in origin and late latency in time (Sur & Sinha, 2009). BCI (Guger et al., 2012; Liu et al., 2012), gaming (Liao et al., 2012), and wellness (Dixit, 2011; Milazzo, Bagade, Banerjee, & Gupta, 2013) use of EEG also rely on prominent sources of electrocortical activity. Investigation of subtle ERPs requires lab conditions with reduced number of noise factors, and advantages of mobile EEG systems are not relevant here. Therefore, it makes sense to compare prominent ERPs across EEG systems.

The variance across sessions, on average 1%, is the smallest source of variance. This 1% means that the systematic effect of knowing which session it is barely helps us to explain the variance. However, this does not mean that there is no variance across sessions. Indeed, adding up Sum of Squares of the three main factors and their interactions do not add up to 100%. Therefore, the residual variance across sessions can substantially contribute to the Error Sum of Squares term. In other words, the variance across sessions would be higher if, for example, subjects were always more anxious in the first than in the second session. Thus, the present results of the variance across sessions suggest that different sessions are interchangeable.

There are other means to compare EEG systems, for example, the correlation of waveforms from two different EEG systems during a simultaneous recording from electrodes located near each other (Gargiulo et al., 2010; Yeung et al., 2015). Many mobile EEG systems have a rigid montage, and it is not possible to overlay montages of two different EEG systems. Moreover, it is problematic to place two electrodes at the same spot, as well as the references. Potential differences can then be attributed either to the systems or to the different positions, which creates the problem. The interpolation approach presented in the current study can address this problem. The approach uses information from several physical electrodes to estimate the ERP at the optimal location as reported in the literature. It allows to have, first, the optimal ERP signal and second, a direct comparison of different EEG systems. To conduct the interpolation, we used standard algorithms supplied by EEGLAB. The interpolation approach allows us to compare virtual channels at same scalp positions directly. However, this approach should be used with caution when there are not enough channels to cover the head for accurate channel interpolation. Therefore, it might be the case that the differences observed for the Emotiv EPOC are still partly due to the lower coverage.

The sensitivity of a given EEG system for a given source inevitably depends on the location of the electrodes. Different systems are developed for different purposes and are not necessarily generic in their usage; certainly a high electrode count is a suitable way to ensure good coverage. For systems with a lower number of electrodes, their spatial distributions are not suitable for all experiments. Interpolating a moderate number of electrodes to a fine mesh might help in some situations. However, they lack flexibility as generic EEG systems. For the Emotiv EPOC especially the users should be aware of the shortcomings and use the system accordingly.

More subjects in the current study would be beneficial, as well as increasing the number of sessions with new montages. However, even with four subjects, four EEG systems, and three sessions with new montages, there are already forty-eight recording sessions. Therefore, a multicenter study involving a larger number of subjects and systems using the present benchmark is a desirable next step.

We compared different EEG systems in this study, but the ground truth of EEG signal remains unknown. For example, the two research-grade systems in the study produce similar results, but it is hard to argue whether these are closer to the unknown ground truth than others or not. There are studies which use a phantom head with controlled dipolar sources of electrical activity embedded in the phantom (Oliveira, Schlink, Hairston, König, & Ferris, 2016a). On the one side, such approach allows us to obtain results closer to the ground truth since the exact electrical stimulation of the phantom can be conducted on different EEG systems. On the other side of the phantom head approach, the real generators of EEG signal in the brain, as well as tissue conductivity and other less predictable properties of natural environment, are missed. While the ground truth of scalp EEG signals is unknown, the expectations from new mobile EEG systems are to produce comparable results to those of research-grade EEG systems, which they do to varying degree.

The mean amplitude values in Fig. 13 are usually larger than they look on the ERPs in Fig. 4, 6-10, panels C-E. To yield a measurement of a dependent variable, we used search intervals for each ERP of the eight sequential parts of a recording session (Fig. 5), but the ERPs in Fig. 4, 6-10 were yielded by a standard averaging procedure of ERPs from the eight sequential parts of a recording session. Therefore, these search intervals made the mean amplitude values in Fig. 13 larger than they look on the ERPs in Fig. 4, 6-10, panels C-E.

We encountered a time displacement of the grand total ERPs between the current study and respective examples from the literature. We think that the reason for that is that different EEG systems were used in the literature examples and in the current study. Indeed, we also encountered a time displacement of 11 ms in the current study between g.Nautilus and the research-grade EEG systems (asalab^{TM} and

actiCAP). Such time displacements may happen due to a variety of technical reasons. The same problem of time alignment between EEG systems was also encountered in other studies (Gargiulo et al., 2010; Ries, Touryan, Vettel, McDowell, & Hairston, 2014). However, these issues did not influence measurements of dependent variables in the study.

The current study proposes an approach to evaluate new mobile EEG systems by means of ERP responses. The developed benchmark of six paradigms and yielded results can be used for evaluation of new EEG systems. In contrast to other studies (M. De Vos et al., 2014; Gargiulo et al., 2010; Guger et al., 2012; Liao et al., 2012; Oliveira et al., 2016b; Yeung et al., 2015), which compare EEG signals from different systems, we match the difference between EEG systems to other factors of variance, like subjects and sessions, in order to estimate the importance of the difference between EEG systems. We also demonstrated the mesh-head-model interpolation approach, which addresses the issue of not overlapping EEG montages of different EEG systems. Results showed that research-grade EEG systems are indeed mature since they have no significantly different means through all paradigms between the two systems. The g.Nautilus with g.SAHARA electrodes and Emotiv EPOC EEG systems used in the current study are in many paradigms as good as current research-grade systems.

3.6 Acknowledgments

The authors would like to thank Marcel Schumacher, Vivien Radke, Ashima Keshava, Maria Sokotushchenko for helping with data acquisition and Allison Moreno-Drexler for proofreading.

Chapter 4. The Price of Saccades

This study has been presented as a poster at the 40th European Conference on Visual Perception (in proceedings).

*Melnik, A., *Schüler, F., Rothkopf, C., & König, P. (2017). The Price of Saccades. European Conference on Visual Perception 2017, Berlin, Germany.
http://journals.sagepub.com/page/pec/collections/ecvp-abstracts/index/ecvp-2017
(*first two authors contributed equally)

AUTHOR CONTRIBUTIONS
Study conception and design: F.S., C.R., & P.K.; acquisition of data: F.S.; analysis and interpretation of data: A.M., F.S.; drafting of manuscript: A.M., F.S.; critical revision: A.M., P.K.

The Price Of Saccades

*,**Andrew Melnik[1], *Felix Schüler[1], Constantin Rothkopf[2] & Peter König[1,3]
[1]University of Osnabrück, Germany; [2]Technische Universität Darmstadt, Germany;
[3]University Medical Center Hamburg-Eppendorf, Germany

* First two authors contributed equally
** andmelnik@uni-osnabrueck.de

Abstract: theories of embodied cognition postulate the world as an external memory. It implies that instead of storing visual information in working memory the information may be equally retrieved by appropriate eye movements. Effectively, humans trade the effort of memorization with the effort of active sampling of visual information. This reasoning predicts that the balance of memorization and active sampling might be shifted by making one or the other more effortful. **Goal:** we intended to quantify the balance between sampling of visual information and use of working memory. **Data:** we analyzed eye-tracking data of 20 participants in a block-copying task [1] where participants had to build a copy of a model displayed on a screen (Fig.1). **Approach:** we used unconstrained (A) and constrained (B) conditions to compare a number of times participants referred to a given model of 8 blocks. In the constrained condition (B), participants waited 700 ms until a model appeared on a screen each time they looked at the model (Fig.2,3). **Results:** in the unconstrained condition (A) participants looked at a model on average 7.9 times, in the constrained condition (B), 5.8 times (Fig.5). **Conclusion:** we demonstrated the shift towards memorization when we introduced a cost for saccades (Fig.4,5), however, our results also showed a suboptimal shift of an adaptive system, as memorization leads to longer performance time (Fig.6).

Experimental setup

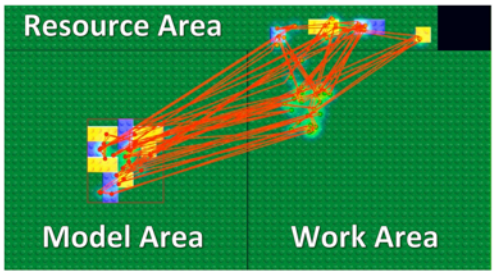


Figure 1. Experimental setup. We divided a computer screen into three main areas: **Model Area**, **Work Area**, and **Resource Area**. Participants had to drag and drop blocks using a computer mouse from Resource Area to Work Area in order to copy a model of eight blocks demonstrated in Model Area. Red lines demonstrate saccades in a single trial.

We asked participants to build a copy of a given model

700 ms delay in Condition B

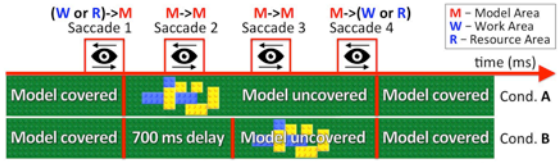


Figure 2. Timeline of conditions A and B. In Unconstrained Condition (A), the model appeared immediately after eye-fixation into Model Area. In Constrained Condition (B), we introduced a 700 ms delay before uncovering of the model.

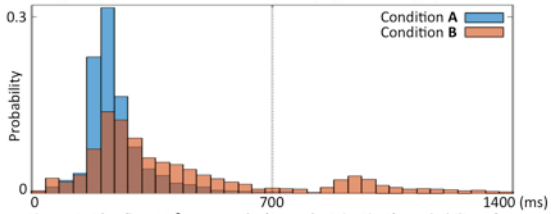


Figure 3. The first M→M saccade (Saccade 2 in Fig.2). Probability of a time interval between Saccade 1 (end) and Saccade 2 (end) (see Fig.2). The sum of the bar heights is 1 for blue and 1 for red.

We excluded M→M saccade occurred in the 700 ms delay period in Cond. B

Markov Chain model of Saccades

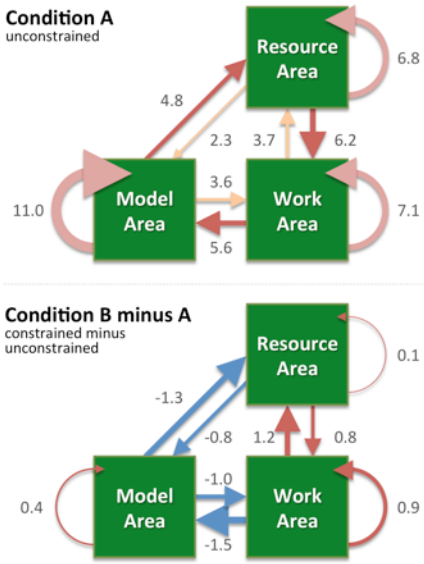


Figure 4. Baseline Condition A and difference between Conditions A and B. Values demonstrate mean number of saccades per trial. Saccades during 700 ms delays in Condition B were not included. Blue arrows represent negative values.

Participants referred to a model less in Cond. B than in Cond. A

Total stats

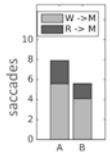


Figure 5. Average number of saccades into Model Area from Work Area (W->M) and from Resource Area (R->M) for conditions A and B.

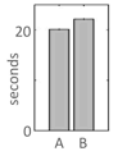


Figure 6. Average duration of a trial in Conditions A (20.1 sec) and B (22.2 sec). The bar for Condition B does not include 700 ms delay periods.

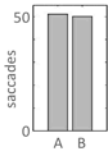


Figure 7. Average number of saccades in a trial in Conditions A (51.2) and B (50.1). The bar for Condition B does not include saccades during 700 ms delays (Fig.2,3).

Condition B (constrained) requires less saccades but more time to accomplish a trial

References: [1] Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. Journal of Cognitive Neuroscience, 7(1), 66-80.

4.1 Abstract

Theories of embodied cognition postulate the world as an external memory. It implies that instead of storing visual information in working memory the information may be equally retrieved by appropriate eye movements. Effectively, humans trade the effort of memorization with the effort of active sampling of visual information. This reasoning predicts that the balance of memorization and active sampling might be shifted by making one or the other more effortful. Here we test this prediction in a block copying task (Ballard et al., 1995).

We analyzed eye-tracking data of 20 participants in a block-copying task where participants had to build a copy of a model displayed on a computer screen. In the experiment, we divided a computer screen into three main areas: Resource Space, Work Space, and Model Space. Participants had to drag and drop blocks from Resource Space to Work Space in order to copy a model of eight blocks in Model Space. Each block was of one of two colors and one of three types (square, vertical, or horizontal rectangle). We used an unconstrained (A) and a constrained (B) conditions to compare a number of times participants referred to a given model of 8 blocks. In the unconstrained condition (A), the model appeared immediately after eye-fixation into Model Space. In the constrained condition (B), we introduced a 0.7 s delay for uncovering of the model.

In the unconstrained condition A, participants made on average 7.9 saccades into Model Area to successfully copy a model of 8 blocks. In the constrained condition B participants made on average 5.6 saccades into Model Area to successfully copy a model of 8 blocks. The mean number of saccades of all types per trial in the condition A was 51.2 and in B 50.1. However, the mean duration of a trial was 2.1 s (9%) longer in the constrained condition B. In the study, we demonstrated the shift towards memorization when we introduced a price of saccades, however, our results also showed a suboptimal shift of an adaptive system, as memorization leads to longer performance time.

4.2 Materials and methods

Participants.
Twenty paid healthy volunteers (seven males), aged 19–27 years (mean = 22.7), participated in the study. All participants had normal (self-reported) or corrected-to-normal vision. We obtained written informed consent from all subjects before the experiment, and the protocol had been approved by the review board of the University of Osnabrück. Experimental procedures conformed to the Declaration of Helsinki and national guidelines.

Stimuli and task.
In the experiment, we divided a computer screen into three main areas: Resource Space, Work Space, and Model Space. We asked participants to copy a model of eight blocks demonstrated in Model Area by dragging and dropping blocks from Resource Area into Work Area. Participants were able to move a block by drag and drop and rotate by a right click on a computer mouse. By clicking on a block of a specific size and color in Resource Area, a new instance of that block was created and attached to a computer mouse cursor. Participants could dispose of a block picked up by mistake in the trash-area (black rectangle in the Resource Area). Each block was of one of two colors (blue or yellow) and one of three types (square, vertical, or horizontal rectangle). Trials ended automatically when participants correctly copied the model. If the model was not copied correctly during a specified period, the trial ended automatically.

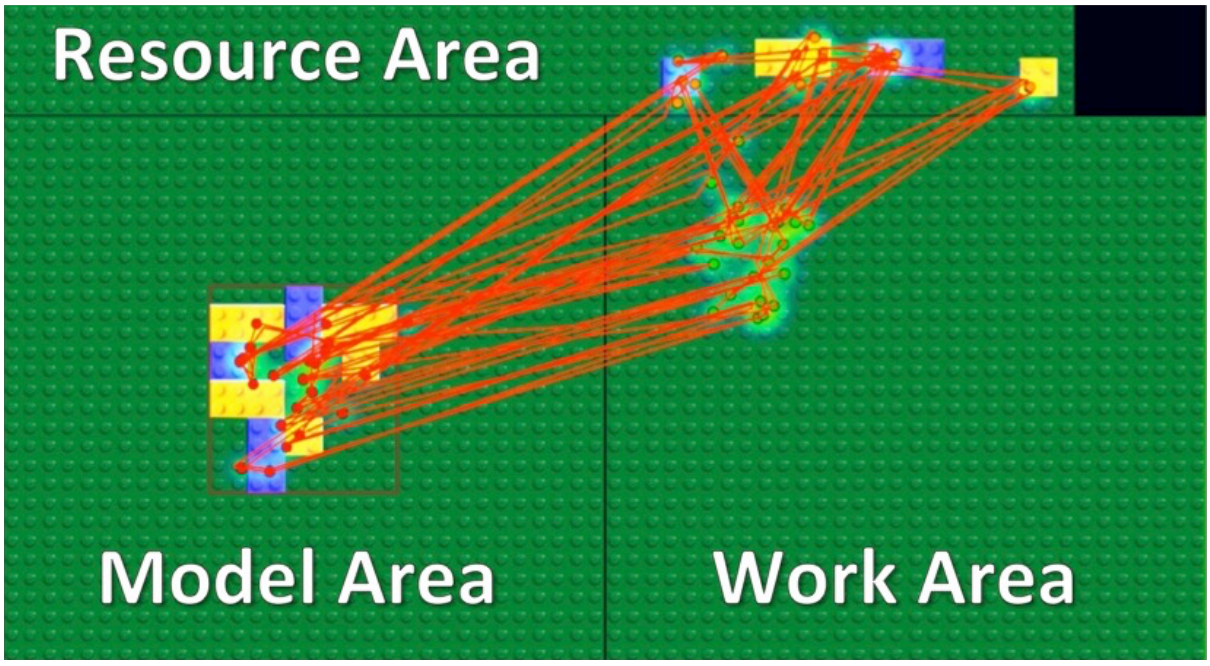


Figure 1. Experimental setup. We divided a computer screen into three main areas: Model Area, Work Area, and Resource Area. Participants had to drag and drop blocks using a computer mouse from Resource Area to Work Area in order to copy a model of eight blocks demonstrated in Model Area. Red lines demonstrate saccades in a single trial.

A model always appeared in the center of Model Area, and its location was outlined by the red rectangle shown in Figure 1. A model was not visible and only appeared while fixations inside the area indicated by the red rectangle (Fig. 1). Therefore, participants could not obtain visual information from peripheral vision while working in Work Area or Resource Area. We hand-designed all models to avoid similarities to familiar object shapes. Our stimuli and task were motivated by a related study with a block-copying task (Ballard et al., 1995) and implemented in a gaze-contingent paradigm with eye and mouse tracking.

In our experimental paradigm, we introduced a price of saccades as a 0.7 s delay for uncovering of the model in a constrained condition B (Fig. 2 Cond. B). We will refer to the condition B as the constrained condition since the 0.7 s waiting time effectively introduced a penalty for sampling the model. Participants had to maintain eye fixations in the area of the red dashed rectangle in Model Space (Fig. 1) for 0.7 s until the model appeared. In an unconstrained condition A (Fig. 2 Cond. A), the model appeared immediately after eye-fixation into red dashed rectangle in Model Space. Each trial started after subjects fixated the centered fixation cross and typical subsequent fixations were made to the model. In both conditions, the model disappeared as soon as participants made a saccade outside of this area.

The design of the experiment and instructions required from participants to be as fast as possible. Furthermore, we encouraged participants by additional monetary reward for fast performance.

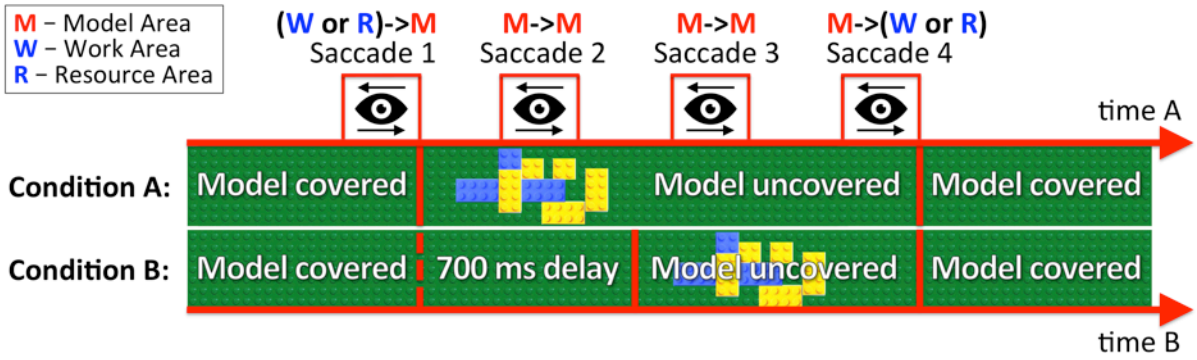


Figure 2. Timeline of conditions A and B. In Unconstrained Condition (A), the model appeared immediately after eye-fixation into Model Area. In Constrained Condition (B), we introduced a 0.7 s delay before uncovering of the model.

The experiment consisted of 80 trials, of which 40 Unconstrained condition trials (A) and 40 Constrained condition trials (B). Participants were randomly assigned to ABBA or BAAB (with 20 trials per block) order of conditions. The 80 trial templates were the same for all participants, but their order was randomized, and they could occur in Unconstrained or Constrained conditions.

Excluded trials.

We recorded 1600 trials in the study (20 participants x 80 trials). However, we excluded 13% of trials in the study (206 out of 1600) based on two factors. The first factor was a duration of a trial. Trials ended automatically if participants did not copy a model in 27 seconds in a condition A, and in 37 seconds in a condition B. Therefore, due to the first factor we excluded 198 trials (87 trials in the condition A and 114 trials in the condition B). The second factor was a number of saccades into Model Area from other areas. We excluded a trial if a subject did less than 3 or more than 24 saccades into Model Area per trial. We excluded 5 trials (2 trials in the condition A and 3 trials in the condition B) based on the second factor. Therefore, we excluded from data analysis 89 trials (11%) in the unconstrained condition A and 117 trials (15%) in the constrained condition B.

Equipment.

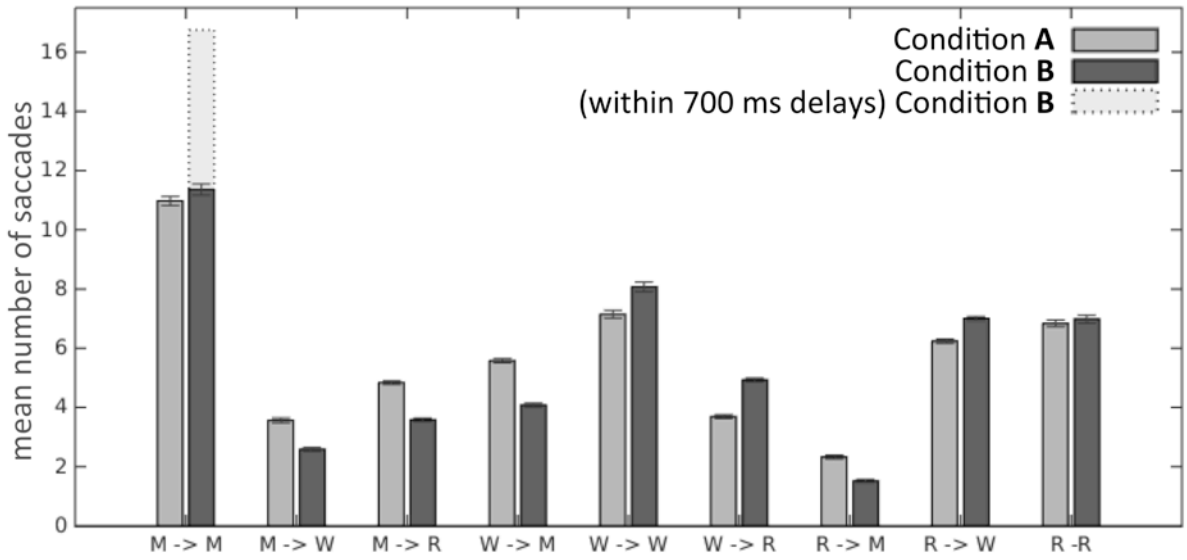
The experiment was set up as a gaze-contingent paradigm with eye and mouse tracking. A remote infra-red eye tracking device (Eyelink 1000, SR Research Ltd., Mississauga, Canada) with 500 Hz sampling-rate was used to track eye-movements, and its standard settings were used to classify fixations, saccades and blinks. A calibration procedure assured that the average fixation-deviation was below 0.5 visual angle. We showed stimuli on a 60Hz, 23" Dell monitor with a resolution of 1920 x 1080 pixels in full-screen mode and the room was darkened to facilitate eye tracking. Subjects had to use a mouse to navigate, rotate and position the blocks and a keyboard was used to enter a nickname or to navigate the instructions.

4.3 Results

Our main hypothesis was that participants would avoid penalized saccades to sample a model of blocks. To answer this question, saccades were classified according to their start- and end-locations (Model Area, Work Area, or Resource Area). We calculated histograms for all of the nine possible combinations of start and end locations (Fig. 3).

Participants were doing MM saccades (from Model Area into Model Area) in condition B during the 0.7 s delay period (Fig. 2). We calculated a mean number of such saccades (Fig. 3) and distribution of time intervals (Fig. 4) between the end of a saccade into Model Area (the activation event of the 0.7 m delay period, Fig. 2) and the end of the first MM saccade in the model area (a new sample of visual information, Fig. 2). The results show that even in the absence of relevant information participants follow a similar visual exploration dynamic. However, we also observed a shift towards a second peak at the 0.92-0.96 s bin (Fig. 4).

We excluded from further data analysis MM saccades in condition B during the 0.7 s delay period (Fig. 2, 4). We assumed that such saccades did not provide relevant to the task information because a model appeared only after the 0.7 s delay period. MM bars in Figure 3 show the influence of the excluded saccades onto the model.



	MM	MW	MR	WM	WW	WR	RM	RW	RR
Condition **A**	11.0	3.6	4.8	5.6	7.1	3.7	2.3	6.2	6.8
Condition **B**	11.4	2.6	3.6	4.1	8.1	4.9	1.5	7.0	7.0
Condition **B** with 700 ms delay saccades	16.8	-	-	-	-	-	-	-	-

Figure 3. Mean number of saccades per trial in conditions A and B. Y-axis represents a mean number of saccades between and within Model Area (M), Work Area (W), and Resource Area (R) in conditions A (Unconstrained) and B (Constrained) per trial (all subjects and all trials). Error bars represent standard error of the mean.

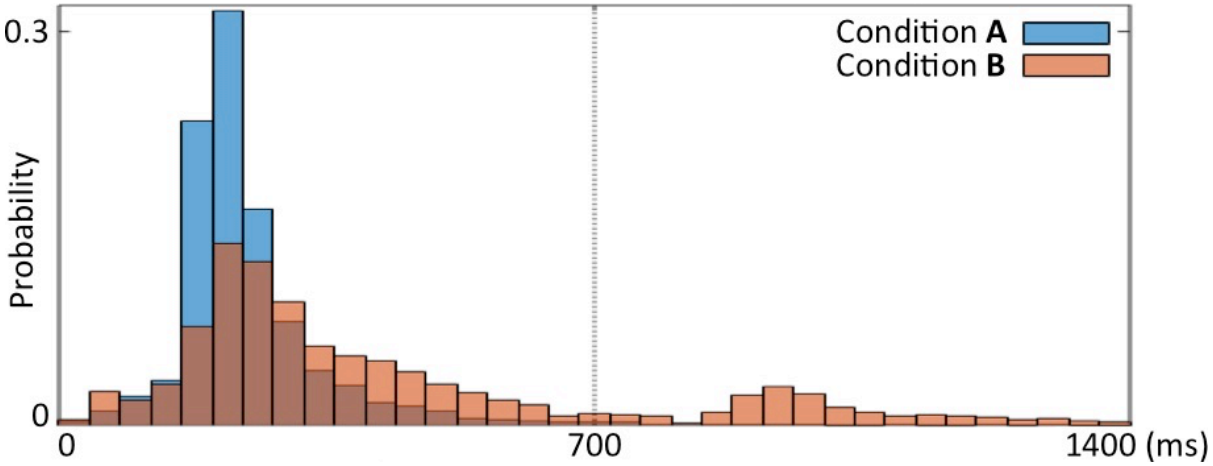


Figure 4. The first MM saccade (Saccade 2 in Fig.2). Probability of a time interval between Saccade 1 (end) and Saccade 2 (end) (see Fig.2). The sum of the bar heights is 1 for blue and 1 for red.

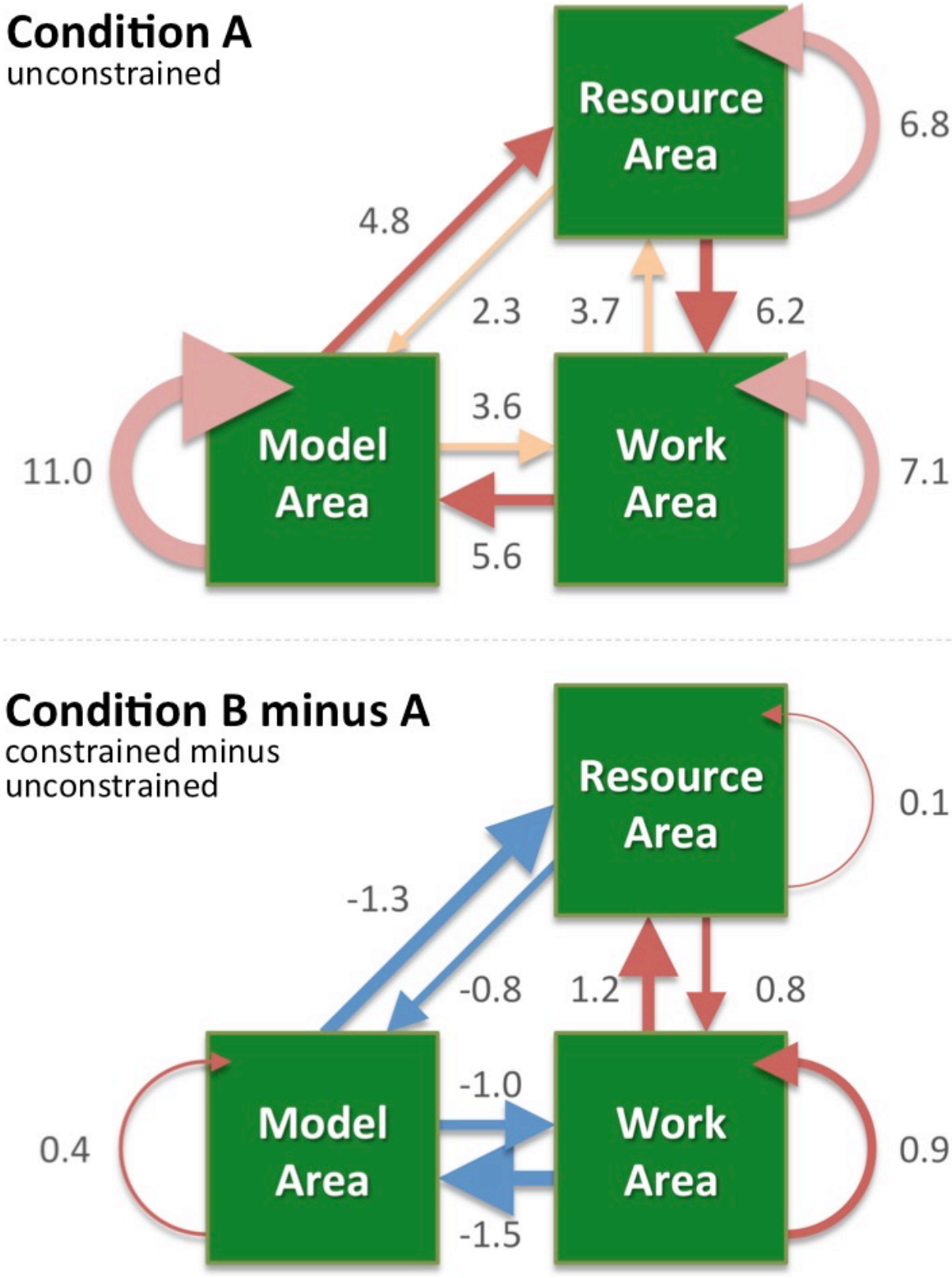


Figure 5. Baseline Condition A and difference between Conditions A and B. Values demonstrate a mean number of saccades per trial. Saccades during 700 ms delays in Condition B were not included. Blue arrows represent negative values.

The Markov Chain model for the Unconstrained condition A (Fig. 5) demonstrates a dominance of the MR - RW - WM cycle. The difference between conditions A and B reviled a compensation against the MR - RW - WM cycle.

We built a Markov Chain model for the unconstrained condition A (Fig. 5) and used it as a baseline. We also built a Markov Chain model for the constrained condition B and subtracted the Markov Chain model A from it. The difference between Markov Chain models for conditions A and B (Fig. 5) reviled the effect of the price of sampling of visual information. Blue arrows in Figure M (condition B minus A) show that participants were trying to avoid sampling of Model Area, and instead, we observe more saccades between Work Area and Resource Area.

The cumulative effect of the price of visual sampling of a model is shown in Figure TA. In the unconstrained condition A, participants made on average 7.9 saccades into Model Area to successfully copy a model of 8 blocks. In the constrained condition B participants made on average 5.6 saccades into Model Area to successfully copy a model of 8 blocks (Fig. 6A). Participants made on average 1.1 saccades (2%) of all types less in the constrained condition B (Fig. 6B). The mean number of saccades of all types per trial in the condition A was 51.2 and in B 50.1 (Fig. 6B). However, the mean duration of a trial was 2.1 s (9%) longer in the constrained condition B (Fig. 6C). Mean duration of a trial in the condition A was 20.1 s and in B 22.2 s (Fig. 6C). The mean duration of a trial in the condition B does not include 0.7 s delay periods.

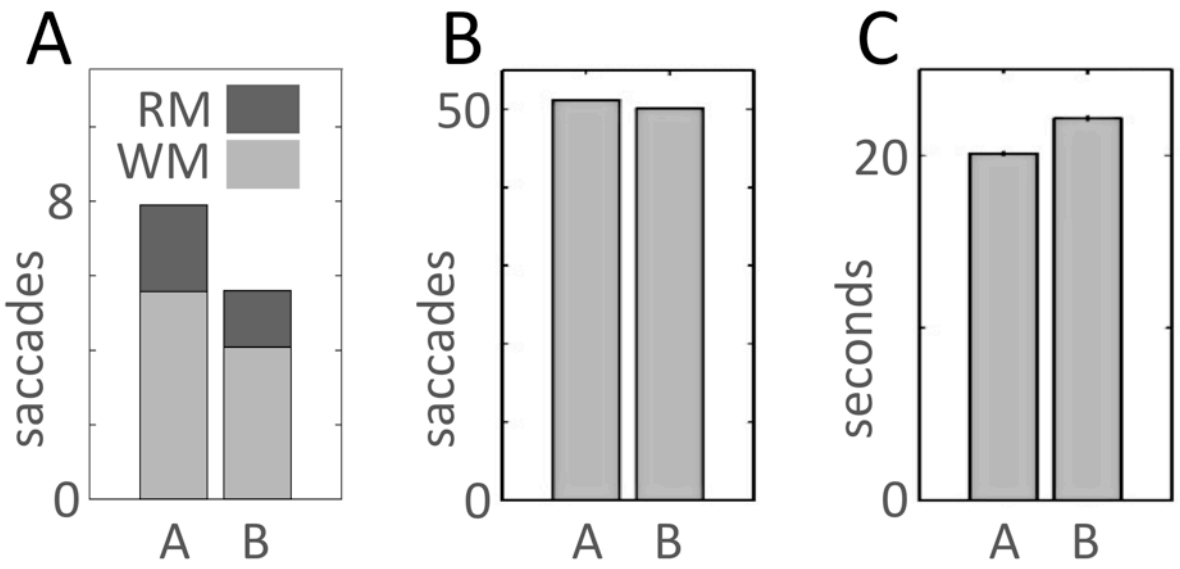


Figure 6. A. Mean number of saccades from Work Area into Model Area (WM) and from Resource Area into Model Area (RM) for conditions A and B. **B.** Mean number of saccades per trial in Conditions A and B. The bar for Condition B does not include MM saccades during 0.7 s delays (Fig. 2,3). **C.** Mean duration of a trial in Conditions A and B. The bar for Condition B does not include 0.7 s delay periods.

4.4 Discussion

In purposeful tasks where the gaze is crucially involved, humans trade the effort of memorization with the effort of active sampling of visual information. To better understand the balance between memorization and active sampling, we devised an experiment, in which participants copied a model of eight blocks shown in Model Area by dragging and dropping blocks from Resource Area into Work Area. We investigated if a parametric change of the price of saccades influences eye-movement behavior. We hypothesized that the balance of memorization and sampling of visual information might be shifted by making one or the other more effortful. In Unconstrained condition A to successfully copy a model of 8 blocks participants made on average 7.9 saccades (Fig. 6A) into Model Area. In Constrained condition B to successfully copy a model of 8 blocks participants made on average 5.6 saccades (Fig. 6A) into Model Area. The results supported our hypothesis. Participants' behavior significantly changed towards increased use of working memory when we introduced the price of saccades (0.7 s delay).

Everyday life mostly consists of purposeful tasks where the gaze is crucially involved. "Human gaze is directed toward regions of the visual scene that are determined primarily by the task requirements" (Rothkopf et al. 2007). "Visual information required for a task is acquired just prior to its use" (Ballard et al. 1995). "95% of fixations are dedicated to completing the immediate sub-task." (Land et al. 1999). "Humans choose not to operate at the maximum capacity of short-term memory but instead seek to minimize its use. Reducing the instantaneous memory required to perform the task can be done by serializing the task with eye movements." (Ballard et al. 1995). "Eye movements serialize complex operations, optimizing performance by deploying foveal acuity as required by the task, as in the case of eye movements to objects 500 ms before reaching." (Pelz & Canosa 2001).

It is unlikely that a complete representation of the visual scene is maintained in memory (Marr, 1982; Nakayama, 1991). Studies suggest that visual working memory has a limited capacity of 3-7 items (Luck & Vogel, 1997; Cowan, 2001; Todd & Marois, 2004) and temporal decay limitations. One reason for such limitation could be a coupling of perceptual and motor representations in the human brain while a participant performing a sensorimotor task (Melnik et al., 2017). In our experiment, the number of blocks in a model was about the level of estimated capacity of the human working memory. However, one block was characterized by several features (color, size, orientation, and location), making it hard for participants to remember more than a few blocks at once. Nevertheless, humans tend to not reach these limits and choose to serialize a task by adding more eye fixations (Ballard et al., 1995).

These fixations allow participants to postpone the acquiring of task-relevant information just before its use. Our results suggest that this is a general phenomenon that can be observed in condition A trials more frequently than in condition B trials as presented in the following sections. Participants evoked a saccade to obtain features of each single block in the unconstrained condition A (7.9 saccades per trial). The change of the strategy in the constrained condition B showed that subjects were able to dynamically optimize the trade-off between sampling of visual information and using of working memory.

The price of a 0.7 s delay period was the central parameter in the design of our study. The price was designed to be high enough that participants tried to avoid it, but without high frustration to pay it. We explored different values for the delay-duration and finally set to the period of 0.7 s. In the study, we used only the 0.7 s delay period, which allowed only for a linear model of the balance between memorization and active sampling of visual information. However, to infer a more precise relation we should vary parameters of the model with smaller steps. Therefore, as a further development of the study, we propose to use a set of delays (e.g., 0.1 s, 0.3 s, 0.5 s, 0.7 s, 0.9 s, etc.). Another parameter to vary is a complexity of displayed models (e.g., 2, 4, 8, 16, 32 blocks). Moreover, we could introduce a parameter of distraction (e.g., auditory or visual distractors) or a parallel subtask. That would allow for building a more precise and comprehensive model of the balance between memorization and active sampling of visual information. Then, it could be possible to see a point of a strategy shift and dynamic adoption of eye-movement behavior.

We set the duration of trials to 27 seconds in the unconstrained condition A and to 37 seconds in the constrained condition B. The duration of trials in the condition B was longer than in condition A to account for the delay periods (0.7 sec). Our goal was to allow participants for completing 90% of trials within a given period. We recorded two pilot subjects without any limitations on the trial duration. Results from the pilot recordings suggested setting the maximum duration of trials in the condition A to 27 seconds. To determine a duration of trials in the condition B we computed the mean number of evoked delays (saccades into Model Area) in the 90% of shortest trials, multiplied it by the delay period (0.7 s) and added this to the duration of trials in the condition A. The average number of penalties in the pilot recordings in the 90% fastest trials was 13.8 which resulted in a maximum duration of 37 seconds for condition B trials (27 s + 13.8 x 0.7 s = 36.66 s., or 37 s). Analysis of the results of our experiment showed that this approach led to the desired outcome. We excluded 13% of trials in the study 206 out of 1600.

In the study, the task was to copy a model of block as fast as possible. One way participants improved performance was copying a model close to Resource Area and Model Area, thus avoiding long saccades.

Many studies have shown that bottom-up saliency and low-level image statistics are good predictors for fixation selection during free viewing (Itti & Koch, 2001; Zhang et al., 2008; Kaspar & König, 2011). However, the presence of a task leads to a strong influence of top-down and task-dependent effects on eye-movement-behavior where the task mainly determines where participants look (Rothkopf et al., 2007; Betz et al., 2010). In the present study, we have demonstrated that a price for certain types of saccades also significantly influence eye-movement-behavior of participants.

In the constrained condition B, participants waited 0.7 s for a model. While waiting for the model, there was no task-relevant information, and no other executable cognitive task was available. Therefore, we expected that participants would stare at the same location on the board due to the lack of available task-relevant information. However, the analyzes of eye-movements reviled a substantial number of MM saccades (Fig. 2, 3, 4) while subjects waited for the model to appear. The purpose of the observed fixations during the waiting-period was not clear. Our eyes are constantly moving, even though we might not always be aware of it. "Fixational eye movements not only help to keep objects of interest in the center of our vision, but also prevent sensory adaptation in our visual path by refreshing our retinal images. Without fixational eye movements, we would be blind during visual fixation, and the world would become visible only when we moved our eyes voluntarily, when we moved our heads or when the world moved in front of us" (Martinez-Conde et al., 2004). We suppose that microsaccades can explain observed eye-movements during delay periods.

We designed this study to understand the trade-off between the gaze and working-memory use in a goal-directed task better. The cost for a new sample of visual information that participants had to "pay" was a short time-delay until a model of blocks appeared on a screen. Participants' behavior significantly changed towards increased use of working memory when we introduced the price of saccades.

Chapter 5. A Cognitive Architecture For Spatial Reasoning

Unpublished work. Melnik, A. (2017). A Cognitive Architecture For Spatial Reasoning.

AUTHOR CONTRIBUTIONS
Study conception and design: A.M.; acquisition of data: A.M.; analysis and interpretation of data: A.M.; drafting of manuscript: A.M.; critical revision: A.M.

5.1 Introduction

Goal-directed behavior requires translation of sensory information into effective motor actions. A simple approach is to translate sensory information into motor outputs directly. Direct translation works well until a certain level of complexity of a sensory input, a motor output, or a task. Development of cognitive architectures pursues to build a model of efficient translation of high-dimensional sensory inputs into goal-directed motor outputs. The problem of deriving "efficient representations of the environment from high-dimensional sensory inputs" (Mnih et al., 2015) and the problem of representation and execution of complex motor actions (Bohg, Morales, Asfour, & Kragic, 2014) are important scientific questions in fields which deal with sensorimotor processing. People can intuitively perform tasks which require intuitive physics knowledge and spatial reasoning. However, such tasks are problematic for robotics, machine learning and artificial intelligence approaches (Weitnauer & Ritter, 2012; Sermanet, Xu, & Levine, 2017).

In the human brain, there are two parallel visual streams: "what" and "where." This fact may be a hint for developing cognitive architectures for spatial reasoning. The idea of deictic codes fits well to the dichotomy of the neocortex. "The word deictic means 'pointing' or 'showing.' Deictic primitives dynamically refer to points in the world with respect to their crucial describing features (e.g., color or shape). The dynamic nature of the referent also captures the agent's momentary intentions." (Dana H. Ballard et al., 1997). Indeed, the approach to process at once all relationships of parts of an image or a scene encounters many complexities. Another approach is to break down a task into a sequence of localization and identification tasks.

Attentional mechanisms allow for the extraction of features that are essential for the task from the current fixation and to select a next fixation point. Meanwhile, additional information can be stored in working memory. Working memory of the human brain can simultaneously hold 3-7 units which are relevant to a task (Cowan, 2001, 2008; Miller, 1956). This capacity limitation of the working memory may not be a shortage but a consequence of the cognitive architecture of the human brain that allows efficient searching, grouping, and processing of information.

How is visual information about objects and environments stored in cognitive architectures? In many tasks, only a few features of an object are important for performing a task. The rest of the object can have a simplified representation. For example, dealing with shapes in a spatial reasoning task does not require representation of a shape as a fully descriptive token, e.g., the Hough transform (D. H. Ballard, 1981). Some important points of the shape can be selected, and their influence on the behavior can be learned. All spatial interactions could be

represented as an interplay of a limited number of important points or features (falling of an object, collision, sliding). Representation of objects, shapes, and environments as constellations of a few important points can dramatically reduce processing complexity of visual information and matches with the fact of capacity limitation of working memory in the human brain.

Objects and environments can have affordances. The concept of affordance (Greeno, 1994; Turvey, 1992) refers to relationships between actions, objects, and effects (Montesano et al., 2008). “When affordances are perceptible, they offer a direct link between perception and action. Complex actions can be understood in terms of groups of affordances that are sequential in time or nested in space” (Gaver, 1991). Perception of a scene is a unified experience in humans and correlates with a location mapping of task-relevant information, which can be immediately accessed by a new fixation. A high level of predictability of future information from the next eye fixation correlates with the feeling of the full perception of an environment. Thus, when any next eye fixation brings no new information, then the environment is fully perceived.

5.2 Methods

Translation of sensory information into objects in a physics engine with a set of features provides additional information on interactions between objects and an environment. In this study, I used the Pymunk physics engine (http://www.pymunk.org) as a basis to work with an environment and objects. This engine allows for the access and use of object related information (e.g., velocity, shape, position, etc.) and the storage of affordances directly in objects. Information provided by the physics engine is useful for building dependencies for predictions, but it is not enough for spatial reasoning.

The ideas of deictic codes and affordances imply breaking down spatial and temporal relationships and actions into a sequence of task related episodes. We assume that each episode refers to a homogeneous type of calculations. For example, a ballistic trajectory in a gravity field may be one episode. Other episodes could be a collision, a sliding over a curve object, etc. Another example is a task to get a billiard ball A into a hole B using a ball C. The task can be broken down into four episodes. (1) hit a ball C. (2) rolling of the ball C in a straight trajectory. (3) a collision of the balls C and A. (4) The ball A gets into the hole B. There is only one type of interaction in each episode. The homogeneity of interactions within an episode makes it reusable in new environments and tasks.

One of the problems of a goal-directed behavior is to find actions which transform an object or an environment from the current state to a goal state. This problem relates to planning and planning means to make reliable predictions based on experience. Episodes allow predictions and a corresponding mapping of actions. Episodes are represented as nodes in a graphical model (Bishop, 2006). Strategic planning correlates with a selection of a path in a graph. Tuning of relationships in an episode (node) correlates with an improvement in tactic.

The physical interactions are complex and barely possible to predict as precise as a real physical experiment. However, it is possible to learn approximate dependencies which are necessary for building a path in a graphical model. Approximate dependencies do not reveal precise action in an episode but, since the number of connected episodes is limited, the approximate correlates successfully highlight a path in a graphical model.

Another source of predictions in episodes is a simulation with the physics engine. It is efficient to use online internal simulations to build a precise temporary mapping for a current situation. An array of simulations for different values of an independent variable pinpoints the best value of a parameter in the episode.

An infinite number of possible objects, environments and their interactions is possible to break down into a finite number of types of episodes. The episodes are interconnected between each other as nodes in a graphical model of goal-directed behavior. An important question is, how to break down processes into episodes? The desirable result is to build a cognitive architecture that can learn and break down processes into episodes in an unsupervised manner.

The simplest sensorimotor primitive is a pointing visual fixation (eye fixation in humans). Each new point-centered fixation refers to a task-relevant hypothesis check. Hypothesis failure can contribute to separation into episodes. Object-related triggers (e.g., collision events) and unpredictable changes of object-related parameters (velocity, acceleration, angle, etc.) provided by the physics engine can also help to break down processes into episodes. Another approach to an automated breaking down into episodes would be to trace a sensorimotor flow in a time window of 1 second in the past and do a 1 second prediction in the future. If predictions based on the past do not correctly predict the future anymore, then a new episode starts from the time point of deviation. It is possible to represent complex shapes as constellations of a few important points. It is also possible to learn episode-related dynamics of such constellation points and categorize them with machine learning approaches. In the current study, I used hard-coded episodes which can learn sensorimotor mappings (current state -> interaction -> goal state).

The implementation of the cognitive architecture was uploaded to github: https://github.com/ndrwmlnk/CA.

5.3 Computation steps and pseudocode

After the acquisition of essential episodes, gaining enough statistics, and learning probabilities of translations between episodes, it is possible to combine the episodes in a sequence of a graphical model which lead to a goal.

Computation steps of the cognitive architecture in Figure 1:

- initialize four objects of a scene (ball (Obj1), wall (Obj2), ground (Obj3), and a spare ball (Obj4)) and allocentric spatial dependencies between them (Figure 1A)
- initialize a task: touch event of Obj1 and Obj2 (the ball and the wall should touch each other)(Figure 1B)
- Obj2 is not moveable, continue with Obj1
- Obj1 is moveable
- determine the context: the object lying on a surface in a gravity space (Figure 1C)
- the context provides two reallocation episodes: a sliding over a line episode (Traj1) and a ballistic episode (Traj2) (Figure 1D)
- check trajectories in both episodes. No obstacle detected (Figure 1E)
- sliding over a line episode requires a change of the velocity vector from V1 (zero) to V2 (directed towards the wall) (Figure 1F)
- velocity change is associated with the collision and gravity episodes (Figure 1F)
- the collision episode gets the current velocity vector (V1), desirable velocity vector (V2), and the context (the object lying on a surface in a gravity space). The collision episode pinpoints the part of Obj1 appropriate for the collision (Figure 1F)
- select the ballistic episode (Figure 1G)
- places Obj4 at a probable starting point (Figure 1H)
- start a simulation of the scene (Figure 1H)
- compare predicted outcomes in episodes with the physics-engine simulation (Figure 1I)
- do adjustments and update mapping in episodes if something deviates from the model.
- store the sequence of episodes as a graphical model with contextual metadata of the scene

Scenes with known solutions have a stored sequence of episodes. Therefore, if a scene is recognized, then the desirable action can be executed, and the real simulation starts.

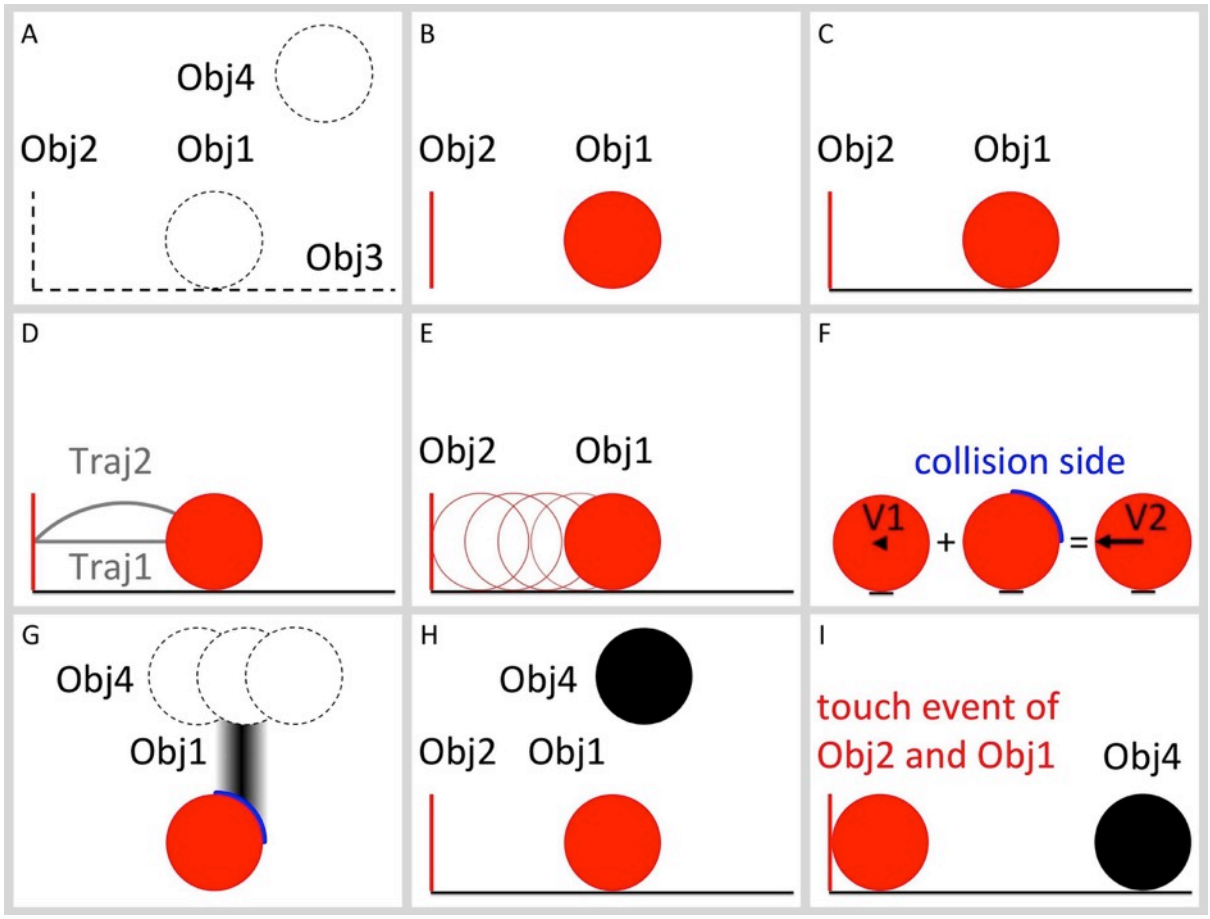


Figure 1. Episodes of spatial reasoning. Task: make the ball hit the left wall. Red lines and balls depict primary objects in the task. Black lines and balls depict auxiliary objects in the task. Dashed units in panel A depict detected but not yet initialized objects. Dashed units in panel G depict a searching for the best position of Obj4. The black gradient in panel G depicts a probability of getting a predicted outcome of the collision.

5.4 Discussion

In the current study, I proposed to break down spatial reasoning into a sequence of episodes where each episode contains a homogeneous type of interactions, e.g., collision, ballistic trajectory, sliding over a line. Episodes substantially simplify prediction dependencies and mapping of actions between states (current and required) in each episode. Another advantage of episodes is suitability for a semi-supervised learning. Homogeneity of dependencies in episodes allows interpolating between learned examples. Thus, several examples can be enough for covering all possible input-output-action combinations and check the predictions by the trial and error method. The episodes are interconnected between each other as nodes in a graphical model of goal-directed behavior. Therefore, spatial reasoning works as a search mechanism in graphical models. A solution for a spatial task is a path through a sequence of nodes between the current state of an object or an environment and a goal-state of the object or the environment. The phenomenon of perceptual understanding of a scene correlates with recognition of a graphical model related to an environment.

Episodes in the cognitive architecture are related to eye-fixations in humans. Eye fixations play an important role in visual processing (Kowler, 2011). They are correlates of planning, reasoning and scene understanding. The episodes are also related to a deictic codes approach (Dana H. Ballard et al., 1997). The approach supports an idea of performing one task at a time: to do inference at a fixation point or to move to a next fixation point. A visual patch surrounding the fixation point is input data for an episode. Therefore, a graphical model provides not only sequential order of episodes but also an allocentric coding of episodes in the environment.

An important scientific question for future studies is how to generate different types of episodes automatically from a sensorimotor flow in an unsupervised or semi-supervised manner. Precoded episodes also seem plausible if we assume that there is a moderate number of types of episodes which can generate an infinite number of variations (children) for different environments and goals.

The cognitive architecture works with objects provided by the Pycharm physical engine. Therefore a translation of visual sensory information into an internal physical model was not required. However, after a short observation of a visual sensory stream, it is possible to approximately grasp basic parameters of relevant objects, like mass, rigidity, collision angle, shape border, etc. These parameters can be transferred into an internal-physics-engine simulation to explore input-output correlations in an episode (e.g. collision).

Pixel-based sensory information is too multidimensional to build direct connections to actions efficiently. It is possible to build deep neural networks which can find efficient strategies based on pixel-based sensory information and reward values (Mnih et al., 2013). However, this approach has two substantial limitations. First, it requires thousands and millions of trials. Second, the learned sensorimotor dependencies are barely possible to use in new environments, and therefore, each new environment requires a new learning procedure. The approach of breaking down a sensorimotor flow into episodes intends to solve both issues. Episodes are reusable in different environments and have to be learned only once.

Goal-directed reasoning occurs at the level of graphical models. Different paths represent different strategies to solve a problem. Improvement of a tactic means improvement in predictions in an episode (node). If some node does not work as it was expected, then either a new strategy represented by a new path should be found, or the node requires several simulations with variance in parameter values to map dependencies better.

Acknowledgments

The author would like to thank Peter König and Holger Finger for useful comments on an earlier version of the manuscript.

Chapter 6. Self-organized Sensorimotor Processing System Based on Context Categorization and Predictions of Consequences of Self-actions

This study has been presented as a poster at Osnabrück Computational Cognition Alliance Meeting 2012 (without proceedings). Osnabrück, Germany.
Melnik, A., Oparin, Y., & König, P. (2012). Self-organized Sensorimotor Processing System Based on Context Categorization and Predictions of Consequences of Self-actions.

AUTHOR CONTRIBUTIONS
Study conception and design: A.M., Y.O.; acquisition of data: A.M.; analysis and interpretation of data: A.M., Y.O.; drafting of manuscript: A.M.; critical revision: A.M., Y.O., P.K.

Self-organized Sensorimotor Processing System Based on Context Categorization and Predictions of Consequences of Self-actions

Andrew Melnik[1], Yury Oparin[2], Peter König[1,3]
Email: andmelnik@uos.de
[1] University of Osnabrück, Germany; [2] Imperial College London, UK;
[3] University Medical Center Hamburg-Eppendorf, Germany.

Abstract

The study investigates how to solve 2 dimensional (2D) sensorimotor tasks using 1D measurements. The question is important for a variety of tasks related to sensorimotor interactions.

Methods

An agent and a reward are placed in a grid environment at different places (Fig. 1A). The agent has three actions: move forward (**F**), turn left (**L**), turn right (**R**). A sensor indicates a change in a distance between the agent and the reward. To achieve the reward the agent learns chains of contexts and sensory-action pairs. When the agent approaches the reward it gets a positive reinforcement.

Figure 1. A. The agent (black) and the reward (green) in a grid environment. **B.** The sensor (red-yellow-green) has three states: approaching (**green**), distance is the same (**yellow**), moving away (**red**). The sensor is updated after the agent commits an action.

Each context has a unique correspondence between an action, a sensory output, and a transition to a new contexts.

When an action leads to an unpredicted result, a new context is set up. Creation of new contexts occur until all occurring combinations of actions, sensory information and resulting contexts are described.

Contexts are connected in a graphical model (Fig. 2B). A sequence of the last five contexts and actions maintains recurrent properties of the system (Fig. 2C). The recurrent mechanism allows to solve ambiguity of the current context based on the recent history.

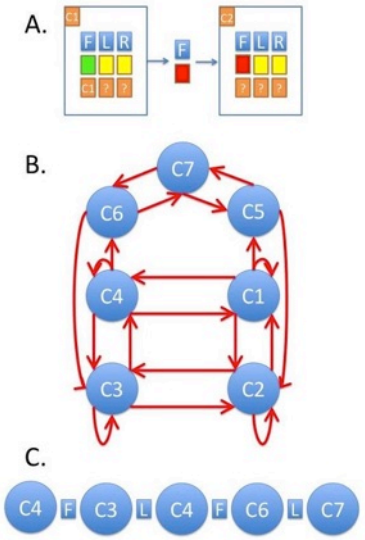


Figure 2. (A) Contexts, (B) graphical models, and (C) recurrent chains.

Results

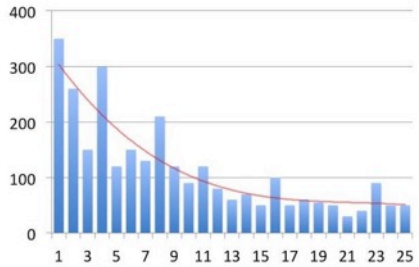


Figure 3. Number of steps required to reach the reward. A new trial starts when a reward cell has been found. X: sequential number of a trial. Y: number of steps per trial. Throughout the trials, the agent learns and exhibits more stable behavior.

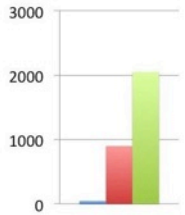


Figure 4. In order to form 2 basic concepts (approaching and moving away) the agent requires about 50 steps (**blue bar**). After that, the agent always moves directly towards the reward. However, near the reward its actions are random. The agent requires about 900 steps to learn transitions between most frequent contexts near the reward cell (**red bar**). To learn the recurrent properties the agent requires about 2000 steps (**green bar**).

Conclusions

The system solves the 2D sensorimotor task using 1D information from the sensor.

The system fits into the extended sensorimotor contingencies (eSMC) theory.

We will apply the investigated principles to computational models of cognitive architectures and autonomous agents.

6.1 Abstract

The study investigates how to solve 2 dimensional (2D) sensorimotor tasks using 1D measurements. The question is important for a variety of tasks related to sensorimotor interactions.

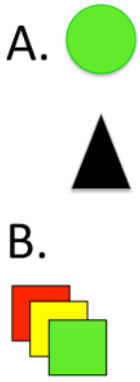


Figure 1. A. The agent (black) and the reward (green) in a grid environment. **B.** The sensor (red-yellow-green) has three states: approaching (green), distance is the same (yellow), moving away (red). The sensor is updated after the agent commits an action.

6.2 Methods

An agent and a reward are placed in a grid environment at different places (Fig. 1A). The agent can perform on of the following actions: move forward (F), turn left (L), turn right (R). A sensor indicates a change in a distance between the agent and the reward. To achieve the reward the agent learns chains of contexts and sensory-action pairs. When the agent approaches the reward it gets a positive reinforcement.

Each context has a unique correspondence between an action, a sensory output, and a transition to a new contexts. When an action leads to an unpredicted result, a new context is set up. Creation of new contexts occur until all occurring combinations of actions, sensory information and resulting contexts are described. Contexts are connected in a graphical model (Fig. 2B). A sequence of the last five contexts and actions maintains recurrent properties of the system (Fig. 2C). The recurrent mechanism allows to solve ambiguity of the current context based on the recent history.

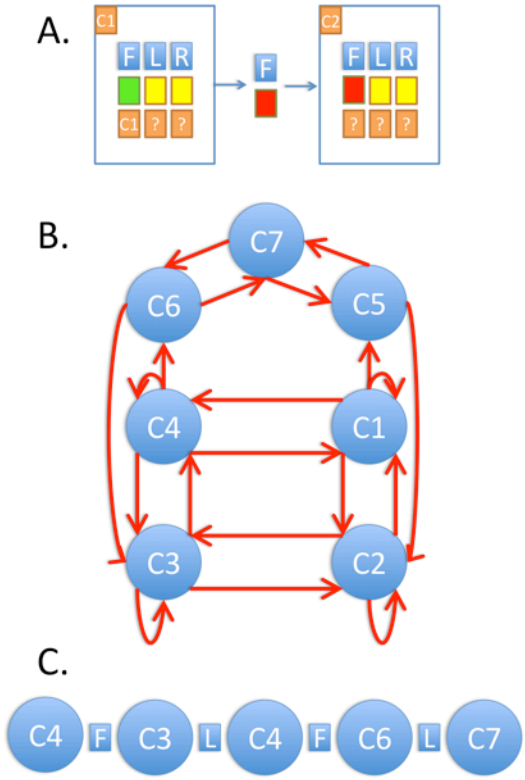


Figure 2. (A) Contexts, **(B)** graphical models, and **(C)** recurrent chains.

6.3 Results

Figure 3 demonstrates that the agent learns and exhibits more stable behavior throughout the trials. Figure 4 demonstrates a number of steps the agent needs to learn different concepts about the environment.

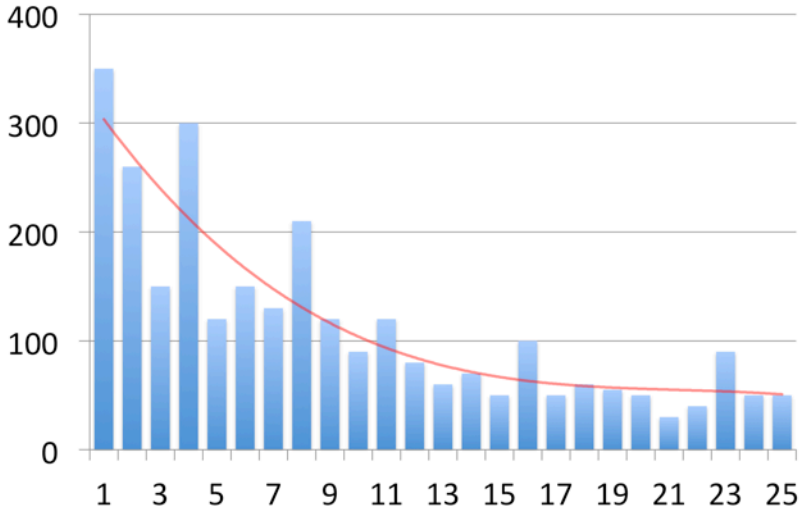


Figure 3. Number of steps required to reach the reward. A new trial starts when a reward cell has been found. X: sequential number of a trial. Y: number of steps per trial. Throughout the trials, the agent learns and exhibits more stable behavior.

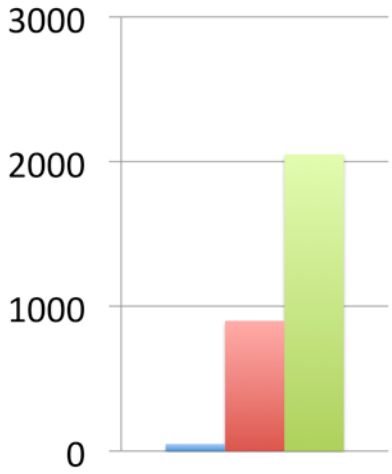


Figure 4. In order to form 2 basic concepts (approaching and moving away) the agent requires about 50 steps (blue bar). After that, the agent always moves directly towards the reward. However, near the reward its actions are random. The agent requires about 900 steps to learn transitions between most frequent contexts near the reward cell (red bar). To learn the recurrent properties the agent requires about 2000 steps (green bar).

6.4 Conclusions

The system solves the 2D sensorimotor task using 1D information from the sensor. The system fits into the extended sensorimotor contingencies (eSMC) theory. The investigated principles are applicable to computational models of cognitive architectures and autonomous agents.

Chapter 7. An architecture of the hierarchically self-learning artificial cognitive agent in a pathfinding task which is utilizing egocentric visual and auditory information

This study has been presented as a poster at the 5th International Conference on Cognitive Systems (in proceedings). Vienna, Austria.

Melnik, A., König, P. (2012). An architecture of the hierarchically self-learning artificial cognitive agent in a pathfinding task which is utilizing egocentric visual and auditory information.
http://cogsys2012.acin.tuwien.ac.at/proceedings.html

AUTHOR CONTRIBUTIONS
Study conception and design: A.M., P.K.; acquisition of data: A.M.; analysis and interpretation of data: A.M.; drafting of manuscript: A.M.; critical revision: A.M., P.K.

An architecture of the hierarchically self-learning artificial cognitive agent in a pathfinding task which is utilizing egocentric visual and auditory information

Andrew Melnik[1], Peter König[1].
[1]University Osnabrück, Germany
Email: andmelnik@uos.de

UNIVERSITÄT OSNABRÜCK

Abstract

The project aims to build a cognitive agent that adapts to an artificial environment through experience and accumulates cognitive abilities in a hierarchical sequence; the higher levels start to learn and function only when the lower levels demonstrate stable activity. The range of what the agent can learn is hardwired by the architecture of the interconnections between the sensory, learning and goal modules (Fig. E). The agent learns strategies to gain more reward. The project also aims to build a model of integration of multisensory information. Therefore, auditory information was represented by a distance value between a reward point and the agent.

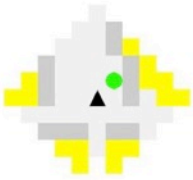
Methods

- ❑ The agent (Fig. E) and the environment (Fig. A) were implemented in Python.
- ❑ The agent gets auditory (distance between a reward point and the agent) and visual sensory information (Fig. B & C).
- ❑ The agent has 3 actions: step forward, turn left, turn right.
- ❑ The modules of the agent use positive and negative reinforcement learning, weighted graphs, and neural networks.

A.

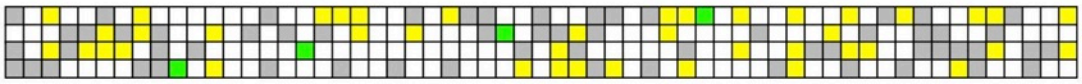
The map of a labyrinth.

B.



Egocentric view. Light gray: free area (integer value=0). Dark gray: wall (integer value=1). Green: reward (integer value=2). Yellow: unknown are behind walls (integer value=3)

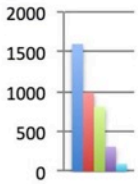
C.



Four sequential egocentric views of the agent (Fig. B) represented as a 4x59 matrix of integers. The color coding is the same as in Fig. B.

Results

D.

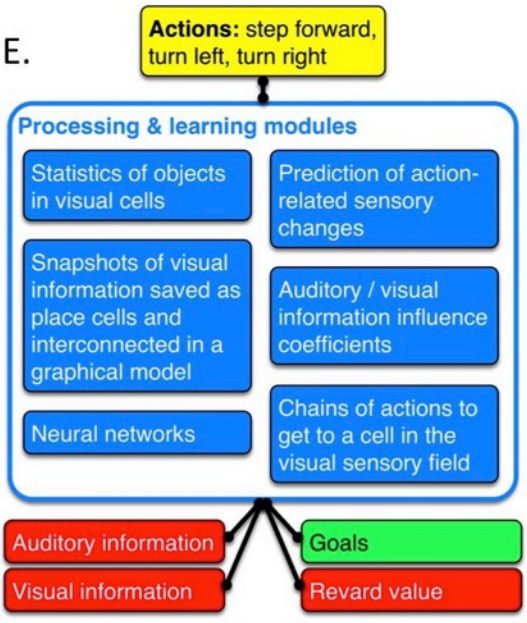


To assess the performance, we counted the average numbers of actions required to reach the goal after each learning level (Fig. D). **Blue:** random actions. **Red:** the agent learned reward properties of objects. **Green:** the agent learned how to move to a certain cell in the egocentric view. **Violet:** the agent learned to move towards unexplored regions when no reward is in the visual sensory field. **Light-blue:** the agent was taught to choose the left path at each path-bifurcation point. The agent and a reward were randomly placed in a 16x16 labyrinth. The bars represent the average value of 20 trials for each learning level.

Conclusions

- The architecture can be extended and embedded in real world systems.
- The agent is able to modulate and integrate multisensory information gained from the environment.
- With the benefit of experience all these mechanisms are reliable and
- The agent can explore a labyrinth with the benefit of the learning levels more efficiently than in a random walk (Fig. D).

E.



The architecture of interconnections between the modules. Connections to the blue outline are shared between all the blue modules.

Acknowledgements

I want to thank my colleagues Robert Märtin, for helpful advices, and Laura Saez, for tips on academic English.

7.1 Abstract

The project aims to build a cognitive agent that adapts to an artificial environment through experience and accumulates cognitive abilities in a hierarchical sequence. The higher levels start to learn and function only when the lower levels demonstrate stable activity. The range of what the agent can learn is hardwired by the architecture of the interconnections between the sensory, learning and goal modules (Fig. 1). The agent learns strategies to gain more reward. The project also aims to build a model of integration of multisensory information. Therefore, auditory information was represented by a distance value between a reward point and the agent.

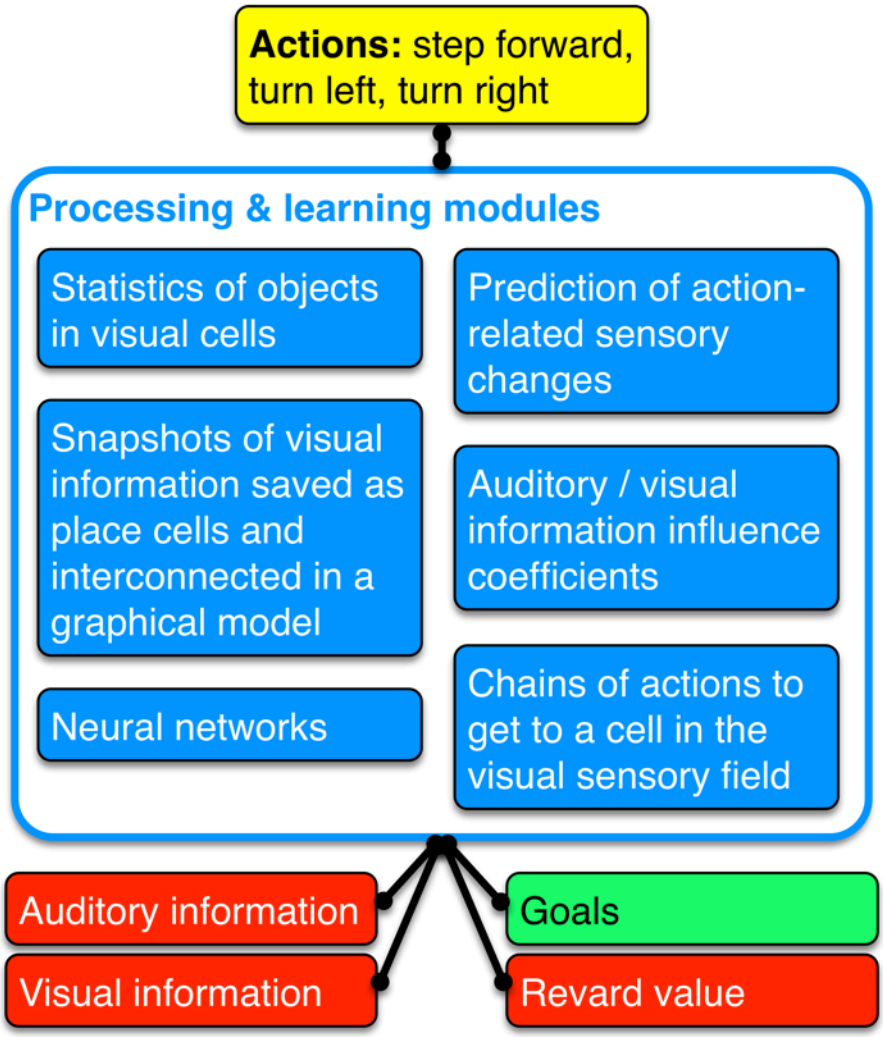


Figure 1. The architecture of interconnections between the modules. Connections to the blue outline are shared between all the blue modules

7.2 Methods

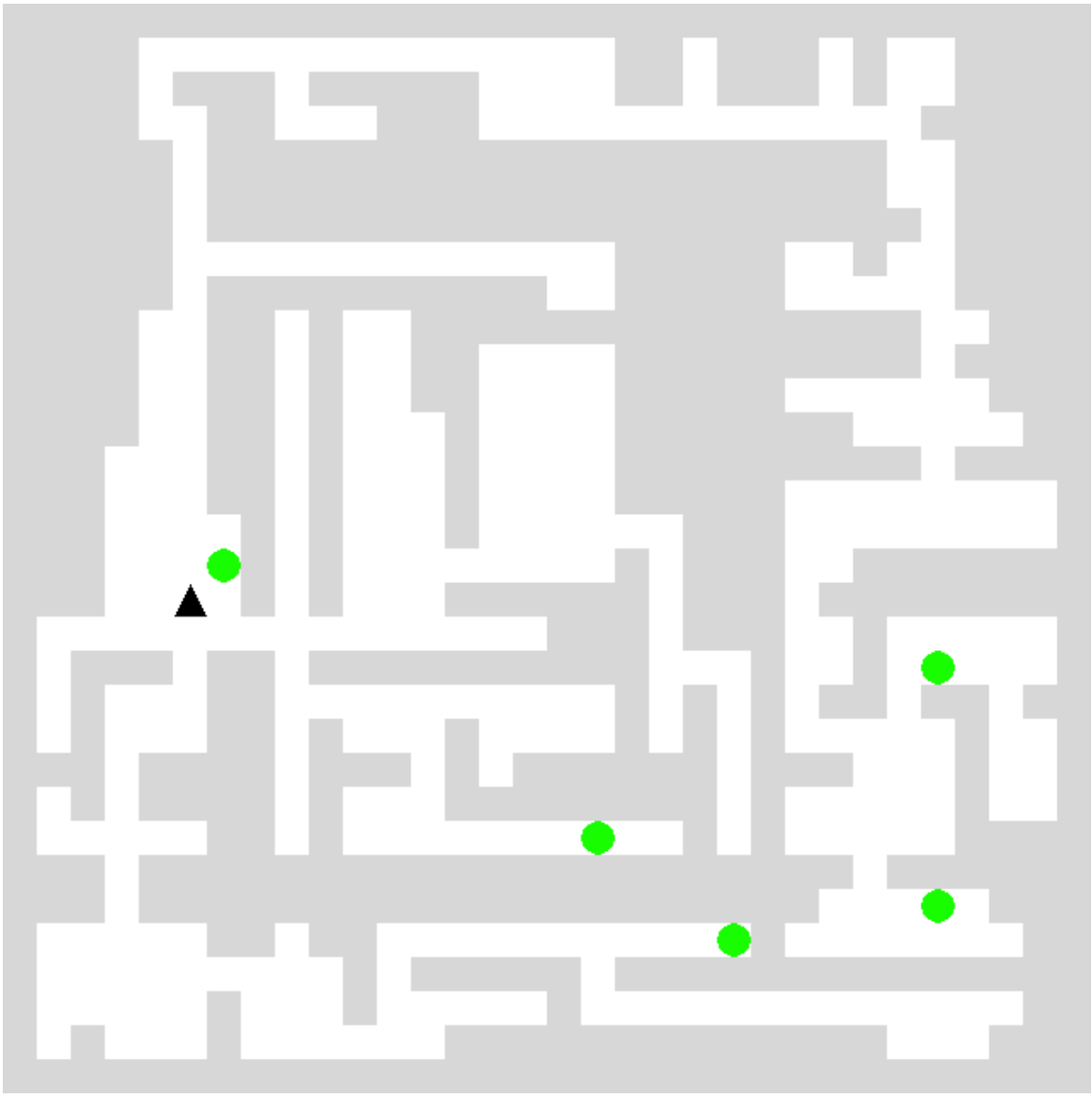


Figure 2. A map of a labyrinth.

The agent (Fig. 1) and the environment (Fig. 2) were implemented in Python. The agent gets auditory (distance between a reward point and the agent) and visual sensory information (Fig. 3 & 4). The agent can perform 3 actions: step forward, turn left, turn right. The modules of the agent use positive and negative reinforcement learning, weighted graphs, and neural networks.

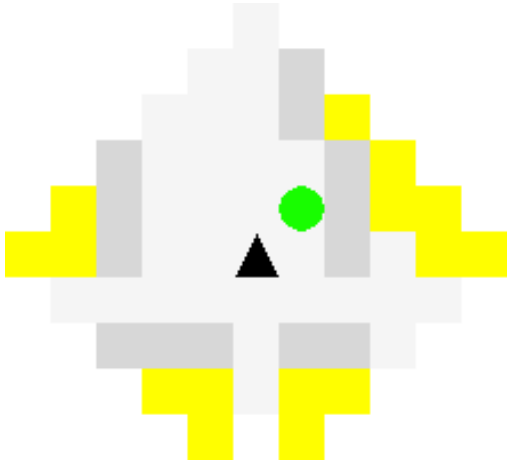


Figure 3. Egocentric view. Light gray: free area (integer value=0). Dark gray: wall (integer value=1). Green: reward (integer value=2). Yellow: unknown are behind walls (integer value=3)

Figure 4. Four sequential egocentric views of the agent (Fig. B) represented as a 4x59 matrix of integers. The color coding is the same as in Fig. B.

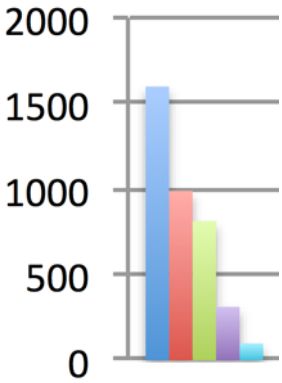


Figure 5. Performance. OY: an average numbers of actions required to reach the goal after each learning level. Blue: random actions. Red: the agent learned reward properties of objects. Green: the agent learned how to move to a certain cell in the egocentric view (Fig. 3). Violet: the agent learned to move towards unexplored regions when no reward is in the visual sensory field. Light-blue: the agent was taught to choose the left path at each path-bifurcation point.

7.3 Results

To assess the performance, we counted the average numbers of actions required to reach the goal after each learning level (Fig. 5). The agent and a reward were randomly placed in a 16x16 labyrinth. The bars in Figure 5 represent the average value of 20 trials for each learning level.

7.4 Conclusions

The architecture can be extended and embedded in real world systems. The agent is able to modulate and integrate multisensory information gained from the environment. The agent can explore a labyrinth with the benefit of the learning levels more efficiently than in a random walk (Fig. 5).

7.5 Acknowledgements

I want to thank my colleagues Robert Märtin, for helpful advices, and Laura Saez, for tips on academic English.

Chapter 8. Discussion and outlook

The motivation for the studies in the preceding chapters was an intention to understand better how sensory information translates into motor actions in the human brain and in cognitive architectures. In Chapter 2, we demonstrated EEG correlates which supported a holistic point of view on sensorimotor processing in the human brain. In Chapter 4, we described a balance between memorization and active sampling of visual information in a sensorimotor task. The study in Chapter 3 was motivated by contemporary EEG experiments under natural conditions while participants are performing sensorimotor tasks with free movements. We compared mobile EEG systems which allow free movements against research-grade EEG systems and proposed a benchmark for testing new EEG systems. Finally, we proposed cognitive architectures which can perform navigational (Chapters 6 and 7) and intuitive physics tasks (Chapter 5). Results of these studies advanced our knowledge of sensorimotor processing in the human brain and in cognitive architectures and exposed open questions for future research.

I discussed results of these studies, their significance, and limitations in relevant chapters of this thesis. Therefore I will focus on methodological limitations, open questions and possible further developments of studies presented in this thesis.

EEG is a valuable tool for studying sensorimotor processing in the human brain. It has an excellent temporal resolution, which is necessary for detecting sensorimotor coupling steps. It would not be possible to distinguish between Sensory, Motor, and Sensorimotor sources (Chapter 2) in the brain with fMRI, due to not sufficient time resolution. EEG recording headsets are allowing for movements during recordings, and it is more convenient for subjects. It would be barely possible to conduct the experiment in Chapter 2, which lasted about two hours, in MEG or fMRI scanner. However, general disadvantages of EEG are a poor spatial resolution, low signal to noise ratio, high susceptibility to muscle activity, and signal artifacts due to changes in contact properties (e.g., jolting). Moreover, a diversity of EEG systems is posing a general question of whether different EEG systems produce EEG data of the same quality (Chapter 3). Comparison of a variance due to EEG systems against variance due to subjects and repeated sessions helped to estimate the quality of EEG data produced by new mobile EEG systems. However, we still do not know the true signal. These limitations restrict possible experimental paradigms and complicate data analysis of EEG recordings.

EEG analysis of channel data does not allow for differentiation between sources of EEG activities in the brain, as a single channel can pick up a mixed signal. ICA transforms EEG data from channel space into source space, which is allowing for differentiation of sources of EEG activity in the brain. However, a general limitation of ICA for large scale studies is a necessity to conduct ICA for each EEG recording in the study separately, and resulting ICs do not coincide between different recordings. Therefore, it is not possible to select the same IC for different EEG recordings and compare activities, as it is possible for EEG channels. This narrows scopes of ICA for EEG data analysis, in many cases, to artifact cleaning applications.

In the study in Chapter 2 ("EEG Correlates Of Sensorimotor Processing: Independent Components Involved In Sensory And Motor Processing"), we demonstrated EEG correlates of sensorimotor processing in the human brain. However, as an extension of this study, we would vary a set of stimuli (e.g., uni-modal audio, uni-modal video, etc.), tasks, and motor response (e.g., go/no go, forced choice categorization, counting, etc.) to identify scopes of the effect. We would also conduct this experiment using MEG and fMRI facilities. MEG provides orthogonal to EEG observations of dipolar activity in the brain. And fMRI provides better spatial resolution than MEG or EEG. Different brain imaging techniques collect complementary information about the brain activity, and, therefore, help to build a more comprehensive model of the cognitive processes in the human brain.

In the study in Chapter 3 ("Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data?"), we analyzed factors of variance using ERP paradigms. However, we did not analyze frequency (Fourier transform) or time-frequency domains. As an extension of this study, we would record and analyze frequency related paradigms (e.g., alpha activity during eye-closed, etc.). Another extension of this study is to conduct more recordings per factor of variance. One option for that could be to reject the factor of Session and focus solely on the factors of Subject and System. Then, in a new study with a similar number of EEG recordings, we could have, e.g., 7x7 design (49 recordings) for ANOVA which embraces a broader sample of EEG systems (7 EEG systems) and subjects (7 subjects).

In the study in Chapter 4 ("The price of Saccades"), we described the balance between memorization and active sampling of visual information. The cost for a new sample of visual information that participants had to "pay" was a short time-delay until a blocks-model appeared on a screen. However, in the study, we used only the 700 ms delay, which is allowing only for a linear model of the tradeoff. Therefore, as a further development of the study, I propose to use a set of delays (e.g., 100 ms, 300 ms, 500 ms, 700 ms, 900 ms, etc.). Another parameter to vary is a complexity of a model of blocks (e.g., 2, 4, 8, 16, 32 blocks). Moreover, we could introduce a parameter of distraction (e.g., auditory or visual distractors) or a parallel subtask.

That would allow for building a more precise and comprehensive model of the balance between memorization and active sampling of visual information. Then, it could be possible to see a point of a strategy shift and dynamic adoption of eye-movement behavior.

Studies of sensorimotor processing on humans provide theoretical knowledge about the cognitive processes in the human brain. This knowledge constitutes a valuable theoretical background for a development of functional artificial cognitive architectures. In the preceding chapters of this thesis (Chapters 5, 6, and 7), we implemented several cognitive architectures and highlighted problems which we encountered.

In the studies in Chapters 5, 6, and 7, we used 2D computer-game environments, because, in comparison to robotic platforms, computer games are more convenient for training and testing of cognitive architectures. Moreover, computer games reflect human cognitive capacities and therefore can serve as valuable benchmarks for cognitive architectures. Generalization and concepts understanding by cognitive agents were some of the generic problems which we encountered in these studies. Approaches to finding a solution for abstract goals is still an open question in the field of cognitive architectures. For examples, abstract goals can be split up into a series of subtasks. A neural network can learn to solve a subtask, and an ensemble of specialized neural networks would be capable of performing more general tasks. Neural networks are capable of function approximation between an input and output (supervised learning or reinforcement learning). However, the input and output are usually carefully selected by a human.

The motivation behind all studies presented in this thesis was to learn more about sensorimotor processing in the human brain and to implement these theoretical outlines into cognitive architectures. The study in Chapter 2 ("EEG Correlates Of Sensorimotor Processing: Independent Components Involved In Sensory And Motor Processing") demonstrated brain areas responsible for sensorimotor coupling. The results support a holistic point of view on sensorimotor processing and related theories (Common Coding, Sensorimotor Contingencies, Embodiment). The study in Chapter 3 ("Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data?") allowed us to quantify variance of EEG measurements, which is an important estimation, as EEG techniques are widely used to investigate sensorimotor processing in the human brain. The study in Chapter 4 ("The price of saccades"), highlighted the balance between memorization and active sampling of visual information and supported a view on the world as an outside memory. In the study in Chapter 5 ("A Cognitive Architecture For Spatial Reasoning"), I proposed a connection between eye fixations in humans and breaking down of long term goals into a sequence of episodes in cognitive architectures. The study in Chapter 6 ("Self-organized Sensorimotor Processing System Based on Context Categorization and

Predictions of Consequences of Self-actions") proposed a way to deal with contextual uncertainty by embracing actions into a perceptual palette. The study in Chapter 7 ("An architecture of the hierarchically self-learning artificial cognitive agent in a pathfinding task which is utilizing egocentric visual and auditory information") proposed a cognitive architecture of an artificial agent with an egocentric view in a pathfinding task. This study also demonstrated an influence of inbuilt "instincts" and strategies onto goal directed performance. Studies presented in this thesis contributed to a deeper understanding of sensorimotor processing in the human brain and in cognitive architectures.

Appendix

Publications, proceedings, posters, unpublished works

PUBLICATIONS IN SCIENTIFIC JOURNALS

(Chapter 2) Melnik, A., Hairston, W. D., Ferris, D. P., & König, P. (2017). EEG correlates of sensorimotor processing: independent components involved in sensory and motor processing. Scientific Reports, 7.
http://dx.doi.org/10.1038/s41598-017-04757-8

(Chapter 3) Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., & König, P. (2017). Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data?. Frontiers in human neuroscience, 11, 150. https://doi.org/10.3389/fnhum.2017.00150

CONFERENCE PROCEEDINGS

(Chapter 4) *Melnik, A., *Schüler, F., Rothkopf, C., & König, P. (2017). The Price of Saccades. European Conference on Visual Perception 2017, Berlin, Germany.
http://journals.sagepub.com/page/pec/collections/ecvp-abstracts/index/ecvp-2017
(*first two authors contributed equally)

(Chapter 7) Melnik, A., König, P. (2012). An architecture of the hierarchically self-learning artificial cognitive agent in a pathfinding task which is utilizing egocentric visual and auditory information. 5th International Conference on Cognitive Systems. Vienna, Austria. http://cogsys2012.acin.tuwien.ac.at/proceedings.html

POSTER PRESENTATIONS (WITHOUT PROCEEDINGS)

(Chapter 6) Melnik, A., Oparin, Y., & König, P. (2012). Self-organized Sensorimotor Processing System Based on Context Categorization and Predictions of Consequences of Self-actions. Osnabrück Computational Cognition Alliance Meeting. Osnabrück, Germany.

UNPUBLISHED WORKS

(Chapter 5) Melnik, A. (2017). A Cognitive Architecture For Spatial Reasoning.

Teaching & Supervision

Co-supervisor of a Master Thesis:
"ICA analysis of multisensory processing in an audio-visual tapping task”
Student: Jessika Schwandt, 2014

Co-teaching a seminar:
"Extending Sensorimotor Contingencies to Cognitive Architectures"
Winter semester 2012/2013

Co-teaching a seminar:
"MATLAB in Cognitive Science"
Winter semester 2012/2013

Co-teaching a seminar:
"Data analysis in Cognitive Science"
Summer semester 2012

Disclaimer

All experiments reported in this thesis conform with the Declaration of Helsinki and have been approved by the ethics committees of the respective institution (University of Osnabrück). I hereby confirm that I wrote this thesis independently and that I have not made use of resources other than those indicated. I have significantly contributed to all materials used in this thesis. Furthermore, this thesis was neither published in Germany nor abroad, except the parts indicated above, and has not been used to fulfill any other examination requirements.

Figure References

All figures presented in this thesis are original work by the authors of the respective chapters.

Acknowledgments

This thesis would not have been possible without the help of many others. First and foremost I want to express my gratitude to my dissertation supervisor, Professor Peter König for encouraging support, continued faith, patience, and contribution. Peter, thank you for many lessons learned, all the great opportunities I was given, and countless POMs we had. I will be forever grateful for it.

I am very much thankful to Dr. David Hairston and Prof. Daniel Ferris whose guidance and suggestions have contributed to studies presented in this thesis. Also, it was a great experience of overseas collaboration.

I would like to thank the members of the NeuroBioPsychology Group as well as other colleagues for the great working environment, critical thinking, and open discussions during Tea Times and seminars. In particular, I wish to thank Jessika Schwandt, Holger Finger, Caspar Goeke, Nicolas Kuske, Benedikt Ehinger, Anna Lisa Gert, Tim Kietzmann, Petr Legkov, Krzysztof Izdebski, Silke Kärcher, Marcel Schumacher, Vivien Radke, Ashima Keshava, Maria Sokotushchenko, Felix Schüler, Allison Moreno-Drexler, Daniel Ditinyak, and Laura Saez. I had the pleasure of working with you guys and learn from and with you. I would also like to thank former and current members of the Graduate Program Committee of the Institute of Cognitive Science and Prof. Peter König for their financial support and possibilities to attend all relevant scientific events.

Finally, I would like to thank my family and in particular my parents, Dmitry and Svetlana, for their moral support, love, and advice.

Bibliography

Abrams, R. A., Meyer, D. E., & Kornblum, S. (1990). Eye-hand coordination: oculomotor control in rapid aimed limb movements. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 248–267. https://doi.org/10.1037/0096-1523.16.2.248

Allen, J. S., Damasio, H., & Grabowski, T. J. (2002). Normal neuroanatomical variation in the human brain: An MRI-volumetric study. *American Journal of Physical Anthropology*, *118*(4), 341–358. https://doi.org/10.1002/ajpa.10092

Andersen, S. K., Hillyard, S. a, & Müller, M. M. (2013). Global facilitation of attended features is obligatory and restricts divided attention. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(46), 18200–7. https://doi.org/10.1523/JNEUROSCI.1913-13.2013

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe? How Can the Human Mind Occur in the Physical Universe?* https://doi.org/10.1093/acprof:oso/9780195324259.001.0001

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception & Psychophysics*, *57*(3), 305–317. https://doi.org/10.3758/BF03213056

Aspinall, P., Mavros, P., Coyne, R., & Roe, J. (2013). The urban brain: analysing outdoor physical activity with mobile EEG. *British Journal of Sports Medicine*, *online*(November 2015), 1–6. https://doi.org/10.1136/bjsports-2012-091877

Bach-y-Rita, P. (1983). Tactile vision substitution: past and future. *The International Journal of Neuroscience*, *19*(1–4), 29–36. https://doi.org/10.3109/00207458309148643

Bach-y-Rita, P., & W. Kercel, S. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2003.10.013

Badillo, L., Ponomaryov, V., Ramos, E., & Igartua, L. (2003). Low noise multichannel amplifier for portable EEG biomedical applications. *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, *4*, 3309–3312. https://doi.org/10.1109/IEMBS.2003.1280852

Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, *13*(2), 111–122. https://doi.org/10.1016/0031-3203(81)90009-1

Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *337*(1281), 331-338-339. https://doi.org/10.1098/rstb.1992.0111

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory Representations in Natural Tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66–80.

https://doi.org/10.1162/jocn.1995.7.1.66

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *The Behavioral and Brain Sciences*, *20*(4), 723-742-767. https://doi.org/10.1017/S0140525X97001611

Banich, M., Milham, M., Atchley, R., Cohen, N., Webb, A., Wszalek, T., … Magin, R. (2000). fMri studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience*, *12*(6), 988–1000. https://doi.org/10.1162/08989290051137521

Battaglia, P. W. (2008). *Probabilistic sensorimotor processing: Overcoming uncertainty and ambiguity to improve behavior*. University of Minnesota.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–32. https://doi.org/10.1073/pnas.1306572110

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2007). *Neuroscience*. *Lippincott Williams & Wilkins*. https://doi.org/978-0878937257

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science : A Journal of the American Psychological Society / APS*, *13*(4), 370–374. https://doi.org/10.1111/j.0956-7976.2002.00466.x

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Betz, T., Kietzmann, T. C., Wilming, N., & Koenig, P. (2010). Investigating task-dependent top-down effects on overt visual attention. Journal of vision, 10(3), 15-15.

Berger, H. (1935). Das Elektrenkephalogramm des Menschen. *Die Naturwissenschaften*, *23*(8), 121–124. https://doi.org/10.1007/BF01496966

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*(June), 16. https://doi.org/10.3389/fninf.2015.00016

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. *Pattern Recognition* (Vol. 4). https://doi.org/10.1117/1.2819119

Bohg, J., Morales, A., Asfour, T., & Kragic, D. (2014). Data-Driven Grasp Synthesis—A Survey. *IEEE Transactions on Robotics*, *30*(2), 289–309. https://doi.org/10.1109/TRO.2013.2289018

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Botvinick, M. M., & Cohen, J. D. (2014). The computational and neural basis of cognitive control: Charted territory and new frontiers. *Cognitive Science*, *38*(6), 1249–1285. https://doi.org/10.1111/cogs.12126

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and

anterior cingulate cortex: An update. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.10.003

Carlton, L. G. (1981). Processing visual feedback information for movement control. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(5), 1019. https://doi.org/10.1037/0096-1523.7.5.1019

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain*. https://doi.org/10.1093/brain/awl004

Chase, W. G., & Simon, H. A. (1973). Chase, W. G., & Simon, H. A. (1973). Perception in chess. Cognitive Psychology, 4, 55–81. *Cognitive Psychology*, *4*, 55–81.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, *58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., & Pesaran, B. (2014). Sensory-motor transformations for speech occur bilaterally. *Nature*, *507*(7490), 94–8. https://doi.org/10.1038/nature12935

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed-processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–361. https://doi.org/10.1037//0033-295x.97.3.332

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and Brain Sciences*, *24*(1), 87-114–85. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? In *Progress in Brain Research* (Vol. 6123, pp. 323–338). https://doi.org/10.1016/S0079-6123(07)00020-9

Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: a motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1673–1692. https://doi.org/10.1037/0096-1523.25.6.1673

Davis, P. A. (1939). Effects of Acoustic Stimuli on the Waking Human Brain. *Journal of Neurophysiology*, *2*(6), 494--499. Retrieved from http://doi.apa.org/?uid=1940-01707-001

De Sanctis, P., Butler, J. S., Malcolm, B. R., & Foxe, J. J. (2014). Recalibration of inhibitory control systems during walking-related dual-task interference: A Mobile Brain-Body Imaging (MOBI) Study. *NeuroImage*, *94*, 55–64. https://doi.org/10.1016/j.neuroimage.2014.03.016

De Vos, M., Kroesen, M., Emkes, R., & Debener, S. (2014). P300 speller BCI with a mobile EEG system: comparison to a traditional amplifier. *Journal of Neural Engineering*, *11*(3), 36008. https://doi.org/10.1088/1741-2560/11/3/036008

Delorme, A., & 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), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0030135

Di Russo, F., Pitzalis, S., Aprile, T., & Spinelli, D. (2005). Effect of practice on brain activity: An investigation in top-level rifle shooters. *Medicine and Science in Sports and Exercise*, *37*(9), 1586–1593. https://doi.org/10.1249/01.mss.0000177458.71676.0d

Dickinson, S. J. (1999). Object representation and recognition. In *What is cognitive science 7* (pp. 172–207).

Dixit, R. (2011). Meditation Training and Neurofeedback Using a Personal EEG Device. *PLoS Computational Biology*, 15–16.

Elliott, D., Helsen, W. F., & Chua, R. (2001). A century later: Woodworth's (1899) two-component model of goal-directed aiming. *Psychological Bulletin*, *127*(3), 342–357. https://doi.org/10.1037/0033-2909.127.3.342

Engel, A. K., Maye, A., Kurthen, M., & König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2013.03.006

Erol, K., Hendler, J., & Nau, D. S. (1994). *Semantics for Hierarchical Task-Network Planning*. *Technical Report CS-TR-3239*.

Fidler, S., & Leonardis, A. (2007). Towards scalable representations of object categories: Learning a hierarchy of parts. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. https://doi.org/10.1109/CVPR.2007.383269

Fitts, P. M., & Seeger, C. M. (1953). S R Compatibility - Spatial Characteristics of Stimulus and Response Codes. *Journal of Experimental Psychology*, *46*(3), 199–210. https://doi.org/Doi 10.1037/H0062827

Fize, D., Boulanouar, K., Chatel, Y., Ranjeva, J. P., Fabre-Thorpe, M., & Thorpe, S. (2000). Brain areas involved in rapid categorization of natural images: an event-related fMRI study. *NeuroImage*, *11*(6 Pt 1), 634–643. https://doi.org/10.1006/nimg.2000.0585

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), 319–325. https://doi.org/10.1002/wcs.1226

Gallese, V., & Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3), 455–79. https://doi.org/10.1080/02643290442000310

Gargiulo, G., Calvo, R. A., Bifulco, P., Cesarelli, M., Jin, C., Mohamed, A., & van Schaik, A. (2010). A new EEG recording system for passive dry electrodes. *Clinical Neurophysiology*, *121*(5), 686–693. https://doi.org/10.1016/j.clinph.2009.12.025

Gaver, W. (1991). Technology affordances. *ACM*, 79–84. https://doi.org/10.1145/108844.108856

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, *30*(1), 535–574.

https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gramann, K., Gwin, J. T., Bigdely-Shamlo, N., Ferris, D. P., & Makeig, S. (2010). Visual evoked responses during standing and walking. *Frontiers in Human Neuroscience*, *4*(October), 202. https://doi.org/10.3389/fnhum.2010.00202

Gramann, K., Jung, T. P., Ferris, D. P., Lin, C. T., & Makeig, S. (2014). Toward a new cognitive neuroscience: modeling natural brain dynamics. *Frontiers in Human Neuroscience*, *8*, 1–3. https://doi.org/10.3389/fnhum.2014.00444

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review*, *101*(2), 336–342. https://doi.org/10.1037/0033-295X.101.2.336

Guger, C., Krausz, G., Allison, B. Z., & Edlinger, G. (2012). Comparison of dry and gel based electrodes for P300 brain-computer interfaces. *Frontiers in Neuroscience*, (MAY). https://doi.org/10.3389/fnins.2012.00060

Gwin, J. T., Gramann, K., Makeig, S., & Ferris, D. P. (2011). Electrocortical activity is coupled to gait cycle phase during treadmill walking. *NeuroImage*, *54*(2), 1289–1296. https://doi.org/10.1016/j.neuroimage.2010.08.066

Hairston, W. D., & Lawhern, V. (2015). How low can you go? Empirically assessing minimum usable DAQ performance for highly fieldable EEG systems. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 9183, pp. 221–231). https://doi.org/10.1007/978-3-319-20816-9_22

Hairston, W. D., Proie, R., Conroy, J., & Nothwang, W. (2015). Batteryless Electroencephalography (EEG): Subthreshold Voltage System-on-a-Chip (SoC) Design for Neurophysiological Measurement.

Hairston, W. D., Whitaker, K. W., Ries, A. J., Vettel, J. M., Cortney Bradford, J., Kerick, S. E., & McDowell, K. (2014). Usability of four commercially-oriented EEG systems. *Journal of Neural Engineering*, *11*(4), 46018. https://doi.org/10.1088/1741-2560/11/4/046018

Hammon, P. S., Makeig, S., Poizner, H., Todorov, E., & de Sa, V. R. (2008). Predicting reaching targets from human EEG. *IEEE Signal Processing Magazine*, *25*(1), 69–77. https://doi.org/10.1109/MSP.2008.4408443

Hansen, P., Kringelbach, M., & Salmelin, R. (2010). *MEG: An introduction to methods*. *MEG: An Introduction to Methods*. https://doi.org/10.1093/acprof:oso/9780195307238.001.0001

Harrison, R. R., & Charles, C. (2003). A low-power low-noise CMOS amplifier for neural recording applications. *IEEE Journal of Solid-State Circuits*, *38*(6), 958–965. https://doi.org/10.1109/JSSC.2003.811979

Hawkins, J., & George, D. (2006). Hierarchical temporal memory: Concepts, theory and terminology. *Whitepaper, Numenta Inc*, 20. https://doi.org/10.1109/IEMBS.2006.260909

Hickok, G., Okada, K., & Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology*, *101*(February 2009), 2725–2732. https://doi.org/10.1152/jn.91099.2008

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, a W., & Evans, a C. (2015). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography*, *22*(2), 324–333. https://doi.org/10.1097/00004728-199803000-00032

Issen, L. A., & Knill, D. C. (2012). Decoupling eye and hand movement control: visual short-term memory influences reach planning more than saccade planning. *Journal of Vision*, *12*(1), 3.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*(Box 1), 304–313. https://doi.org/10.1038/nrn2332

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. Nature reviews. Neuroscience, 2(3), 194.

Kärcher, S. M., Fenzlaff, S., Hartmann, D., Nagel, S. K., & König, P. (2012). Sensory Augmentation for the Blind. *Frontiers in Human Neuroscience*, *6*(March), 37. https://doi.org/10.3389/fnhum.2012.00037

Kaspar, K., & Koenig, P. (2011). Viewing behavior and the impact of low-level image properties across repeated presentations of complex scenes. Journal of Vision, 11(13), 26-26.

Kaspar, K., König, S., Schwandt, J., & König, P. (2014). The experience of new sensorimotor contingencies by sensory augmentation. *Consciousness and Cognition*, *28*(1), 47–63. https://doi.org/10.1016/j.concog.2014.06.006

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*. https://doi.org/10.1016/S0959-4388(99)00028-8

Kerns, J. G. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science*, *303*(5660), 1023–1026. https://doi.org/10.1126/science.1089910

König, P., Wilming, N., Kaspar, K., Nagel, S. K., & Onat, S. (2013). Predictions in the light of your own action repertoire as a general computational principle. *Behavioral and Brain Sciences*, *36*(3), 219–220. https://doi.org/10.1017/S0140525X12002294

Kowler, E. (2011). Eye movements: the past 25 years. *Vision Research*, *51*(13), 1457–83. https://doi.org/10.1016/j.visres.2010.12.014

Kowler, E., & Anton, S. (1987). Reading twisted text: Implications for the role of saccades. *Vision Research*, *27*(1), 45–60. https://doi.org/10.1016/0042-6989(87)90142-8

Kraus, N., & Nicol, T. (2009). Auditory Evoked Potentials. In *Encyclopedia of Neuroscience* (pp. 214–218). https://doi.org/10.1007/978-3-540-29678-2_433

Lacadie, C. M., Fulbright, R. K., Rajeevan, N., Constable, R. T., & Papademetris, X. (2008). More accurate Talairach coordinates for neuroimaging using non-linear registration. *NeuroImage*, *42*(2), 717–725. https://doi.org/10.1016/j.neuroimage.2008.04.240

Laird, J. E. (2012). *The Soar cognitive architecture*. MIT Press.

Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A.

C., … Mazziotta, J. C. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. In *Human Brain Mapping* (Vol. 5, pp. 238–242). https://doi.org/10.1002/(SICI)1097-0193(1997)5:4<238::AID-HBM6>3.0.CO;2-4

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., … Fox, P. T. (2000). Automated Talairach Atlas labels for functional brain mapping. *Human Brain Mapping*, *10*(3), 120–131. https://doi.org/10.1002/1097-0193(200007)10:3<120::AID-HBM30>3.0.CO;2-8

Lau, T. M., Gwin, J. T., & Ferris, D. P. (2012). How Many Electrodes Are Really Needed for EEG-Based Mobile Brain Imaging? *Journal of Behavioral and Brain Science*, *2*(3), 387–393. https://doi.org/10.4236/jbbs.2012.23044

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain*. https://doi.org/10.1093/brain/awt162

Liao, L.-D., Chen, C.-Y., Wang, I.-J., Chen, S.-F., Li, S.-Y., Chen, B.-W., … Lin, C.-T. (2012). Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors. *Journal of NeuroEngineering and Rehabilitation*. https://doi.org/10.1186/1743-0003-9-5

Lin, R., Lee, R.-G., Tseng, C.-L., Wu, Y.-F., & Jiang, J.-A. (2006). Design and implementation of wireless multi-channel EEG recording system and study of EEG clustering method. *Biomedical Engineering - Applications, Basis and Communications*, *18*(6), 276–283.

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, *38*(5), 701–711. https://doi.org/10.1016/S0028-3932(99)00106-2

Lippa, Y. (1996). A Referential coding Explanation for Compatibility Effects of Physically Orthogonal Stimulus and Response Dimensions. *The Quarterly Journal of Experimental Psychology Section A*, *49*(July 2014), 950–971. https://doi.org/10.1080/713755676

Liu, Y., Jiang, X., Cao, T., Wan, F., Mak, P. U., Mak, P. I., & Vai, M. I. (2012). Implementation of SSVEP based BCI with Emotiv EPOC. In *Proceedings of IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems,VECIMS* (pp. 34–37). https://doi.org/10.1109/VECIMS.2012.6273184

Luck, S. J. (2005). An Introduction to the Event-Related Potential Technique. *Monographs of the Society for Research in Child Development*, *78*(3), 388. https://doi.org/10.1118/1.4736938

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.03.008

Makeig, S., Delorme, A., Westerfield, M., Jung, T. P., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2004). Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biology*, *2*(6). https://doi.org/10.1371/journal.pbio.0020176

Makeig, S., Gramann, K., Jung, T. P., Sejnowski, T. J., & Poizner, H. (2009). Linking

brain, mind and behavior. *International Journal of Psychophysiology*, *73*(2), 95–100. https://doi.org/10.1016/j.ijpsycho.2008.11.008

Makeig, S., J. Bell., A., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electroencephalographic Data. *Advances in Neural Information Processing Systems*, *8*(3), 145–151. https://doi.org/10.1109/ICOSP.2002.1180091

Makeig, S., & Onton, J. (2011). *ERP Features and EEG Dynamics*. *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0035

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. Nature reviews. Neuroscience, 5(3), 229.

Mcdowell, K., Lin, C. T., Oie, K. S., Jung, T. P., Gordon, S., Whitaker, K. W., … Hairston, W. D. (2013). Real-world neuroimaging technologies. *IEEE Access*, *1*, 131–149. https://doi.org/10.1109/ACCESS.2013.2260791

Melnik, A., Hairston, W. D., Ferris, D. P., & König, P. (2017). EEG correlates of sensorimotor processing: independent components involved in sensory and motor processing. Scientific Reports, 7.

Melnik, A., Legkov, P., Izdebski, K., Kärcher, S. M., Hairston, W. D., Ferris, D. P., & König, P. (2017). Systems, Subjects, Sessions: To What Extent Do These Factors Influence EEG Data? *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00150

Meyer, D. E., Abrams, R. a, Kornblum, S., Wright, C. E., & Smith, J. E. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review*, *95*(3), 340–370. https://doi.org/10.1037/0033-295X.95.3.340

Milazzo, J., Bagade, P., Banerjee, A., & Gupta, S. K. S. (2013). bHealthy. *Proceedings of the 4th Conference on Wireless Health - WH '13*, 1–2. https://doi.org/10.1145/2534088.2534095

Milham, M. P., & Banich, M. T. (2005). Anterior cingulate cortex: An fMRI analysis of conflict specificity and functional differentiation. *Human Brain Mapping*, *25*(3), 328–335. https://doi.org/10.1002/hbm.20110

Milham, M. P., Banich, M. T., & Barad, V. (2003). Competition for priority in processing increases prefrontal cortex's involvement in top-down control: An event-related fMRI study of the stroop task. *Cognitive Brain Research*, *17*(2), 212–222. https://doi.org/10.1016/S0926-6410(03)00108-3

Milham, M. P., Banich, M. T., Claus, E. D., & Cohen, N. J. (2003). Practice-related effects demonstrate complementary roles of anterior cingulate and prefrontal cortices in attentional control. *NeuroImage*, *18*(2), 483–493. https://doi.org/10.1016/S1053-8119(02)00050-2

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., & Cohen, N. J. (2002). Attentional Control in the Aging Brain: Insights from an fMRI Study of the Stroop Task. *Brain and Cognition*, *49*(3), 277–296. https://doi.org/10.1006/brcg.2001.1501

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our

capacity for processing information. *Psychological Review*, *101*(2), 343–352. https://doi.org/10.1037/h0043158

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning. *arXiv Preprint arXiv: …*, 1–9. https://doi.org/10.1038/nature14236

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. a, Veness, J., Bellemare, M. G., … Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Mohan, S., & Laird, J. E. (2011). An Object-Oriented Approach to Reinforcement Learning in an Action Game. *Proceedings of the Seventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-11)*, 164–169. Retrieved from http://www.shiwali.me/content/mohan_aiide_2011.pdf

Montesano, L., Lopes, M., Bernardino, A., & Santos-Victor, J. (2008). Learning object affordances: From sensory - Motor coordination to imitation. *IEEE Transactions on Robotics*, *24*(1), 15–26. https://doi.org/10.1109/TRO.2007.914848

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Sawamoto, N., … Shibasaki, H. (1999). Transient neural activity in the medial superior frontal gyrus and precuneus time locked with attention shift between object features. *NeuroImage*, *10*(2), 193–199. https://doi.org/10.1006/nimg.1999.0451

Nagel, S. K., Carl, C., Kringe, T., Robert, M., Märtin, R., & König, P. (2005). Beyond sensory substitution—learning the sixth sense. *Journal of Neural Engineering*, *2*(4), R13–R26. https://doi.org/10.1088/1741-2560/2/4/R02

Newell, A. (1994). *Unified theories of cognition*. Harvard University Press.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology*, *46*(3), 461–488. https://doi.org/10.1037/h0084327

O'Regan, J. K., & Noe, a. (2001). What it is like to see: A sensorimotor theory of perceptual experience. *Synthese*, *129*(1), 79–103. https://doi.org/Doi 10.1023/A:1012699224677

Ojeda, A., Bigdely-Shamlo, N., & Makeig, S. (2014). MoBILAB: an open source toolbox for analysis and visualization of mobile brain/body imaging data. *Frontiers in Human Neuroscience*, *8*(March), 121. https://doi.org/10.3389/fnhum.2014.00121

Oliveira, A. S., Schlink, B. R., Hairston, W. D., König, P., & Ferris, D. P. (2016a). Induction and separation of motion artifacts in EEG data using a mobile phantom head device. *Journal of Neural Engineering*, *13*(3), 36014. https://doi.org/10.1088/1741-2560/13/3/036014

Oliveira, A. S., Schlink, B. R., Hairston, W. D., König, P., & Ferris, D. P. (2016b). Proposing Metrics for Benchmarking Novel EEG Technologies Towards Real-World Measurements. *Frontiers in Human Neuroscience*, *10*, 188. https://doi.org/10.3389/fnhum.2016.00188

Onton, J., Delorme, A., & Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *NeuroImage*, *27*(2), 341–356.

https://doi.org/10.1016/j.neuroimage.2005.04.014

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. https://doi.org/10.1016/S1388-2457(00)00527-7

Ord, J. K. (1967). Handbook of the Poisson Distribution. *Journal of the Operational Research Society*. https://doi.org/10.1057/jors.1967.84

Pardo, J. V, Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences*, *87*(1), 256–259. https://doi.org/10.1073/pnas.87.1.256

Paus, T., Petrides, M., Evans, a C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. *Journal of Neurophysiology*, *70*(2), 453–469.

Pavlov, I. P. (1927). *Conditioned Reflexes*. *Oxford University Press* (Vol. 17). https://doi.org/10.2307/1134737

Pelz, J. B. (1995). *Visual representations in a natural visuo-motor task*. University of Rochester.

Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. Vision research, 41(25), 3587-3596.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. *Attention and Performance V*, (JULY 1975), 669–682. Retrieved from papers2://publication/uuid/5E57ECE7-3B3B-4C34-9C4F-CDF5620D8711

Prinz, W. (1990). Relationships Between Perception and Action. https://doi.org/10.1007/978-3-642-75348-0

Prinz, W. (1992). Why don't we perceive our brain states? *European Journal of Cognitive Psychology*, *4*(1), 1–20. https://doi.org/10.1080/09541449208406240

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. https://doi.org/10.1080/713752551

Ravasz, E., & Barabási, A.-L. (2003). Hierarchical organization in complex networks. *Physical Review E*, *67*(2), 26112. https://doi.org/10.1103/PhysRevE.67.026112

Ries, A. J., Touryan, J., Vettel, J., McDowell, K., & Hairston, W. D. (2014). A Comparison of Electroencephalography Signals Acquired from Conventional and Mobile Systems. *Journal of Neuroscience and Neuroengineering*, *3*(1), 10–20. https://doi.org/10.1166/jnsne.2014.1092

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(21), 9475–9489. https://doi.org/10.1016/S0377-2217(02)00363-6

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. *Vision Research*, *51*(12), 1297–1311.

https://doi.org/10.1016/j.visres.2011.04.003

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. Journal of vision, 7(14), 16-16.

Scott A. Huettel; Allen W. Song; Gregory McCarthy, Huettel, Scott; Song, A., Beckmann, C. F., Smith, S. M., Huettel, S. a., Song, A. W., & McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. *Book* (Vol. 23).

Senevirathna, B., Berman, L., Bertoni, N., Pareschi, F., Mangia, M., Rovatti, R., … Abshire, P. (2016). Low cost mobile EEG for characterization of cortical auditory responses. *Proceedings - IEEE International Symposium on Circuits and Systems*, *2016–July*, 1102–1105. https://doi.org/10.1109/ISCAS.2016.7527437

Sermanet, P., Xu, K., & Levine, S. (2017). Unsupervised Perceptual rewards for imitation learning. In *International Conference on Learning Representation* (pp. 1–20).

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*(4), 1916–36. https://doi.org/10.3410/f.1001494.23207

Silton, R. L., Heller, W., Towers, D. N., Engels, A. S., Spielberg, J. M., Edgar, J. C., … Miller, G. A. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior cingulate cortex during top-down attentional control. *NeuroImage*, *50*(3), 1292–1302. https://doi.org/10.1016/j.neuroimage.2009.12.061

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., … Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. https://doi.org/10.1038/nature16961

Sipp, A. R., Gwin, J. T., Makeig, S., & Ferris, D. P. (2013). Loss of balance during balance beam walking elicits a multifocal theta band electrocortical response. *Journal of Neurophysiology*, *110*(9), 2050–60. https://doi.org/10.1152/jn.00744.2012

Snider, J., Plank, M., Lee, D., & Poizner, H. (2013). Simultaneous neural and movement recording in large-scale immersive virtual environments. *IEEE Transactions on Biomedical Circuits and Systems*, *7*(5), 713–721. https://doi.org/10.1109/TBCAS.2012.2236089

Sridharan, M., Wyatt, J., & Dearden, R. (2010). Planning to see: A hierarchical approach to planning visual actions on a robot using POMDPs. *Artificial Intelligence*, *174*(11), 704–725. https://doi.org/10.1016/j.artint.2010.04.022

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*. https://doi.org/10.1037/h0054651

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nature Reviews. Neuroscience*, *6*(5), 363–75. https://doi.org/10.1038/nrn1666

Sun, R. (2003). A tutorial on CLARION 5.0. *Cognitive Science Department, Rensselaer Polytechnic Institute*.

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Ind.Psychiatry J*,

18(0972–6748 (Print)), 70–73. https://doi.org/10.4103/0972-6748.57865

Swick, D., & Jovanovic, J. (2002). Anterior cingulate cortex and the Stroop task: neuropsychological evidence for topographic specificity. *Neuropsychologia*, *40*(8), 1240–1253. https://doi.org/10.1016/S0028-3932(01)00226-3

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(13), 4240–9. https://doi.org/10.1016/j.neuropsychologia.2011.02.038

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(11), 4567–70. https://doi.org/10.1073/pnas.0811155106

Turken, a U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature Neuroscience*, *2*(10), 920–924. https://doi.org/10.1038/13224

Turvey, M. T. (1992). Affordances and Prospective Control: An Outline of the Ontology. *Ecological Psychology*, *4*(3), 173–187. https://doi.org/10.1207/s15326969eco0403_3

VanRullen, R., & Thorpe, S. J. (2001). The Time Course of Visual Processing: From Early Perception to Decision-Making. *Journal of Cognitive Neuroscience*, *13*(4), 454–461. https://doi.org/10.1162/08989290152001880

Vogt, S., Taylor, P., & Hopkins, B. (2003). Visuomotor priming by pictures of hand postures: Perspective matters. *Neuropsychologia*, *41*(8), 941–951. https://doi.org/10.1016/S0028-3932(02)00319-6

Vos, P. G., Mates, J., & van Kruysbergen, N. W. (1995). The Perceptual Centre of a Stimulus as the Cue for Synchronization to a Metronome: Evidence from Asynchronies. *The Quarterly Journal of Experimental Psychology*, *48A*(4), 1024–1040. https://doi.org/10.1080/14640749508401427

Wagner, J., Solis-Escalante, T., Scherer, R., Neuper, C., & Müller-Putz, G. (2014). It's how you get there: walking down a virtual alley activates premotor and parietal areas. *Frontiers in Human Neuroscience*, *8*(February), 93. https://doi.org/10.3389/fnhum.2014.00093

Warchall, J., Balkan, Ö., Mercier, P., Garudadri, H., Hairston, W. D., & Theilmann, P. (2016). A Multi-Channel EEG System Featuring Single- Wire Data Aggregation via FM-FDM Techniques, 526–529.

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. *Machine Learning*, *8*(3–4), 279–292. https://doi.org/10.1007/BF00992698

Weitnauer, E., & Ritter, H. (2012, September). Physical bongard problems. In Ifip international conference on artificial intelligence applications and innovations (pp. 157-163). Springer, Berlin, Heidelberg.

West, R. (2003). Neural correlates of cognitive control and conflict detection in the Stroop and digit-location tasks. *Neuropsychologia*, *41*(8), 1122–1135. https://doi.org/10.1016/S0028-3932(02)00297-X

Workman, J., & Weyer, L. (2007). *Practical Guide to Interpretive Near-Infrared Spectroscopy*. *CRC Press*. https://doi.org/10.1002/anie.200885575

Wright Jr, R. S., Haemel, N., Sellers, G. M., & Lipchak, B. (2010). *OpenGL SuperBible: comprehensive tutorial and reference*. Pearson Education.

Wyczesany, M., & Ligeza, T. S. (2014). Towards a constructionist approach to emotions: verification of the three-dimensional model of affect with EEG-independent component analysis. *Experimental Brain Research*, *233*(3), 723–733. https://doi.org/10.1007/s00221-014-4149-9

Yankov, D., & Keogh, E. (2006). Manifold clustering of shapes. In *Proceedings - IEEE International Conference on Data Mining, ICDM* (pp. 1167–1171). https://doi.org/10.1109/ICDM.2006.101

Yeung, A., Garudadri, H., Van Toen, C., Mercier, P., Balkan, O., Makeig, S., & Virji-Babul, N. (2015). Comparison of foam-based and spring-loaded dry EEG electrodes with wet electrodes in resting and moving conditions. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, *2015–Novem*, 7131–7134. https://doi.org/10.1109/EMBC.2015.7320036

Zenker, F., & Gärdenfors, P. (2015). *Applications of conceptual spaces: The case for geometric knowledge representation*. *Applications Of Conceptual Spaces: The Case For Geometric Knowledge Representation*. https://doi.org/10.1007/978-3-319-15021-5

Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). SUN: A Bayesian framework for saliency using natural statistics. Journal of vision, 8(7), 32-32.

Zhang, F., Leitner, J., Milford, M., Upcroft, B., & Corke, P. (2015). Towards Vision-Based Deep Reinforcement Learning for Robotic Motion Control. *Acra*. Retrieved from http://arxiv.org/abs/1511.03791