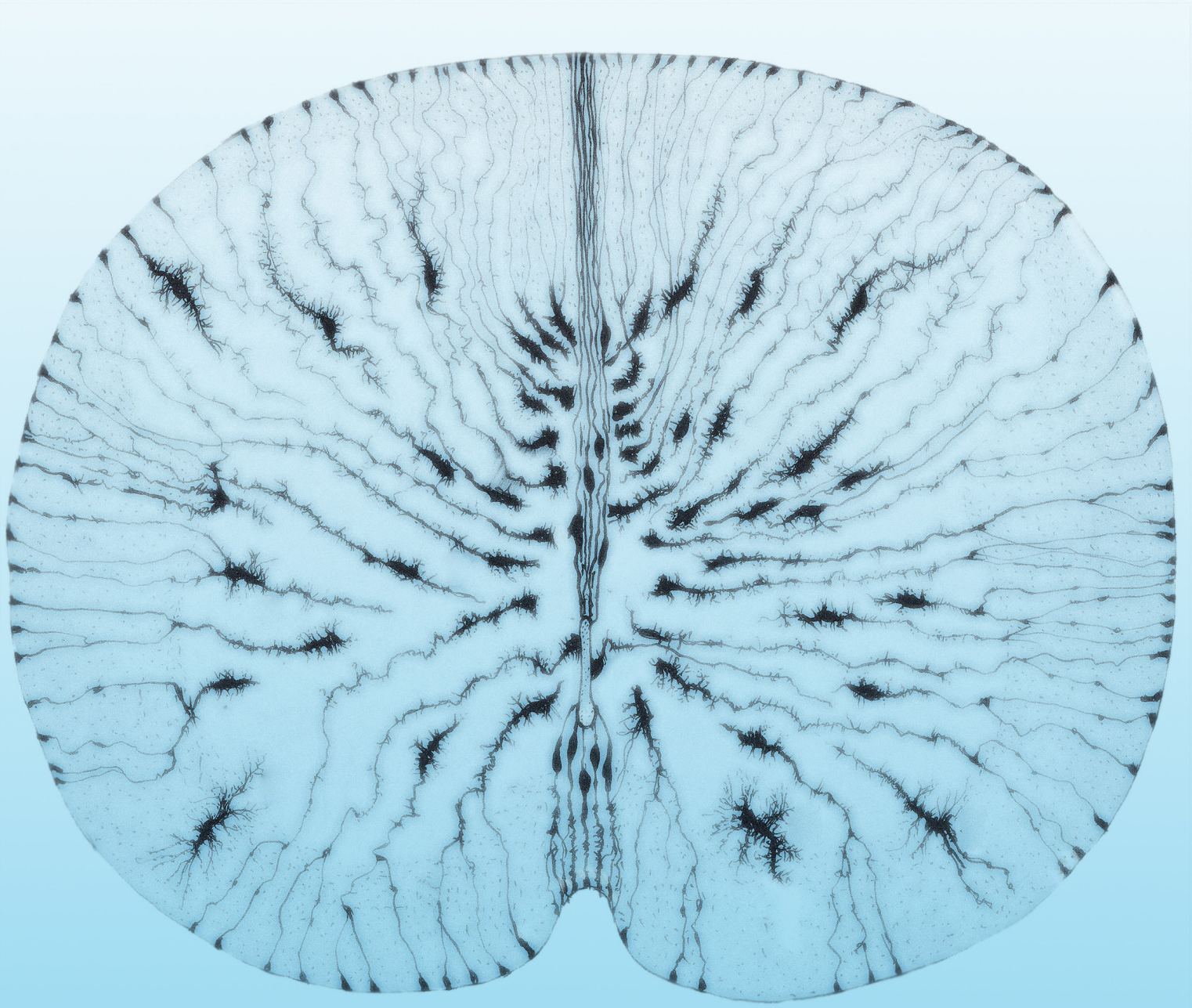
dbc
amsterdam
brain &
cognition

February 2019 | Issue 8

Journal



SEX HORMONES
How we came to understand them and what went wrong

NEUROSCIENTIFIC METHODS
Finding our way through the brain

MODULARITY OF COGNITION
Where do you think?

The ABC Journal committee

Eylül Turan, Steven Voges, Sammy Millard, Sven Wientjes, Iris Bosch, Nikos Kolonis, Marianne de Heer Klots, Rose Nasrawi (Not pictured: Christina Bruckmann, Sylvia Edwards, Feline de Wit)

ABOUT THE COVER IMAGE:
The cover image is a modified, stylized version of one of Santiago Ramon y Cajal's sketches. Ramon y Cajal proposed the neuron doctrine, the hypothesis that the neuron is the functional unit of the brain. His work was instrumental in establishing neuroscience as a distinct academic field. His drawings illustrating different brain cells from different animal brains are still used today for educational purposes.

FROM THE EDITORS

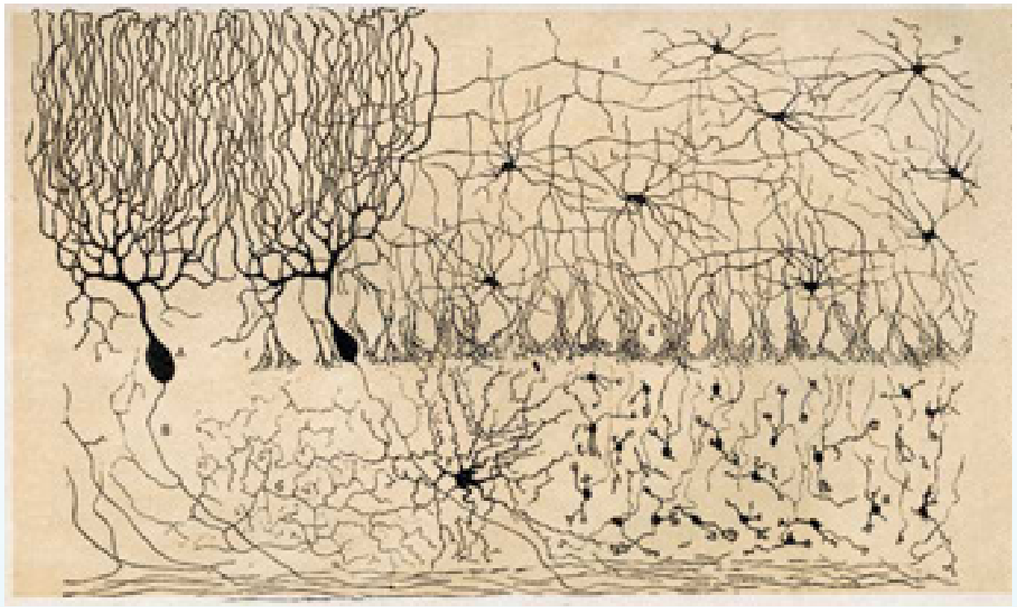
Nowadays in neuroscience, it can sometimes feel impossible to keep up with all the new research being produced. It is estimated that every year more than 2 million scientific papers are published globally, with a growth rate of 8-9% each year (Bornmann & Mutz, 2014; AJE Annual Publishing Review, 2016). As scientists and researchers, it is particularly common for us to get caught up in trying to navigate this deluge of available research in order to stay up to date with the latest and greatest of it, that we never really find the time to take a step back, look at science from a historical perspective and appreciate the progress that has been made.

This is exactly what we are aiming to do with the current issue of the Amsterdam Brain and Cognition Journal. We wanted to take a moment to look back at some of the most important advances and discoveries within the fields of neuro and cognitive science that have shaped our current understanding of the fields. Original pieces from our editors will be delving deeper into some of those topics, including the modularity of cognition (p. 5), the development of sex hormones research (p. 16) and the evolution of neuroscientific methdos (p. 27).

However, this issue of the ABC does not only dwell in the past, it also provides an opportunity for us to look towards the future. In this issue you will find four excellent research articles that were selected among many, from the students of the Brain and Cognitive Sciences Master's at the University of Amsterdam.

Last but not least, with this issue we are welcoming eight new members to the editorial team! We would like to thank them all for their enthusiasm and hard work on making this current issue a reality. Our team is now stronger than ever and we are all working together to make this journal the best it can be.

On behalf of the editorial team,
Sammy Millard and Nikos Kolonis



Ramón y Cajal's drawing of the cells of the chick cerebellum (from 'Estructure de los centros nerviosos de las aves', Madrid, 1905).

To download a digital version of this or any of our previous issues visit:
https://student.uva.nl/bcs/content/az/abc-journal/abc-journal.html

If you are interested in having your research article published in our next issue, please read through the submission guidelines that can be found on the same website.

If you have an interesting idea and would like to contribute with a thought-provoking original piece, follow our Facebook page to get updated on future themes of the Journal or send us an email with your idea at abcjournal.uva@gmail.com.

TABLE OF CONTENTS

MODULARITY OF COGNITION

Where do you think?

by Sven Wientjes

When we think about our human cognition, introspectively it can seem as if it is 'unified'. We do not feel as if we are a collection of special-purpose systems that can interact. Rather we feel like one person, one process that is in control. However, ever since the cognitive revolution lifted the stigma on forming theories about the functioning of cognitive processes, the tendency in cognitive psychology is to describe solely special-purpose systems (miller, 2003). The term that became linked to the 'segregation' of both mental and neural functions is modularity.

The philosopher Jerry Fodor (1983) has put forward influential claims about the modularity of cognition. He claims modularity is expected to be more prevalent at the 'periphery' of the cognitive system (e.g. sensory input processing and motor output processing). How more 'central' parts of cognition are organized, is less obvious. These processes might require such tight coupling of many functions, that segregating them into modules becomes meaningless.

Modularity in cognitive processes could be inferred from behavioural measures. One example of a tour-de-force theory of cognitive processing is the ACT-R cognitive architecture (Anderson, 1996). Here, different processing modules are proposed, which perform separate functions (e.g. visual encoding, memory retrieval, working memory storage, etc.). Every process takes a specified amount of time. Modifying properties of the tasks that are modelled changes the sequence of steps the model goes through, and thus leads to different responses and reaction times. These could be compared to human performance. Recently the ACT-R community has also tried linking their cognitive modules to specific brain regions (although not explicitly claiming these brain regions implement exactly these modules, see Anderson (2007), p. 81-86).

It is widely accepted that there is some type of modularity at the neural level. The brain is in some way organized, as we can for example find areas that relate to sensory processing, while others are for motor output (cf. Fodors peripheral argumentation). How fundamental and how clear this divide is, is what is up for debate. One typical example of modular organization is the distinction between the Fusiform Face Area (FFA) and the Parahippocampal Place Area (PPA), the former encoding information relevant for facial recognition, while the latter encodes specific scenes. The FFA has been called a module for face perception (Kanwisher, McDermott & Chun, 1997), even though activity in this area can increase with expertise in recognizing any arbitrary novel object category (Gauthier et al., 1999). Recently there have even been successful attempts at decoding viewed scenes from only FFA activity, even though the PPA was the proposed 'module' for scene recognition (Johnson & Johnson, 2014).

The idea that distinct brain regions perform distinct functions has come more and more under fire with time. Some even mockingly call it a 'new phrenology' (Uttal, 2001; Anderson, 2014). Interestingly, it was Ramon y Cajal's work on the 'neuron doctrine', which established the neuron as the essential computational device of the brain, that led to a frantic search for the localization of function (Finger, 1994) even though Cajal himself did not support even the definition of a psychological 'faculty' as it was called at the time, let alone its assignment to any localized neural 'organ'. Cajal believed only the relations between many neurons could give rise to functions, essentially dissolving the debate on localizing function.

In fMRI studies that try to localize functions, there is often primacy of the psychological taxonomy of mental function, that has been inherited from cognitive psychology. This is potentially problematic. As Russell Poldrack (2010) describes, if fMRI was invented in the 1860s, it is very likely some of the mental faculties of Phrenology might reliably correlate with the engagement of some brain region. Accepting the general terminology of cognitive psychology could lead to comparable errors.

> "It might have been possible for our (cortical) brains to be organized in a completely distributed manner, just like a typical PDP model. The evidence shows this is not the case."

But the brain is clearly organized by the different sensory modalities, right? This is what is often assumed as 'common knowledge' about the brain, that we have a visual cortex, an auditory cortex, a tactile sensory cortex, and other areas specific to one sense. Pascual-Leone and Hamilton (2001) however, build the case that the brain itself is in principle 'metamodal'. It only appears modal because of functional suppression. One study they cite is Bach-y-Rita and Kercel (2003), where blind participants got a tactile device installed on their lower back that applied pressure correlated to a camera mounted on their front. After the participants got used to this tactile stimulation, they started to report that the sensory information came from 'in front' of them: from the location of the objects instead of from the lower back! Not only somatosensory areas are used for this processing, but also 'visual' areas. This is called 'crossmodal plasticity': the ability of brain regions to adapt and contribute to processing of stimuli outside their 'typical' modality. Brain regions might get input from more than one sense, but during development one form of input comes to 'dominate' the processing of those regions.

By focusing more on how the brain learns as opposed to how it typically operates, considerations of inherent modularity or possible metamodality can be uncovered more clearly. The Parallel Distributed Processing (PDP) or 'connectionism' framework models neural networks that perform cognitive functions (Rumelhart & McClelland, 1986). These networks are often trained to be optimized for certain tasks using algorithms such as backpropagation. Inspecting the organization of typical neural networks trained like this reveals a one-to-many and many-to-one mapping of neural computation to behavior. Each behavior is represented in multiple sites and each site subserves multiple behaviors (Mesulam, 1990). Cognition is 'distributed' over the network as a whole. It makes little sense to divide these networks into regions and find their unique functional contribution.

But obviously there is wide diversity of patterns of activity in the brain. There also seems to be at least in some cases a logical pattern or 'gradient' (e.g. orientation coding for lines in visual cortex). What, if not functional specialization, explains this diversity? Anderson (2014, p. 36) suggests that we should use a different terminology. Instead of functional specialization, think of functional differentiation. Instead of real functions, we should think of functional biases or functional profiles. To make this concrete, Poldrack, Halchenko, and Hanson (2009) gave participants a wide set of typical cognitive psychology tasks while measuring fMRI activity. Using a machine learning approach on all ~214.000 measured voxels they could achieve an accuracy of 90% in labelling which task the participant was engaged in at any point. When trying to understand what information is used by the machine learning algorithm, they applied dimension reduction to ultimately obtain only 6 variables, which performed labelling nearly as accurate as the full 214.000 voxel set. They linked the 6 dimensions, which are summaries of large-scale brain activity, back to the functions that have been reported to correlate with these brain regions in previous studies. The psychological content of these dimensions is then revealed to be quite heterogeneous. The first dimension relates to speech, hearing, working memory and control (inhibition). The second is related to many concepts involved in language, but also quite strongly in response inhibition and spatial processing. The dimensions of brain activity do have some sort of 'personality', but it is unfortunately not so simple that they perform exactly one function.

It might have been possible for our (cortical) brains to be organized in a completely distributed manner, just like a typical PDP model. The evidence shows this is not the case. Network models of brain activity and connectivity show clear modular organization. This type of modularity however, is strictly in the graph-theoretical sense (Sporns, 2010 p. 113). This means a community of nodes (neurons) that share many connections between themselves, but relatively few towards nodes outside the community. Localizationist accounts of brain function based on this type of modularity are considered a failure (Sporns 2010, p. 71-72). Modern network neuroscience focuses on how cognitive function emerges from the brain. The observed modularity stems from evolutionary constraints in their vision. There is degeneracy in brain organization, meaning many different organizations can support similar behavior. Evolution would then find an organization that is optimal with respect to e.g. metabolic cost and volume (the brain needs to fit in the skull after all). This is called the 'wiring minimization hypothesis'. Ruppin, Schwartz and Yeshurun (1993) show that focusing on short-range over long-range connections supports a volume-efficient architecture, while also resulting in a modular organization. Modularity also isolates evolutionary mutations, allowing modules to evolve somewhat independently (Wagner, Pavlicev & Cheverud, 2007). This is important: if the brain was one fully integrated processor and all cognition was widely distributed, mutations would have massive effects and often disrupt the functioning of an organism. Allowing modules to evolve independently improves the selectable phenotypic variation, which is beneficial. This is called the 'evolvability' of an organism (Kirschner and Gerhart, 2005).

To this day, different researchers have very different opinions and intuitions on the degree to which the brain is specialized and in which way its organization can be mapped to our psychology. The answers are not clear, but hopefully this brief overview can provide some guidance through this messy but interesting question.

REFERENCES

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. American Psychologist, 51(4), 355–365. doi:10.1037/0003-066x.51.4.355

Anderson, J.R (2007) How can the human mind occur in the physical universe? Oxford: Oxford University Press.

Anderson, M. L. (2014). After phrenology. Cambridge, MA: MIT Press.

Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human–machine interface. Trends in cognitive sciences, 7(12), 541-546.

Fodor, J.A. (1983). The modularity of mind: An essay on faculty psychology. Cambridge, MA: MIT Press.

Finger, S. (1994). Origins of neuroscience: A history of explorations into brain function. New York: Oxford University Press.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform'face area'increases with expertise in recognizing novel objects. Nature neuroscience, 2(6), 568.

Johnson, M. R., & Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. Frontiers in human neuroscience, 8, 59.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. Journal of neuroscience, 17(11), 4302-4311.

Kirschner, M., & Gerhart, J. (1998). Evolvability. Proceedings of the National Academy of Sciences, 95(15), 8420-8427.

Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. Annals of Neurology, 28(5), 597-613.

Miller, G.A. (2003). The cognitive revolution: A historical perspective. Trends in Cognitive Sciences, 7(3), 141-144.

Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. Progress in brain research, 134, 427-445).

Poldrack, R. A., Halchenko, Y. O., & Hanson, S. J. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. Psychological science, 20(11), 1364-1372.

Poldrack, R. A. (2010). Mapping mental function to brain structure: how can cognitive neuroimaging succeed?. Perspectives on Psychological Science, 5(6), 753-761.

Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition. volume 1. foundations.

Ruppin, E., Schwartz, E. L., & Yeshurun, Y. (1993). Examining the volume efficiency of the cortical architecture in a multi-processor network model. Biological cybernetics, 70(1), 89-94.

Sporns, O. (2010). Networks of the Brain. Cambridge, MA: MIT press.

Uttal, W. R. (2001). The new phrenology: The limits of localizing cognitive processes in the brain. The MIT press.

Wagner, G. P., Pavlicev, M., & Cheverud, J. M. (2007). The road to modularity. Nature Reviews Genetics, 8(12), 921.

Spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations

Lara Engelbert

Supervisor: Mascha Linszen PhD UMC Groningen

University Medical Center Groningen

ABSTRACT

Several authors have stated that there is an association between hearing impairment and complex auditory hallucinations. Such hearing impaired patients must deal with the additional burden of auditory hallucinations in their daily life. The underlying neuronal mechanisms of complex auditory hallucinations in hearing impaired patients remain poorly understood and insights from neuroimaging studies are missing. However, understanding the neuronal mechanisms of complex auditory hallucinations has the potential to improve treatment of auditory hallucinations in this specific patient population in the long- run. Research has indicated that aberrant spontaneous neuronal activity might underlie deafferentation and distorted top-down processes, which may be responsible for complex auditory hallucinations in patients with hearing impairment. We included 50 participants to address this question, consisting of 18 patients with hearing impairment and complex auditory hallucinations (n= 15 tinnitus), 12 patients with hearing impairment without complex auditory hallucinations (n=10 tinnitus) and 20 healthy controls. Resting state fMRI scans were acquired and individual (fractional) amplitude of low frequency fluctuations maps were computed and analyzed to measure spontaneous neuronal activity. Patients with hearing impairment and complex auditory hallucinations showed aberrant spontaneous neuronal activity in the cerebellum, frontal operculum cortex, anterior cingulate gyrus, thalamus, occipital lobe and precuneus as compared to healthy controls. In addition, patients with hearing impairment and complex auditory hallucinations and patients with hearing impairment without complex auditory hallucinations showed overlap in aberrant spontaneous neuronal activity patterns in brain regions such as the cerebellum and the anterior cingulate gyrus. We propose that complex auditory hallucinations may be part of a spectrum on which this phenomenon shares aberrant spontaneous activity patterns with the symptom of tinnitus but is further marked by aberrant top-down processes indicated by aberrant spontaneous neuronal activity in frontal regions.

KEYWORDS

auditory hallucinations, hearing impairment, tinnitus, amplitude of low frequency fluctuations, fractional amplitude of low frequency fluctuations

INTRODUCTION

Hallucinations can occur in any sensory modality, across several neurodegenerative and psychiatric diseases, as well as among healthy individuals (Sommer et al., 2012; Hunter et al., 2006). However, not only neurodegenerative or psychiatric conditions are associated with hallucinations. Studies have indicated an association between sensory deprivation in a specific modality and hallucinations (Teunisse et al., 1996). Especially, several authors have suggested an association between hearing impairment and the experience of auditory hallucinations (Sommer, 2014; Thewissen et al., 2005; Linszen, Brouwer, Heringa & Sommer, 2016). Thus, some hearing impaired patients are confronted with the additional burden of complex auditory hallucinations in their daily lives, i.e. hearing music or voices. However, the underlying neuronal mechanisms of auditory hallucinations among patients with hearing impairment are strikingly understudied in cognitive neuroscience. Clarifying the involved neuronal mechanisms could help to improve our understanding and future treatments for complex auditory hallucinations in this specific patient population.

Approximately 500 million people worldwide are affected by hearing impairment (Stevens et al. 2013). Recently, a cross-sectional study showed that auditory hallucinations occurred in 16.2% of the adult population of patients with hearing impairment, which was significantly more than in the control group (5.8%; Linszen et al., 2018). Some researchers have suggested that deafferentation, i.e. the loss of sensory input in a modality often caused by a distortion in sensory fibers, might underlie complex auditory hallucinations, and that this deafferentation patterns are indicated by spontaneous neuronal activity (Vanneste et al., 2013). Evidence from several resting state fMRI studies illustrated that (fractional) amplitude of low frequency fluctuations ((f)ALFF) within the 0.01-0.1 Hz frequency band are a reliable instrument to measure spontaneous neuronal activity (Alonso-Solís et al., 2017; Chen et al., 2015; Song et al., 2011). Examining spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations by using (f)ALFF measurements might provide us with new insights in aberrant neuronal mechanisms of this specific patient population. Therefore, the present study focused on the examination of spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations.

Auditory hallucinations are often described as perceiving different types of sound, e.g. hearing voices or music (Vanneste et al,2011; Teunisse & Olde Rikkert, 2012; Linszen et al., 2016). Multiple different cognitive mechanisms, such as a change in top-down/bottom-up balance and deafferentation might underlie the experience of auditory hallucinations. Several authors suggested that deafferentation, i.e. reduced auditory input which leads to missing sensory information in brain regions such as the auditory cortex, might underlie the perception of complex auditory hallucinations in patients with hearing impairment (Linszen, Brouwer, Heringa & Sommer, 2016; Sanchez et al., 2011; Braun et al., 2003). In line with this, Linszen and colleagues (2016) suggested that the threshold for neuronal firing within brain regions, which are missing sensory input due to hearing impairment, may eventually decrease. As a result, spontaneous neuronal activity would be more likely to reach the threshold for neuronal firing and eventually cause auditory hallucinations. This is in line with the disinhibition model which assumes that hallucinations are the result of brain activity which emerges due to reduced sensory input (David, 1999). Alongside auditory hallucinations, patients with hearing impairment often experience tinnitus, which is defined as perceiving a sound in the absence of an acoustic source (Kumar et al., 2014; Nam, 2005). Linszen and colleagues (2018) found that 87.5% of patients with hearing impairment and auditory hallucinations, and 77.5% of patients with hearing impairment without complex auditory hallucinations experienced tinnitus. Some authors have suggested that tinnitus could be viewed as a simple auditory hallucination and might be caused by similar deafferentation patterns (Teunisse & Olde Rikkert, 2012). This raises the question if auditory hallucinations are part of a spectrum from simple (e.g. a high frequency tone, tinnitus) to complex (e.g. hearing voices or music) misperceptions. However, research so far is still unclear about the neurobiological basis of auditory hallucinations in patients with hearing impairment and the difference between tinnitus and complex auditory hallucinations. Like more complex auditory hallucinations, tinnitus is strongly associated with deafferentation due to hearing impairment (Hoare et al., 2012). For example, Vanneste and colleagues (2013) performed source-localized EEG on patients with chronic musical hallucinations and patients with tinnitus. The authors found that simple (i.e. perception of tinnitus) and complex (i.e. perception of music) auditory hallucinations shared neurobiological mechanisms, especially similarities in theta-gamma activity in the auditory cortex and beta activity in the dorsal anterior cingulate cortex and anterior insula. In following work, De Ridder and colleagues (2014) stated that filling in missing information to compensate for reduced or missing sensory information activates brain regions such as the anterior cingulate and insula, which are thought to be involved in salience and stimulus detection processes. Their work suggests that auditory hallucinations and tinnitus might share neuronal mechanisms (Vanneste et al., 2013; De Ridder et al., 2014). Moreover, Ghazaleh et al. (2017) used fMRI to assess patients with unilateral hearing loss and tinnitus while stimulating their unaffected ear with sounds. The authors found increased spontaneous and driven neuronal activity in the auditory thalamus. Patients in the present study experienced hearing loss on both ears. However, the same spontaneous neuronal activity in the thalamus might be involved in bilateral hearing loss. For example, Eggermont and Roberts (2012) stated in their review on underlying mechanisms of tinnitus that the thalamocortical input arriving from a damaged ear might be involved in the perception of tinnitus. Especially, laterally disinhibition of the auditory cortex due to thalamocortical dysrhythmia might lead to low frequency fluctuations in the auditory cortex (Ramírez et al., 2009). This could lead to a lowered threshold for spontaneous neuronal firing in regions such as the auditory cortex, subsequently facilitating spontaneous neuronal activity in this region, eventually causing also complex auditory hallucinations in hearing impaired patients. Thus, deafferentation may induce aberrant neuronal activity in the auditory thalamus, which in turn lowers the threshold for spontaneous low frequency oscillations in the auditory cortex. Evidence for spontaneous neuronal firing comes from recent neuroimaging studies which investigated aberrant spontaneous activity patterns within the low frequency band (i.e. 0.01-0.1 Hz) by calculating (fractional) amplitude of low frequency fluctuation (fALFF/ALFF) maps (Alonso-Solís et al. 2017; Chen et al., 2015). These measurements are strongly associated with spontaneous neuronal activity (Song et al., 2011) and have been used in patient populations with tinnitus (Chen et al., 2015) or schizophrenic patients and auditory hallucinations (Alonso-Solís et al., 2017). However, to our knowledge, there is no study yet, which investigated spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations.

Even though complex auditory hallucinations and tinnitus might share neuronal mechanisms associated with deafferentation, this is still a matter of debate and further research is needed (Nam, 2005). Especially, aberrant spontaneous neuronal activity might be region specific for complex auditory hallucinations. Aberrant top-down processes, i.e. the continuous influence of higher cognitive functions on sensory information, may be related to the experience of complex auditory hallucinations in patients with hearing impairment. According to the Bayesian theorem, the brain constantly computes predictions about the environment and updates those based on sensory input to reduce environmental uncertainty (De Ridder, Vanneste

& Freeman, 2014). Thus, a change in bottom-up and top-down balance can strongly influence how sensory input is experienced. Higher cognitive processes might create and add elements to existing external stimuli (Mason & Brady, 2009), potentially leading to complex auditory misperceptions (De Ridder et al., 2014). This is in line with Powers, Kelley and Corlett (2016) who stated that hallucinations can be understood as the product of top-down effects on perception. Filling in information to compensate for missing sensory input would reduce uncertainty about the external stimuli (Shahin et al., 2009). Thus, spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations might not be restricted to the auditory cortex, caused by deafferentation, but could extend to frontal regions. Aberrant low frequency fluctuations in the frontal cortex, indicating spontaneous neuronal activity, might imply distorted top-down processes, which add or create auditory elements to compensate for reduced sensory input within brain regions such as the thalamus and the auditory cortex.

The present study

Evidence from neuroimaging studies regarding spontaneous neuronal activity, potentially underlying deafferentation or top-down processes in patients with hearing impairment and complex auditory hallucinations, is missing. Therefore, the present study used resting state fMRI data to examine ALFF/fALFF values as an indication of spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations (HI-H; n=15 experienced tinnitus), patients with hearing impairment without complex auditory hallucinations (HI; n=10 experienced tinnitus), and healthy controls (HC). To our knowledge, research on hallucinations did not examined spontaneous neuronal activity in patients with hearing impairment and complex auditory hallucinations so far. Therefore, the present research was primarily exploratory, and we were mainly interested in *how spontaneous brain activity in the HI-H group differs from spontaneous brain activity in the HC group*. Furthermore, by including the HI group, we had the opportunity to investigate *how spontaneous brain activity in the HI-H group differs from spontaneous brain activity in the HI group*.

Several authors have suggested that complex auditory hallucinations can be understood as the product of deafferentation (Linszen, Brouwer, Heringa & Sommer, 2016; Sanchez et al., 2011; Braun et al., 2003). Deafferentation is thought to cause aberrant spontaneous neuronal activity which can be measured with (f)ALFF values. Brain regions such as the auditory cortex or the thalamus are directly influenced by hearing impairment, i.e. less sensory information is send to these regions. Therefore, we expected to find significantly more spontaneous neuronal activity in the auditory cortex and thalamus in hearing impaired patients with complex auditory hallucinations as compared to healthy controls. In addition, some researchers stated that complex auditory hallucinations could be the result of top-down processes which add or create new elements to eventually reduce uncertainty about distorted external stimuli (Mason & Brady, 2009; De Ridder et al., 2014; Powers et al., 2016). These top-down processes might be reflected in aberrant spontaneous neuronal activity in frontal regions. Therefore, we expected to find significantly more aberrant spontaneous activity in frontal regions in the hearing impaired patient group with complex auditory hallucinations as compared to healthy controls.

Previous research has raised the question if tinnitus and complex auditory hallucinations share underlying neuronal mechanisms (Nam et al., 2005). Deafferenation is considered as a potentially underlying mechanisms of both complex auditory hallucinations (Linszen, Brouwer, Heringa & Sommer, 2016; Sanchez et al., 2011; Braun et al., 2003) and tinnitus (Teunisse & Olde Rikkert, 2012). Additionally, research has shown that patients with tinnitus as well as patients with complex auditory hallucinations show similar brain activity patterns in the auditory cortex (Vanneste et al., 2013). Therefore, we expected to find an overlap in increased (f)ALFF values in the auditory cortex between patients with hearing impairment and complex auditory hallucinations and patients with hearing impairment without complex auditory hallucinations.

METHOD

Design

We used an observational, between-subjects design to examine spontaneous neuronal activity in hearing impaired patients with complex auditory hallucinations, hearing impaired patients without complex auditory hallucinations and healthy controls. Resting state fMRI data of both hearing impaired patient groups was collected as part of the study “Understanding hallucinations II fMRI and EEG”. Scans of healthy controls were derived from the “Spectrum” study. Both studies were approved by the Local Research Ethics Committee from the University Medical Center Utrecht. All participants gave informed consent before their participation.

Participants

Twenty-five patients with hearing impairment and complex auditory hallucinations (HI-H) and 16 patients with hearing impairment without complex auditory hallucinations (HI) were recruited from the audiological centre at the University Medical Center Utrecht. Derived from recent clinical tone audiometric measures, the High Fletcher Index (HFI, mean hearing loss in dB for tones on 1,2 and 4 kHz) served as an indication for hearing loss. A value of 125 dB was assigned in case of complete deafness. Patients underwent a semi-structured interview, used in previous studies (Teunisse & Olde Rikkert, 2012), consisting out of 14 items on tinnitus and spontaneous acoustical phenomena to identify auditory hallucinations and distinguish them from tinnitus, imagery and illusions. Patients in the HI-H group had to have experienced an auditory hallucination at least once within the past month and a HFI ≥ 25dB in the worst ear. Patients in the HI group had no auditory hallucinations within the last two years (or not more than one episode of an hallucination, longer than two years ago) and a HFI ≥ 25dB in the best ear. Twenty healthy control participants (HC) were recruited via the website www.verkenuwgeest.nl (“explore your mind”). Participants in all three groups were older than 18, spoke Dutch on a sufficient level and were mentally competent. Seven patients from the HI-H group, three patients from the HI group and one patient from the HC group were excluded because of insufficient scan quality. Another patient from the HI group was excluded because the patient did not meet the inclusion criteria any longer. Especially, the patient was neither a case, i.e. had complex auditory hallucinations at least once in the last four weeks, nor a control, i.e. the patient had not more than one episode of a hallucination within the last two years. The three groups were matched on age, and the proportions of gender and handedness did not differ significantly from each other between the groups (table 1).

MRI acquisition

Blood oxygenation level-dependent sensitive resting state fMRI scans of eight minutes duration were acquired with a Philips Achieva 3.0 Tesla scanner (Philips Medical Systems, Best, The Netherlands) at the University Medical Center Utrecht (3D-PRESTO pulse sequence with parallel imaging (SENSE) in two directions, using commercial 8-channel SENSE headcoil, full brain coverage within 609 ms, TR/TE = 21.75/32.4 ms, field of view (FOV) 224 mm x 256 mm x 160 mm, matrix = 64 mm x 64 mm x 40 mm, number of slices (coronal)= 40, 4 x 4 x 4 mm voxels, flip angle =10°). A total of 600 volumes were scanned and used for data analysis. For anatomical reference, high resolution T1-weighted images (TR= 10 ms, TE = 4.6 ms, FOV =240 mm/100%, voxel size = 0.75 mm x 0.75 mm x 0.80 mm, reconstruction matrix = 200 x 320 x 320, flip angle = 90°) were acquired. Participants were asked to lie as still as possible in the scanner, with their eyes closed, and to stay awake during the scan.

Functional data processing

Functional neuroimaging data was processed using FMRIB’s Software Library (FSL) version 5.0.4. The processing pipeline was carried out using FEAT and consisted of non-brain removal using BET, motion correction using MCFLIRT, and spatial smoothing (5 mm full-width at half maximum (FWHM) Gaussian kernel). Registration to standard space was done using FLIRT (2mm Montreal Neurological Institute (MNI) standard space) and refined using FNIRT non-linear registration. Grand mean intensity normalization of the entire dataset was done with a single multiplicative factor. Linear detrending and band pass filtering (0.01-0.08 Hz) was applied using the “RESTing-state fMRI data analysis toolkit” (created by Song Xiaowei, http://resting-fmri.sourceforge.net).

ALFF and fALFF calculation

Research has indicated that amplitude of low frequency fluctuation (ALFF) and fractional amplitude of low frequency fluctuation (fALFF) values are reliable instruments to investigate spontaneous neuronal activity (Song et al., 2011). Whereas ALFF values reflect the total power within the low frequency range (0.01-0.1 Hz), i.e. the intensity of low frequency oscillations, fALFF values are the total power within the low frequency range, divided by the power detectable in the entire frequency range (Song et al., 2011). Therefore, fALFF values reflect the relative contribution of low frequency oscillations to the entire frequency domain. fALFF measurements have an increased specificity and sensitivity concerning the detection of spontaneous neuronal activity in grey matter (Song et al., 2011; Zou et al., 2008). On the other hand, ALFF measurements are known for having an improved test-retest reliability as compared to fALFF measurements (Child Mind Institute, https://fcp-indi.github.io/docs/user/alff.html). Therefore, both measurements have their benefits and reflect spontaneous neuronal activity on different scales, i.e. only within the low frequency (ALFF) or in comparison with the entire frequency domain (fALFF). (f)ALFF values were computed using REST. First, time courses were transformed to the frequency domain using a Fast Fourier Transform (FFT). The square root of the power spectrum was computed and averaged across the 0.01-0.08 Hz interval (Alonso-Solís et al., 2017; Yu et al., 2014). The average square root was divided by the individual global mean ALFF value to reduce effects of variability among subjects. The resulting ALFF maps (per subject) were used for further statistical group analysis. fALFF was computed as the ratio of the power spectrum of the low frequency interval 0.01-0.08 Hz to that of the entire frequency band. Therefore, bandpass filtering was not applied before the calculation of fALFF values (Song et al., 2011).

Statistical analysis

Demographic data of the participants was analysed using X2 – test for proportions (gender, handedness, tinnitus), a one-way analysis of variance (ANOVA) between subjects for means (age) ($p < 0.05$) and an independent t-test for differences in HFI between HI-H and HI group using SPSS 22 software (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.). A non- parametric group-level between subjects (group as between-subjects factor) whole brain analysis was performed to examine the effect of group on ALFF/fALFF values using FSL’s randomise command in combination with Threshold-Free Cluster Enhancement (TFCE, Family Wise Error (FWE)-corrected at $p < 0.05$). In addition, a conjunction analysis for the contrasts HI-H>HC and HI>HC as well as HI-H<HC and HI<HC with a cluster-wise threshold of $z>2.6$ was performed to examine a possible overlap in spontaneous neuronal activity patterns between the HI-H and the HI group.

RESULTS

Clinical characteristics

Clinical and demographical characteristics of the three groups are presented in table 1. The proportions of gender and handedness were equal in all three groups and participants were matched on age. Furthermore, the participants in both hearing impaired patient groups were matched on hearing loss (HFI_best and HFI_worst) and the proportions of patients who experienced tinnitus were equal in both groups.

Table 2 shows the categorized content of the complex auditory hallucinations of participants from the HI- H group. Most of the patients experienced musical hallucinations or non-verbal sounds such as the sound of a door bell or a telephone.

Table 1. Demographics

	Group			
	Healthy controls	Hearing impairment	Hearing impairment + Complex auditory hallucinations	Statistical comparison
Number Participants	N=20	N=12	N=18	
Gender	11 females	7 females	14 females	n.s[a]
Age	M=54 SD=7.3	M=61 SD=12.3	M=56 SD=12.2	n.s[b]
Handedness	17 right handed	5 right handed	14 right handed	n.s[a]
HFI_best	-	M=49.4 SD=21.7	M=40.6 SD=21.5	n.sc
HFI_worst	-	M=59.3 SD=21	M=59.5 SD=24.8	n.sc
Tinnitus	-	N=10	N=15	n.s.a

Notes. n.s. no significant effect; [a]Comparison between groups using χ2-test; [b]Comparison between groups using one-way ANOVA, [c]Comparison between groups using independent t-test, unpaired, 2-tailed.

ALFF Analysis

First, the effect of group on ALFF values was examined. Brain regions which showed significantly larger ALFF values in the HI-H group as compared to the HC group included the cerebellum (bilateral), left temporal pole, left anterior cingulate gyrus, left parahippocampal gyrus, frontal orbital cortex and left inferior temporal gyrus (table 3, figure 1). A significant increase in ALFF values in the HC group as compared to the HI-H group were found in the left and right occipital lobe, the putamen, the left temporal occipital fusiform cortex, the left precuneous cortex and the right lingual gyrus (table 4, figure 1).

Significantly higher ALFF values in the HI as compared to the HC group were found in the cerebellum (bilateral) and in the left temporal pole (appendix A, figure 2). The left cuneal cortex, the left insula, the right lingual gyrus, the left temporal occipital fusiform cortex,

Table 2. Categorized content of auditory hallucinations in HI-H group

	Categorization		
	Music	Verbal	Non-verbal
Music (radio)	X		
Music, Murmuring crowd	X		X
Radio, voice of famous news moderator	X	X	
Music, children shouting name, door bell	X	X	X
Calling name		X	
Music	X		
Fly			X
Airplanes, storm, Murmuring crowd			X
Sword fight. Music	X		X
Melody, instrumental	X		
Music, Murmuring crowd	X		X
Flight of Birds			X
Music, helicopter	X		X
Music	X		
Music	X		
Murmuring crowd			X
Noise, tap			X
Voices (cannot understand), door bell telephone			X

Notes. The categorization of auditory hallucinations into musical, verbal and non-verbal auditory hallucinations is based on work by Blom and Sommer (2010). The categories are based on the content of the auditory hallucinations. We decided to categorize the hearing of many voices, without being able to understand what those voices are saying, as non-verbal hallucinations.

Table 3. Group level contrast of HI-H > HC for ALFF

Peak MNI coordinates					
Cluster size	Peak t-value	x	y	z	Brain region
490	6.29	-36	-42	-46	Left cerebellum
366	6.53	24	-40	-54	Right cerebellum
211	4.92	-14	28	14	Anterior left cerebral white matter
97	4.31	-10	-16	-40	Brainstem
44	4.16	4	18	22	Anterior left cingulate gyrus
10	4.22	-36	6	-48	Left temporal pole
9	4.14	56	-44	-28	Right inferior temporal gyrus
8	3.89	-52	-60	-38	Left cerebellum
7	3.84	-28	-26	-30	Left parahippocampal gyrus
6	3.83	-16	24	-12	Frontal orbital cortex
6	3.91	-50	-44	-22	Left inferior temporal gyrus
6	3.91	12	-22	-40	Brainstem

Notes. Table shows significant clusters (p<0.05, FWE corrected) of whole brain analysis.

the left putamen and the left precentral gyrus showed higher ALFF values in the HC group as compared to the HI group (appendix A, figure 2). We did not find a significant difference in ALFF values between HI-H and HI group.

fALFF Analysis

Brain regions which showed significant higher fALFF values for the HI-H group as compared to the HC group included the left cerebellum, the frontal operculum cortex, the right and left thalamus and

Table 4. Group level contrast of HI-H < HC for ALFF

Peak MNI coordinates					
Cluster size	Peak t-value	x	y	z	Brain region
1505	5.72	-10	-78	8	Left Occipital Lobe
240	7.31	-32	-28	2	Left cerebral white matter/Putamen
53	4.62	-36	-52	-18	Left Temporal Occipital Fusiform Cortex
29	3.83	-24	-50	48	Left Superior Parietal Cortex
16	3.83	-26	-20	52	Left Precentral Gyrus
13	4.08	20	-74	26	Right Cuneal Cortex
13	3.93	-12	-52	8	Left Precuneus Cortex
10	3.99	18	-66	-2	Right Lingual Gyrus
9	3.72	-24	-34	56	Left Postcentral Gyrus
9	3.98	16	-78	16	Right Occipital Lobe

Notes. Table shows significant clusters (p<0.05, FWE corrected) of whole brain analysis

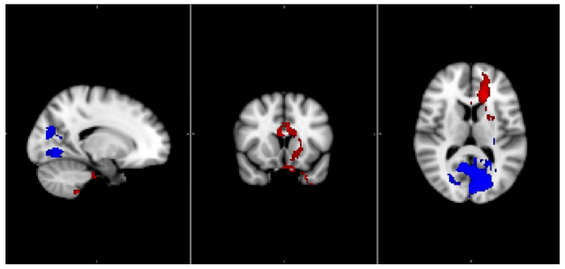


Figure 1. **Regions associated with increased/decreased ALFF values in the HI-H group as compared to HC group.** Red colored regions showed a significant increase in ALFF values in the HI-H group. Blue colored regions showed a significant decrease in ALFF values in the HI-H group. The shown activations are thresholded t-stat images from the non-parametric TFCE- based test (p < 0.05, FWE corrected).

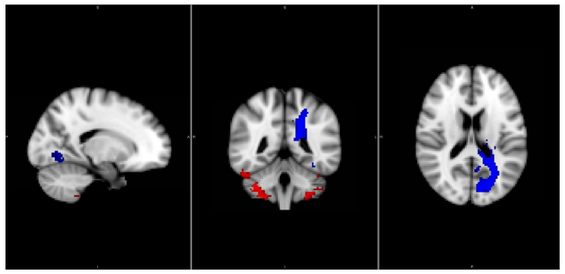


Figure 2. **Regions associated with increased/decreased ALFF values in the HI group as compared to HC group.** Red colored regions showed a significant increase in ALFF values in the HI group. Blue colored regions showed a significant decrease in ALFF values in the HI group. The shown activations are thresholded t-stat images from the non-parametric TFCE- based test (p < 0.05, FWE corrected).

the left posterior temporal cortex (table 5, figure 3). No significantly higher fALFF values were found for the HC group as compared to the HI-H group.1).

The left superior parietal lobule, the right lateral occipital cortex,

Table 5. Group level contrast of HI-H > HC for fALFF

Peak MNI coordinates					
Cluster size	Peak t-value	x	y	z	Brain region
265	5.08	-10	-80	-28	Left Cerebellum
249	5.56	36	26	6	Frontal Operculum Cortex
148	6.51	16	-44	32	Right Posterior Cingulate Gyrus
84	4.98	-20	-70	-48	Left Cerebellum
70	4.57	-20	-14	8	Right Thalamus
54	4.13	32	-58	28	Right Superior Lateral Occipital Cortex
46	4.85	2	-54	-38	Cerebellum
44	4.48	-16	-36	40	Left posterior Cingulate Gyrus
41	4.44	-18	-16	8	Left Thalamus
36	4.75	38	-10	32	Right Precentral Gyrus
29	4.18	-12	-46	24	Left Posterior Cingulate Gyrus
29	5.21	-30	-40	-16	Left Posterior Temporal Fusiform Cortex
25	4.28	-34	-74	16	Left Lateral Occipital Cortex

Notes. Table shows significant clusters (p<0.05, FWE corrected) of whole brain analysis.

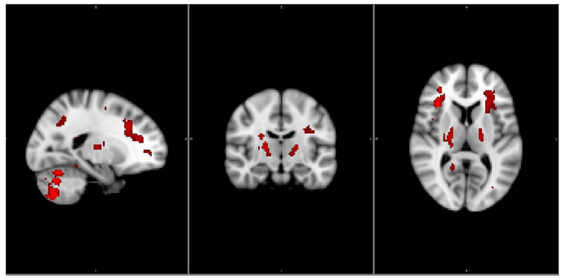


Figure 3. **Regions associated with increased fALFF values in the HI-H group as compared to HC group.** Red colored regions showed a significant increase in fALFF values in the HI-H group. The shown activations are thresholded t-stat images from the non-parametric TFCE-based test (p < 0.05, FWE corrected).

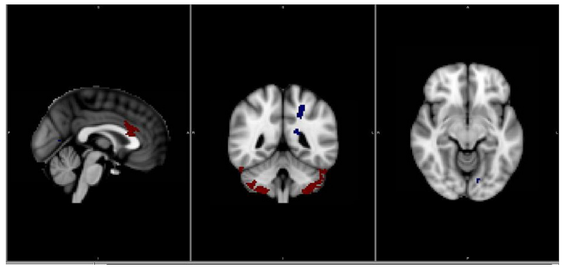


Figure 4. **Conjunction analysis for ALFF values of (HI-H > HC) ∩ (HI > HC) and (HI-H < HC) ∩ (HI < HC).** Red colored regions showed overlap in cluster activation for an increase in ALFF values for the two contrasts HI-H > HC and HI > HC. Blue colored regions showed overlap in cluster activation for a decrease in ALFF values for the two contrasts HI-H < HC and HI < HC. The shown activations are thresholded z-stat images from the conjunction analysis in FSL (z > 2.6).

the right frontal pole and the right lingual gyrus showed, among others, higher fALFF values for the HI group as compared to the HC group (appendix B). Significant higher fALFF values for the HC group as compared to the HI group were found in the left cerebellum (appendix B). We did not find a significant difference in fALFF values between HI-H and HI group.

Table 6. Conjunction analysis ALFF

Peak MNI coordinates					
Cluster size	Peak z score	x	y	z	Brain region
(HI-H > HC) ∩ (HI > HC)					
229	5.1	30	-42	-54	Right Cerebellum
162	5.2	-28	-42	-52	Left Cerebellum
7	3.83	4	18	22	Right Anterior Cingulate Gyrus
(HI-H < HC) ∩ (HI < HC)					
266	4.94	-18	-76	24	Left Cuneal Cortex
116	4.61	-12	-72	0	Left Lingual Gyrus
19	3.86	-18	-50	20	Left Posterior Cingulate Gyrus
10	3.71	-20	-58	28	Left Superior Precuneous Cortex

Notes. Table shows significant clusters (p<0.05, FWE corrected) of whole brain analysis.

Conjunction Analysis

The conjunction analysis revealed significantly larger ALFF values in both groups in the right and left cerebellum, and the right anterior cingulate gyrus (table 6, figure 4). Regions which showed significantly larger ALFF values in the HC group as compared to the HI-H and HI group included the cuneal cortex, the left lingual gyrus and the left posterior cingulate gyrus (table 6, figure 4). We did not find a significant overlap in fALFF values between HI-H and HI group.

DISCUSSION

Our results demonstrate that aberrant spontaneous neuronal activity in various regions differentiate patients with hearing impairment and complex auditory hallucinations from healthy controls. For example, we found an increase in spontaneous neuronal activity in frontal regions such as the frontal operculum cortex in patients with hearing impairment and complex auditory hallucinations. These spontaneous brain activity patterns might indicate an involvement of top-down cognitive processes in the perception of complex auditory hallucinations. Furthermore, our findings indicate that there is no significant difference in spontaneous neuronal activity between patients with hearing impairment and complex auditory hallucinations, and patients with hearing impairment without complex auditory hallucinations. Surprisingly, we found aberrant spontaneous activity for both patient groups in the bilateral cerebellum and right anterior cingulate gyrus but not in the auditory cortex.

The first aim of the present study was to determine how spontaneous neuronal activity in the HI- H group differs from spontaneous neuronal activity in the HC group. Our results indicate that patients who suffer from complex auditory hallucinations had significantly more spontaneous neuronal activity in regions such as the bilateral cerebellum, the left temporal pole and the anterior cingulate gyrus as compared to healthy controls. Furthermore, findings demonstrate a significant decrease in spontaneous neuronal activity in regions including the right lingual gyrus, the left temporal occipital fusiform cortex, left putamen, precentral gyrus, bilateral occipital lobe and left superior parietal cortex in the HI-H as compared to the HC group. In addition, patients with hearing impairment and complex auditory hallucinations showed increased fALFF values in the cerebellum, which is in line with the ALFF analysis. Notably, fALFF values also revealed an increase spontaneous neuronal activity in the right frontal operculum cortex, the thalamus (bilateral) and the left posterior temporal cortex in hearing impaired patients with complex auditory hallucinations as compared to healthy controls. Therefore, these regions probably significantly contribute to the overall spontaneous neuronal activity in the entire frequency domain of these patients. However, no significant decreases of fALFF values in the HI-H group as compared to the HC group were found. Thus, our results indicate that there is no decreased spontaneous neuronal activity in the low frequency domain in the hearing impaired patient group with complex auditory hallucinations, as compared to healthy controls, that significantly influenced the overall spontaneous brain activity in the whole frequency domain.

ALFF/fALFF results indicate spontaneous neuronal activity in the cerebellum in the HI-H group. To our knowledge, this activation pattern was not found in previous examinations of spontaneous neuronal activity of patients with complex auditory hallucinations. However, in line with our findings for the HI group, Chen and colleagues (2015) found larger ALFF values in the cerebellum of patients with tinnitus. Therefore, the spontaneous activity we found in the cerebellum in the hearing impaired patient group with complex auditory hallucinations might rather underlie the perception of tinnitus than complex auditory hallucinations since the majority of patients experienced tinnitus as well (n=15 tinnitus). Moreover, research has indicated that structural changes in the cerebellum might underlie auditory hallucinations in patients with schizophrenia (Cierpka et al., 2017). However, Chen et al. (2015) did not find significant differences in grey matter volume of tinnitus patients as compared to healthy controls. There is no research, yet, on structural changes in e.g. the cerebellum of patients with hearing impairment and complex auditory hallucinations and if these might influence aberrant spontaneous neuronal activity remains unknown. Thus, future studies should examine if there are structural changes in the cerebellum of patients with hearing impairment and complex auditory hallucinations and if these are correlated with an increase of ALFF/fALFF values within this brain region.

Furthermore, our findings demonstrate increased ALFF values in the anterior left cingulate gyrus, and increased fALFF values in the left posterior temporal fusiform cortex in the HI-H group. Bonilha and colleagues (2017) suggested that these regions might serve as a focal point for integration of auditory and conceptual processing. Moreover, research has indicated that activation in the anterior cingulate is associated with salience and stimulus detection processes (De Ridder et al., 2014). Thus, the increased spontaneous neuronal activity we found within these regions might indicate that patients with hearing impairment and complex auditory hallucinations engage in a top-down process in which missing or reduced sensory auditory input is detected and higher cognitive processes construct additional elements eventually leading to auditory hallucinations (De Ridder et al.,2014). Increased fALFF values in the right frontal operculum cortex further indicate the involvement of top-down mechanism in the HI-H group. For example, Eggermont and Roberts (2012) stated that sensory input which arrives from a damaged ear might engage frontal networks to create a more accurate auditory perception. Furthermore, the results of the fALFF analysis showed spontaneous neuronal activity in the thalamus (bilateral) in the HI-H group. Therefore, aberrant thalamocortical rhythms due to a distortion of incoming sensory input, might trigger spontaneous neuronal activity in frontal networks such as the frontal operculum cortex. This is in line with Powers and colleagues (2016) who stated that uncertainty evoked by distorted sensory input is compensated by the engagement of frontal networks, aimed at constructing more accurate auditory perceptions.

The aberrant spontaneous activity we found in the thalamus in the HI-H group, is in line with previous research in which the authors found larger fALFF values in the thalamus of schizophrenic patients with consistent auditory hallucinations (Alonso-Solís et al., 2017). Even if Alonso-Solís and colleagues (2017) examined fALFF values in a different patient group, the increased fALFF values in the thalamus in the HI-H group might indicate that the thalamus plays a crucial role for the emergence of complex auditory hallucinations among various diagnostic groups. However, this assumption remains hypothetical and future studies which investigate spontaneous neuronal activity patterns across different diagnostic groups are needed.

In line with previous research on tinnitus and schizophrenic hallucinating patients, we found decreased ALFF values in the HI-H group in the bilateral occipital lobe (Chen et al., 2015; Alonso-Solís et al., 2017), the left precuneus cortex and the right lingual gyrus (Alonso-Solís et al., 2017; Hare et al., 2017). In combination with the increased fALFF/ALFF values found in the anterior and posterior cingulate gyrus, these results indicate aberrant spontaneous neuronal activity in the default network in patients with hearing impairment and complex auditory hallucinations. However, it remains unclear if this aberrant spontaneous neuronal activity is solely attributable to the perception of auditory hallucinations or rather reflects spontaneous activity related to hearing loss or tinnitus. For example, Yang and colleagues (2014) found decreased ALFF values in patients with unilateral hearing loss in the bilateral precuneus.

The second aim of the present study was to determine how spontaneous neuronal activity in the HI-H group differs from spontaneous neuronal activity in the HI group. Our results show that there were no significant differences in ALFF/fALFF values between patients with hearing impairment and complex auditory hallucinations and patients with hearing impairment without complex auditory hallucinations. Moreover, our results from the conjunction analysis indicate that both groups share patterns of spontaneous neuronal activity in the bilateral cerebellum and the right anterior cingulate gyrus as compared to healthy controls. Moreover, HI-H and HI group showed decreased ALFF values in the left cuneal cortex, the left lingual gyrus and the left posterior cingulate gyrus. The proportions of patients who experienced tinnitus were equal in both patient groups. Therefore, the two groups might share spontaneous activity patterns because they were equally affected by the perception of tinnitus. In addition, aberrant spontaneous activity patterns might have reflected tinnitus only, and not complex auditory hallucinations in the HI-H group. However, ALFF values extended to frontal regions such as the frontal orbital cortex in the HI-H group. There was also a decrease in ALFF values which extended to regions such as the bilateral occipital lobe and the left superior parietal cortex in the HI-H group as compared to the HI group. Even though these spontaneous neuronal activity patterns did not significantly differ with the HI group, these results indicate that aberrant spontaneous activity patterns in the HI-H group extended to different regions. This indicates that other cognitive processes might be distorted in the HI-H as compared to the HI group. (see figure 1 and figure 2). Moreover, results showed a significant increase of ALFF values in the left temporal pole in the HI-H and HI group (table 3, appendix A). This is in line with previous work by Alonso-Solís and colleagues (2017) who found significant larger ALFF values in schizophrenic patients with auditory hallucinations in this brain area. Thus, although participants in both patient groups in the present study were equally affected by the perception of tinnitus, we found spontaneous brain activity patterns which are comparable to other patient groups with auditory hallucinations. Therefore, our results, in line with previous work (Alonso-Solís et al., 2017), indicate that tinnitus and complex auditory hallucinations might share spontaneous activity patterns, for example in the left temporal pole.

In the end, we propose that aberrant spontaneous neuronal activity is associated with the experience of complex auditory hallucinations. We suggest a spectrum model on which both, tinnitus and auditory hallucinations, are marked by aberrant spontaneous neuronal activity patterns in certain brain areas such as the anterior cingulate gyrus and the cerebellum. On this spectrum, aberrant spontaneous activity appears to extend to subcortical regions such as the thalamus and cortical regions such as the frontal operculum and frontal orbital cortex in hearing impaired patients with complex auditory hallucinations as compared to healthy controls.

Limitations

The composition of the three different groups reveals some limitations. First, healthy controls were expected to hear significantly better than the participants in the two patient groups, but only the exact data of hearing loss of the participants in the HI-H and the HI group was available. However, the question raises if the differences we found between HI-H and the HC group are the result of hearing loss only, and thus not reflect spontaneous neuronal activity related to complex auditory hallucinations. Second, not all patients in the HI-H and HI group experienced tinnitus (15 patients in the HI-H group; 10 patients in the HI group, table 1). Therefore, participants without tinnitus in the two patient groups might have influenced the results. Furthermore, there was no data available concerning tinnitus duration, i.e. if the participants continuously perceived a tinnitus. In addition, we were not able to check for differences between educational levels between all three groups because the educational level of the healthy controls was unknown. Therefore, we could not exclude the possibility that these factors differed among the participants and influenced the results. Finally, the participant numbers between the three different groups were not equal, due to e.g. exclusions based on scan quality or inclusion criteria. Thus, there was a power difference between the three groups which might have influenced the results. The optimal solution for these limitations would be a four groups design in which all participants are affected by an equal amount of hearing loss. The four groups should contain a group with complex auditory hallucinations and tinnitus, complex auditory hallucinations without tinnitus, no complex auditory hallucinations and tinnitus and no complex auditory hallucinations and no tinnitus. However, it remains difficult to find a group of patients that is hearing impaired (with and without complex auditory hallucinations) but does not perceive tinnitus, as indicated by recent work in which 77.5% of patients with hearing impairment without auditory hallucinations and 87.5% of patients with hearing impairment and complex auditory hallucinations also perceived tinnitus (Linszen et

al., 2018). In addition, 25 participants are considered to be an adequate group size for a resting state fMRI study (Chen et al., 2017) and future studies should uphold these standards. Moreover, additional information about tinnitus duration and education should be gathered to exclude a possible influence of these factors on spontaneous neuronal activity.

Conclusion

We identified aberrations in low frequency fluctuations in frontal, posterior and subcortical regions in patients with hearing impairment and complex auditory hallucinations. In addition, we found an overlap in aberrant neuronal activity patterns in the cerebellum and the anterior cingulate gyrus as well as posterior regions in patients with hearing impairment and complex auditory hallucinations and patients with hearing impairment without complex auditory hallucinations. Our findings suggest that aberrant low frequency fluctuations in these regions might be an underlying cause of both complex auditory hallucinations and tinnitus in patients with hearing impairment. However, future research is needed to clarify if aberrant spontaneous neuronal activity in patients with hearing impairment reflects hearing loss itself, tinnitus or underlying neuronal mechanisms of complex auditory hallucinations.

REFERENCES

Alonso-Solís, A., Vives-Gilabert, Y., Portella, M. J., Rabella, M., Grasa, E. M., Roldán, A., & Alvarez, E. (2017). Altered amplitude of low frequency fluctuations in schizophrenia patients with persistent auditory verbal hallucinations. Schizophrenia research, 189, 97- 103.

Bonilha, L., Hillis, A. E., Hickok, G., Den Ouden, D. B., Rorden, C., & Fridriksson, J. (2017). Temporal lobe networks supporting the comprehension of spoken words. Brain, 140(9), 2370-2380.

Braun, C. M., Dumont, M., Duval, J., Hamel-Hébert, I., & Godbout, L. (2003). Brain modules of hallucination: an analysis of multiple patients with brain lesions. Journal of Psychiatry and Neuroscience, 28(6), 432.

Chen, Y. C., Xia, W., Luo, B., Muthaiah, V. P., Xiong, Z., Zhang, J., & Teng, G. J. (2015). Frequency-specific alternations in the amplitude of low-frequency fluctuations in chronic tinnitus. Frontiers in neural circuits, 9, 67.

Chen, X., Lu, B., & Yan, C.G. (2017). Reproducibility of R-fMRI metrics on the impact of different strategies for multiple comparison correction and sample sizes. Human Brain Mapping 00:00-00

Cierpka, M., Wolf, N. D., Kubera, K. M., Schmitgen, M. M., Vasic, N., Frasch, K., & Wolf, R. C. (2017). Cerebellar Contributions to Persistent Auditory Verbal Hallucinations in Patients with Schizophrenia. The Cerebellum, 16(5-6), 964-972.

Corlett, P.R., & Fletcher, P.C. (2014). Computational psychiatry: a Rosetta Stone linking the brain to mental illness. Lancet Psychiatry 1 (5): 399-402.

David, A. S. (1999). Auditory hallucinations: phenomenology, neuropsychology and neuroimaging update. Acta Psychiatrica Scandinavica, 99(s395), 95-104

De Ridder, D., Vanneste, S., & Freeman, W. (2014). The Bayesian brain: phantom percepts resolve sensory uncertainty. Neuroscience & Biobehavioral Reviews, 44, 4-15.

Eggermont, J. J., & Roberts, L. E. (2012). The neuroscience of tinnitus: understanding abnormal and normal auditory perception. Frontiers in systems neuroscience, 6, 53.

Ghazaleh, N., Van Der Zwaag, W., Clarke, S., Van De Ville, D., Maire, R., & Saenz, M. (2017). High-Resolution fMRI of Auditory Cortical Map Changes in Unilateral Hearing Loss and Tinnitus. Brain topography, 30(5), 685-697.

Hare, S. M., Ford, J. M., Ahmadi, A., Damaraju, E., Belger, A., Bustillo, J., & van Erp, T. G. (2017). Modality-dependent impact of hallucinations on low-frequency fluctuations in schizophrenia. Schizophrenia bulletin, 43(2), 389-396.

Hoare, D. J., Edmondson-Jones, M., Sereda, M., Akeroyd, M. A., & Hall, D. (2012). Amplification with hearing aids for patients with tinnitus and co-existing hearing loss. The Cochrane database of systematic reviews, 1.

Hunter, M. D., Eickhoff, S. B., Miller, T. W. R., Farrow, T. F. D., Wilkinson, I. D., & Woodruff, P. W. R. (2006). Neural activity in speech-sensitive auditory cortex during silence. Proceedings of the National Academy of Sciences of the United States of America, 103(1), 189-194.

Linszen, M. M., Brouwer, R. M., Heringa, S. M., & Sommer, I. E. (2016). Increased risk of psychosis in patients with hearing impairment: Review and meta-analyses. Neuroscience & Biobehavioral Reviews, 62, 1-20.

Linszen, M. M. J., van Zanten, G. A., Teunisse, R. J., Brouwer, R. M., Scheltens, P., & Sommer, I.E. (2018). Auditory hallucinations in adults with hearing impairment: a large prevalence study. Psychological medicine, 1-8.

Kumar, S., Sedley, W., Barnes, G. R., Teki, S., Friston, K. J., & Griffiths, T. D. (2014). A brain basis for musical hallucinations. Cortex, 52, 86-97.

Mason, O. J., & Brady, F. (2009). The psychotomimetic effects of short-term sensory deprivation. The Journal of nervous and mental disease, 197(10), 783-785.

Nam, E. C. (2005). Is it necessary to differentiate tinnitus from auditory hallucination in schizophrenic patients?. The Journal of Laryngology & Otology, 119(5), 352-355.

Powers, A. R., Kelley, M., & Corlett, P. R. (2016). Hallucinations as top-down effects on perception. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1(5), 393-400.

Ramírez, R. R., Kopell, B. H., Butson, C. R., Gaggl, W., Friedland, D. R., & Baillet, S. (2009, September). Neuromagnetic source imaging of abnormal spontaneous activity in tinnitus patient modulated by electrical cortical stimulation. In Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE (pp.1940-1944).

Sanchez, T. G., Rocha, S. C. M., Knobel, K. A. B., Kii, M. A., Santos, R. M. R. D., & Pereira, C. B. (2011). Musical hallucination associated with hearing loss. Arquivos de neuro-psiquiatria, 69(2B), 395- 400.

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. Neuroimage, 44(3), 1133-1143.

Sommer, I. E., Clos, M., Meijering, A. L., Diederen, K. M., & Eickhoff, S. B. (2012). Resting state functional connectivity in patients with chronic hallucinations. PLoS One, 7(9), e43516.

Sommer, I. E., Roze, C. M., Linszen, M. M. J., Somers, M., & van Zanten, G. A. (2014). Hearing loss; the neglected risk factor for psychosis. Schizophrenia research, 158(1), 266-267.

Song, X. W., Dong, Z. Y., Long, X. Y., Li, S. F., Zuo, X. N., Zhu, C. Z., & Zang, Y. F. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. PloS one, 6(9), e25031.

Stevens, G., Flaxman, S., Brunskill, E., Mascarenhas, M., Mathers, C. D., & Finucane, M. (2011). Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. The European Journal of Public Health, 23(1), 146-152.

Teunisse RJ, Cruysberg JR, Hoefnagels WH, Verbeek AL and Zitman FG (1996) Visual hallucinations in psychologically normal people: Charles Bonnet's syndrome. The Lancet 347, 794–797.

Teunisse, R. J., & Olde Rikkert, M. G. O. (2012). Prevalence of musical hallucinations in patients referred for audiometric testing. The American Journal of Geriatric Psychiatry, 20(12), 1075- 1077.

Thewissen, V., Myin-Germeys, I., Bentall, R., de Graaf, R., Vollebergh, W., & van Os, J. (2005). Hearing impairment and psychosis revisited. Schizophrenia Research, 76(1), 99-103.

Vanneste, S., Van de Heyning, P., & De Ridder, D. (2011). The neural network of phantom sound changes over time: a comparison between recent-onset and chronic tinnitus patients. European Journal of Neuroscience, 34(5), 718-731.

Vanneste, S., Song, J. J., & De Ridder, D. (2013). Tinnitus and musical hallucinosis: the same but more. Neuroimage, 82, 373-383.

Yang, M., Chen, H. J., Liu, B., Huang, Z. C., Feng, Y., Li, J., & Zhu, X. (2014). Brain structural and functional alterations in patients with unilateral hearing loss. Hearing research, 316, 37-43.

Yu, R., Chien, Y. L., Wang, H. L. S., Liu, C. M., Liu, C. C., Hwang, T. J., & Tseng, W. Y. I. (2014). Frequency-specific alternations in the amplitude of low-frequency fluctuations in schizophrenia. Human brain mapping, 35(2), 627-637.

Zou, Q. H., Zhu, C. Z., Yang, Y., Zuo, X. N., Long, X. Y., Cao, Q. J., & Zang, Y. F. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting- state fMRI: fractional ALFF. Journal of neuroscience methods, 172(1), 137-141.

Supplementary material for this article can be found in the digital version of the issue.

SEX HORMONES

How we came to understand them and what went wrong

by Eylül Turan

"When you meet a human being, the first distinction you make is 'male' or 'female', and you are accustomed to making the distinction with unhesitating certainty," Freud said once (as cited in Oudshoorn, 1994). Early twentieth century scientists, however, became increasingly confused by their own assumptions regarding male and female characteristics. They focused on sex hormones but encountered difficulties in making accurate definitions of male and female sex hormones. This essay explores how our understanding of sex hormones was shaped throughout the history of neuroscience and life sciences, and presents animal studies that focused on sex hormones and how their results translated rather misleadingly to humans due to a biased testing across males and females. It is inspired by Nelly Oudshoorn's book Beyond the Natural Body: An Archeology of Sex Hormones.

First, it is important to establish the difference between the terms "sex" and "gender". Rubin (1975) restricted the concept of sex to biological sex, specified by anatomical, hormonal or chromosomal criteria. Gender, on the other hand, refers to all other socially constructed characteristics attributed to males and females (e.g. social roles and psychological or behavioral characteristics).

Since the early decades of twentieth century, the prevailing view was that the concept of body was hormonally constructed. In other words, sex hormones would give rise to a natural body that would be either male or female. Scientists were in search for ovaries, testes and urines that would allow them to construct the hormonal body. Throughout their search however, they were criticized by some (e.g. Oudshoorn, 1994) claiming that the very notion of a "natural" body is a constructed one (i.e. constructed by scientists, clinicians, drug industries). Others accepted the notion of a "natural" body and its constituents (i.e. sex hormones), but were critical of how sex hormones were being tested and how results from these studies were interpreted (e.g. Berry and Zucker, 2011).

"Such conceptualization was widely accepted by the day's cultural notions regarding masculinity and femininity: females and males as opposite categories, rather than being two independent or complementary dimensions."

But how did the concept of sex hormones even arise? In his book History of Animals, Aristotle was one of the first to relate the presence of ovaries to female sexual development. At the beginning of the nineteenth century, gynecological textbooks defined ovaries as "the organ of crisis which is missing in the male body" (as cited in Gallagher and Laqueur 1987, p.27). On the other hand, the idea that the testes are related to male sexuality, bravery and longevity has its roots in Greek and Roman thinking.

These and many other pre-scientific ideas led to more scientific developments. Back in 1905, a professor at the University College in London, named Ernest H. Starling, introduced the term "hormone". In the following years, with the contribution of other scientists, sex hormones were defined as "the chemical messengers of masculinity and femininity" (Oudshoorn, 1994, p.17). Thus, the scientific conceptualization of sex and sex hormones remained very simple and straightforward with only having two sex hormones, one for each sex. Such conceptualization was widely accepted by the day's cultural notions regarding masculinity and femininity: females and males as opposite categories, rather than being two independent or complementary dimensions (Lewin, 1984).

In the 1930s a wave of criticism regarding terminology and classification arose. Especially scientists in Amsterdam repeatedly expressed their discontent about the use of the terms "male" and "female" sex hormones. Ian Hacking once wrote controversially, "We did not find sex hormones somewhere in a lost corner, like a desert island lost in the mist. We ourselves called sex hormones into existence" (as cited in Oudshoorn, 1994, p.43).

Initially, the study of sex hormones was dominated by clinicians and it was only during the 1910s that laboratory scientists entered the field, becoming the dominant researchers in the 1920s. Laboratory scientists introduced new methods such as reproductive behavior tests on guinea-pigs and rats, and focused on standardization procedures. They transferred the study of sex hormones from the clinic to the laboratory, and the clinic simply became an audience for the products they developed (Oudshoorn, 1994). Following this transformation, a landmark moment for the field was the first Conference on the Standardization of Sex Hormones in 1932. Pharmaceutical companies now also depended on the laboratories' work. For example, in the Netherlands, the Organon (a pharmaceutical company) personnel were trained by Ernst Laqueur's laboratory at the University of Amsterdam. With the laboratory as a bridge, the clinic and the pharmaceutical interest matched each others' needs and this collaboration led female sex hormones to be the subject of not only science but also of the business industry. In the early 1930s, the female body became the focus for hormone therapy, with female sex hormones being used universally to cure various diseases such as eczema and diseases of joints (both diseases were thought of being related to dysfunctioning ovaries). Sex hormones entered the market as pills controlling fertility, or taken for menstrual and menopausal problems, but not for contraception and male menopause: marketing of male sex hormones could not reach the same success.

Probably the most revolutionary discovery regarding sex hormones is the fact that they are not restricted to humans, or even to living organisms. As Robert Frank (1929) wrote, the female sex hormone could be found in the animal and even in the vegetable kingdom (as cited in Oudshoorn, 1994). Laboratory scientists were especially interested in animal testing and utilized mice and rat for their experiments.

In the 1990s, several surveys indicated that for animal testing there is a non-negligible male bias in the literature. For example, in 1984 78% and in 1991 81% of the studies reported in the journal of Behavioral Neuroscience used male animals as their subjects (Sechzer et al., 1994).

Berkley (1992) noted that in four neuroscience journals, 57% of single-sex studies focused on males, while only 17% were done using females (others did not indicate subject sex). As animal models were widely used for developing new treatments, these findings brought to question the reliability of female health-care medications. Unfortunately, the situation did not improve, as in 2009 studies of single-sex mammals in Behavioral Neuroscience had a male bias of 65% (Beery & Zucker, 2011). In fact, in 2009, 8 out of the 10 biological disciplines had a male-bias with neuroscience being the most pronounced one (a ratio of 5.5 males to 1 female). The sex of subjects was not reported in 22-42% of neuroscience, physiology and interdisciplinary biology journals. Additionally, 75% of the studies in three highly cited immunology journals did not specify the sex of the animals that they used. For human studies, although male bias was evident in some fields, including interdisciplinary neuroscience, compared to non-human animal research the number of articles reported on both sexes was higher. Interestingly, while male-only studies have increased since 1969 in non-human animal research, human research showed an opposite trend: the majority of studies since 1993 investigated both sexes. This raises questions regarding the reliability of how results from non-human animal studies translate to humans.

For instance, although women are diagnosed with anxiety disorder twice as often as men (Bekker and van Mens-Verhulst, 2007), most of the animal studies that focus on anxiety and anxiolytic drugs use male rats (Palanza, 2011). Further, although women experience stroke events more often than men over the course of their lives (Reeves et al., 2008), studies in several journals that focused on animal models of stroke (e.g. Journal of Stroke and Cerebrovascular Disease) reported results that were male biased: 65% focused on males, none were on females, 10% were on both sexes, and 25% did not specify sex (Beery and Zucker, 2011). Last but not least, despite the fact that women are at greater risk than men of suffering from clinical pain conditions (Fillingim et al., 2009), 79% of the studies reported in the Journal of Pain between 1996 and 2005 included only male subjects (Mogil and Chanda, 2005).

"Although women are diagnosed with anxiety disorder twice as often as men, most of the animal studies that focus on anxiety and anxiolytic drugs use male rats."

Berkley (1992) as well as Sechzer and colleagues (1994) encouraged journals to make sex-specification part of their policy, and to advise against generalizing one-sex studies to the opposite-sex. Unfortunately, these recommendations have been neglected by most of the scientific community. Even today, journal policies do not greatly encourage study of both sexes and accurate reporting of subject sex.

There are two arguments stated for not studying females in animal research. Some researchers believe that drugs might have adverse effects on females and therefore generalize findings from males to females, while others simply consider males as representative of the human species and accept them as the norm (Marts and Keitt, 2004). Many researchers find it time and resource consuming to control for female estrous/menstrual cycles or say that after understanding the phenomenon in males, they'll check whether it's there in females (McCarthy et al., 2012). McCarthy and colleagues (2012) believe that the main reason actually arises due to misconceptions that it is challenging to do it the right way, and misconceptions of thinking that comparisons between females and males would not provide valuable insights.

As Mogil and Chanda (2005) stated it seems that "basic scientists are shirking their responsibilities to half of the human population by avoiding the use of direct animal models of them" (Mogil and Chanda, 2005, p. 4). It should be noted, however, that there have been encouraging developments over the past few years such as the establishment of the Organization for the Study of Sex Differences which met in 2007 for the first time and the initiative of a new journal: Biology of Sex Differences.

REFERENCES

Beery, A. K., & Zucker, I. (2011). Sex bias in neuroscience and biomedical research. *Neuroscience & Biobehavioral Reviews, 35*(3), 565-572.

Bekker, M. H., & van Mens-Verhulst, J. (2007). Anxiety disorders: sex differences in prevalence, degree, and background, but gender-neutral treatment. *Gender medicine, 4*, S178-S193.

Berkley, K. J. (1992). Vive la difference. *Trends Neurosci, 15*(9), 331-332.

Fillingim, R. B., King, C. D., Ribeiro-Dasilva, M. C., Rahim-Williams, B., & Riley III, J. L. (2009). Sex, gender, and pain: a review of recent clinical and experimental findings. *The journal of pain, 10*(5), 447-485.

Gallagher, C., & Laqueur, T. W. (Eds.). (1987). *The making of the modern body: Sexuality and society in the nineteenth century (No. 1)*. Univ of California Press.

Lewin, M. (1984). Psychology measures femininity and masculinity, 2: From "13 gay men" to the instrumental-expressive distinction. *In the shadow of the past: Psychology portrays the sexes*, 179-204.

Marts, S. A., & Keitt, S. (2004). Foreword: a historical overview of advocacy for research in sex-based biology. In *Advances in Molecular and Cell Biology* (Vol. 34, pp. v-xiii). Elsevier.

McCarthy, M. M., Arnold, A. P., Ball, G. F., Blaustein, J. D., & De Vries, G. J. (2012). Sex differences in the brain: the not so inconvenient truth. *Journal of Neuroscience, 32*(7), 2241-2247.

Mogil, J. S., & Chanda, M. L. (2005). The case for the inclusion of female subjects in basic science studies of pain. *Pain, 117*(1), 1-5.

Oudshoorn, N. (1994). *Beyond the Natural Body: An Archeology of Sex Hormones.*

Palanza, P. (2001). Animal models of anxiety and depression: how are females different?. *Neuroscience & Biobehavioral Reviews, 25*(3), 219-233.

Reeves, M. J., Bushnell, C. D., Howard, G., Gargano, J. W., Duncan, P. W., Lynch, G., Khatiwoda, A. & Lisabeth, L. (2008). Sex differences in stroke: epidemiology, clinical presentation, medical care, and outcomes. *The Lancet Neurology, 7*(10), 915-926.

Rubin, G. (1975). *The Traffic in Women: Notes on the "Political Economy" of Sex. Teoksessa Rayna R. Reiter (toim.), toward an anthropology of Women.*

Sechzer, J. A., Rabinowitz, V. C., Denmark, F. L., McGINN, M. F., Weeks, B. M., & Wilkens, C. L. (1994). Sex and Gender Bias in Animal Research and in Clinical Studies of Cancer, Cardiovascular Disease, and Depression a. *Annals of the New York Academy of Sciences, 736*(1), 21-48.

Longitudinal Assessment of Radiation Therapy Effects on White Matter Structures: a Diffusion MRI Study

Victor A. Bodiut*

Supervisors: Szabolcs David, dr. Alexander Leemans

PROVIDI Lab, Image Sciences Institute, University Medical Center Utrecht

ABSTRACT

Purpose: To demonstrate the long-term, regional sensitivity of white matter (WM) structures to radiation exposure during clinical protocols of radiation therapy (RT), using diffusion MRI (dMRI).

Methods and Materials: The sample consisted of 38 patients (20 females), showing clinical heterogeneity (e.g., age, pathology, tumor location, treatment planning). The imaging dataset included 38 individual CT scans and 393 MRI / dMRI scans assessed pre-operative, post-operative pre-RT, and post-RT longitudinally. Following a pre-processing pipeline tuned for standard neuro-oncological settings, we performed an atlas-based ROI analysis and estimated DTI-based metrics for all datapoints. We achieved the end stage of having the ROI delineation, ROI volume, DTI metrics (i.e., FA, MD, AD and RD) and RT dose in the same space.

Results: FA showed a stronger pre – post RT decreasing trend in WM that received higher doses compared to WM that received lower doses. Longitudinally, a similar trend but a shift in amplitude was observed (i.e., lower FA at higher doses). MD showed a clear increasing trend in WM that received higher doses, compared to WM that received lower doses. This was again the most evident between the pre- and first post-RT assessments, but persisted in time. The volume of ROIs showed variability (i.e., up to 60-80% change) both within and between patients.

Conclusions: The volume of structures is changing in time, and this must be accounted for due to partial voluming effects. Results, overlooking this, are prone to error. Then, there seems to be a longitudinal effect (consistent with literature) of dose levels on DTI metrics of WM structures, with higher doses leading to a decrease in FA and increase in MD. In summary, when using an adequate and robust analysis pipeline, dMRI proves to be insightful in longitudinal neuro-oncological settings.

KEYWORDS

radiation therapy, white matter structures, dMRI, brain tumor, cognitive impairment

INTRODUCTION

Radiation therapy (RT) is a common treatment procedure for both primary and metastatic tumors in the brain, often in combination with surgery and chemotherapy. As radiation is not selective to tumor cells but targets all cells in the process of replication, the efficacy of RT is hindered by the radio-resistance of healthy tissue following ionization (Cox, Stetz, & Pajak, 1995; Dawson & Jaffray, 2007; Kelley, Knisely, Symons, & Ruggieri, 2016). The effect of radiation on brain tissue is dynamic and involves structures outside the targeted tumor volume, directly or indirectly (Schultheiss, Kun, Ang, & Stephens, 1995; Tofilon, Fike, Tofilona, & Fikeb, 2016). Despite the substantial advances in RT technology and application (i.e., increased precision and conformality, intensity-modulated techniques, fractionated stereotactic radiosurgery; Baskar, Lee, Yeo, & Yeoh, 2012; Bucci, Bevan, & Roach, 2005), regional sensitivity to radiation dosing is not well documented. This is especially true for white matter (WM) structures. Current clinical protocols include guidelines of maximum dosing for brain parenchyma and several organs at risk for which the structure-function relationship is well-described (e.g., brain stem, optic chiasm, hippocampi; Scoccianti et al., 2015). However, regional constraints on WM structures are not a standard consideration in part because the focus is mainly placed on its connective or supportive role rather than its function, and in part because of the methodological limitations of WM-related imaging (Connor et al., 2016; Gondi et al., 2014; Nagesh et al., 2008).

Radiation-induced WM damage has been reported to include axonal injury, demyelination, neuro-inflammation, and necrosis (Kumar et al., 2000; Nagesh et al., 2008; Wang et al., 2009). Importantly, these structural deficits seem to correlate in time with both verbal and non-verbal functional cognitive impairment, including executive functioning, working memory, visuospatial processing, and decision making (Douw et al., 2009; Greene-Schloesser et al., 2012; Kerchner et al., 2012). Temporally, the radiation-induced cognitive impairment has been divided into three phases post-RT: acute (<2 weeks), early-delayed (2 weeks to 3-4 months) and late-delayed (6+ months). Notably, while acute and early-delayed damage seems to be usually transient, late-delayed damage is usually permanent. This progressive decline affects the physical and mental health of long-term survivors and impairs their quality of life (Greene-Schloesser et al., 2012). Late-delayed cognitive impairment currently occurs in 50-90% of survivors (Johannesen, Lien, Hole, & Lote, 2003; Meyers & Brown, 2006), and this population is increasing with RT advancement (Krex et al., 2007; Weller et al., 2005). Besides high doses (e.g., 60 Gy total) known to be dangerous, lower doses (e.g., 20 Gy total) have been reported to cause late-delayed damage as well (Chapman et al., 2012). Therefore, information on regional WM sensitivity to RT dosage in the long-term is crucial to the RT planning and patient well-being. The fact that current clinical protocols are not pragmatically informed by these findings is worthy of attention.

Diffusion MRI (dMRI) has proven to be a useful non-invasive imaging technique to study WM structure in the brain in the past years. The fundamental principle behind dMRI is that the Brownian motion of water molecules that is dependent on the surrounding tissue structure can be imaged. As opposed to grey matter (GM) where the structures tend to be more spherical or isotropic, WM structures tend to be more elongated or anisotropic. This is reflected in the diffusion metrics, making the technique the most sensitive to WM anatomy (for comprehensive descriptions see Jones, 2010; Basser & Jones, 2002; Jones & Leemans, 2011). Neuro-oncological research has placed dMRI as a potential technique for tumor diagnosis (Kono et al., 2001), surgical planning (Nimsky et al., 2005a, 2005b), pre-treatment prediction of tumor response (Mardor et al., 2004), monitoring early efficacy of treatment (Chenevert et al., 2000), early WM damage post-radiation (Kumar et al., 2008; Nagesh et al., 2008; Price et al., 2003), and more recently late-delayed effects of RT on WM (Chapman et al., 2012; Connor et al., 2016, 2017; Zhu et al., 2016).

Thus, we aimed to build upon the recent relevant findings and study the long-term effects of brain RT on WM structures using dMRI. Specifically, we were interested in how the susceptibility of WM structures to radiation varies across regions and dose levels on the long-term.

METHODS AND MATERIALS

Sample and treatment (see Table 1 for an overview)

The sample consisted of 38 patients (20 females) who were treated at the University of Texas MD Anderson Cancer Center, Houston, Texas. The dataset showed heterogeneity on several levels, including age (i.e., range 24-88, Mdn = 57), pathology (glioblastoma – 62%), anatomical location of tumor, type of surgery, radiation technique (IMRT – 74%), dosage (60Gy in 30 fractions – 59%), and chemotherapy (Temodar – 64%).

All patients without any MRI / dMRI scans post-RT were excluded from the final analysis (N = 6), resulting in analyzable data from 32 patients.

CT acquisition

The 38 baseline CT scans (used for RT planning) were acquired on a Philips Brilliance Big Bore scanner, with a tube potential of 120 kVp, using a matrix size of 512 × 512 × 87 and 0.98 × 0.98 × 3.0 mm voxel size.

MRI acquisition (see Table 2 in Appendix for an overview)

MR imaging was acquired at multiple timepoints for each patient: pre-operative, post-operative pre-RT, and post-RT longitudinally (see Table 3 in Appendix for an overview). The dataset included 393 usable MRI / dMRI scans (range 2-43, Mdn = 8 per patient).

The MRI acquisition used a T1-weighted sequence on a GE Medical Systems scanner of type Signa HDxt in 91% of cases (SIGNA EXCITE - 5%; Signa HDx - 4%; Optima MR450w - 1%) with an 8HRBRAIN coil in 97% of cases, 256 × 256 matrix size and 220 × 220 mm2 field of view (FOV). The other acquisition settings showed greater heterogeneity both between and within patients in terms of field strength (i.e., 1.5T - 52%, 3T - 48%), TE, TR, and voxel dimensions (i.e., 0.85 × 0.85 × 6.5 mm3 - 57%, 0.85 × 0.85 × 3.5 mm3 - 35%, varied for the rest).

The dMRI images were acquired using 27 gradient directions with a diffusion weighting factor of b = 1200 s/mm2 in 93% of cases and b = 1000 s/mm2 for the rest, with one non-diffusion weighted gradient direction of b = 0 s/mm2 (see Figure 1) .

The diffusion data was processed using the MATLAB-based (MATLAB Release 2014b, The MathWorks, Inc., Natick, MA) software ExploreDTI (Leemans, Jeurissen, Sijbers, & Jones, 2009).

To increase registration performance and reduce computational power, we designed a cropping algorithm based on the standardized Hounsfield scaling of CT acquisition (i.e., with values around -1000 for air, 0 for water and 200+ for bone; see also Feeman, 2010) that outputs a bounding box of the brain. The same cropping settings were used for the relevant dose map for each patient (i.e., having the dose map in the same native space).

Then, the DWI data at each timepoint was corrected for motion and eddy-current distortions (with the b = 0 s/mm2 image as the reference for the latter) and registered to the CT space using rigid registration based on mutual information (Elastix; Klein, Staring, Murphy, Viergever, & Pluim, 2010). The quality and spatial alignment of registration was visually inspected (i.e., overlays, movie

*FULL LIST OF AUTHORS: Bodiut, V.A., David, S., De Luca, A., Philippens, M.E.P., Elhalawani, H., Ding, Y., Mohamed, A., Chung, C., Radwan, S., Fuller C.D., Verhoeff, J.J.C., Leemans, A.

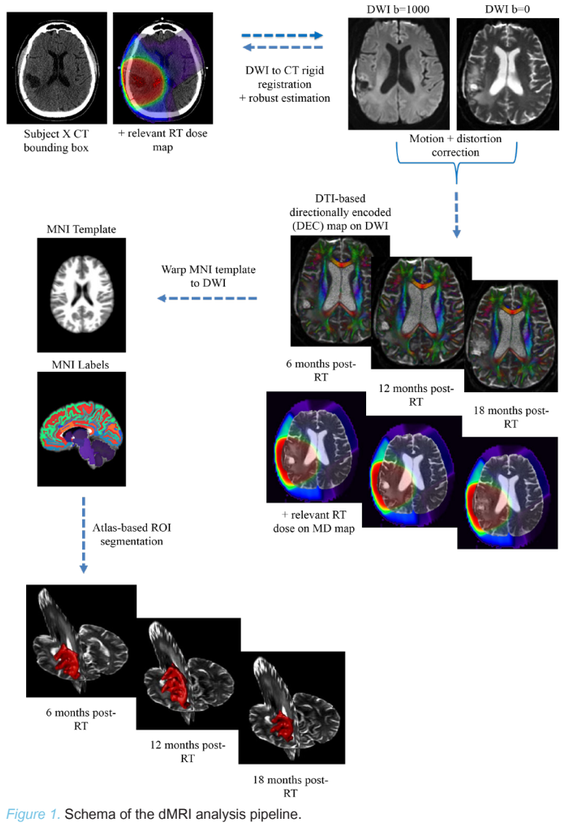


Figure 1. Schema of the dMRI analysis pipeline.

loops, physically implausible signal maps, residual maps; also see Leemans & Jones, 2009; Tournier, Mori, & Leemans, 2011; Vos et al., 2017). Initially, this registration step was planned to use the MRI images at each timepoint as targets (to take advantage of the anatomical contrast information), but the heterogeneity in MRI acquisition was simply too high (as described previously). Nevertheless, registering to CT space proved to be both feasible and adequate quality-wise.

Following, a robust estimation of the diffusion tensor (DTI) was done with the in-house algorithm REKINDLE (robust extraction of kurtosis indices with linear estimation; Tax, Otte, Viergever, Dijkhuizen, & Leemans, 2015).

We then performed an atlas-based ROI analysis (Mori et al., 2008) on the diffusion data by warping the WM template in MNI152 (Montreal Neurological Institute) space from FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) on the individual scans. This method provided DTI metrics based on pre-defined volumetric ROIs in the pre-registered native space, aiming for a high spatial alignment between anatomical and scalar data for each datapoint. As the dose maps were also in the native space, all relevant information (i.e., ROI delineation, ROI volume, DTI metrics and RT dose) was available in the same space.

Therefore, we computed the following metrics per ROI: average fractional anisotropy (FA; ranging from 0 to 1, +SD), average mean diffusivity (MD; multiplied by 10k, +SD), average axial diffusivity (AD; multiplied by 10k, +SD), average radial diffusivity (RD; multiplied by 10k, +SD), volume (in mm3), mean dose (+SD), max dose (+SD), and median dose (+SD).

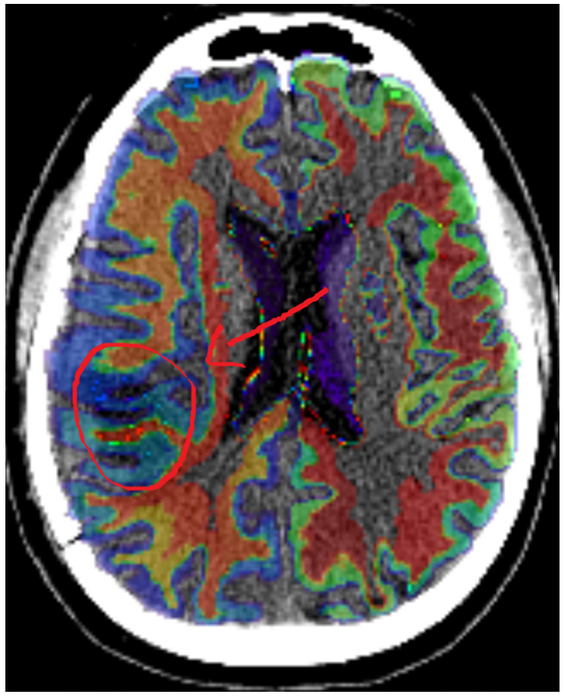


Figure 2. Atlas labels overlaid on CT for one patient, used for the manual exclusion of ROIs overlapping the tumor location.

To avoid the bias of remaining cancerous tissue or liquid (i.e., post-operatively) on the DTI metrics, all ROIs overlapping the tumor location were entirely excluded manually (N = 301; M = 10 per patient; see Figure 2).

EXPLORATORY RESULTS

For a proof of concept, we selected data from one patient with homogeneous acquisition settings (i.e., field strength, voxel dimensions, bvalue, etc.). Then, we applied a dichotomous selection between WM ROIs that received a low mean dose (i.e., lower than 10Gy; N = 7) and a high mean dose (i.e., higher than 40Gy; N = 5) and investigated the changes between the two groups in FA, MD and volume. No statistical analysis was performed at this point.

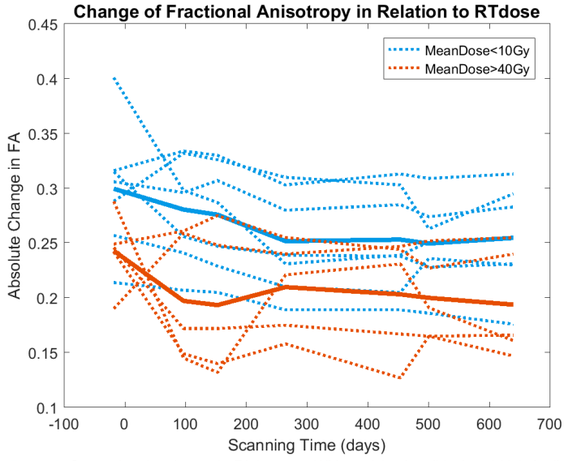


Figure 3. Comparing the change in FA between WM receiving low (blue) and high (red) mean RT dose. The dashed lines represent FA of individual ROIs, while the solid lines represent the average FA across ROIs, per group. overlapping the tumor location.

FA showed a stronger decreasing trend in WM that received higher doses compared to WM that received lower doses between the pre- and post-RT assessments. Across time, the trend was similar between groups, but a clear shift in FA amplitude was observed (i.e., lower FA at higher doses; see Figure 3).

MD showed a clear increasing trend in WM that received higher doses, compared to WM that received lower doses. Once again, this was the most evident between the pre- and post-RT assessments, but persisted in time (see Figure 4).

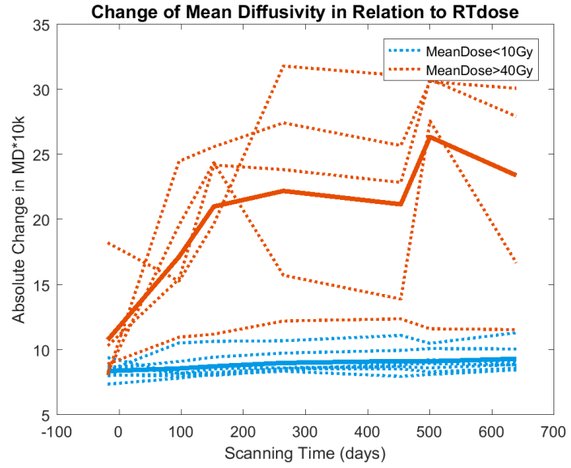


Figure 4. Comparing the change in MD between WM receiving low (blue) and high (red) mean RT dose. The dashed lines represent MD of individual ROIs, while the solid lines represent the average MD across ROIs, per group.

The volume of ROIs showed variability (i.e., up to 60-80% change) both within and between patients. No clear trend in relation to RT dose was observed, but while no ROIs showed a decrease in volume in the lower dose group, several ROIs showed shrinkage in the higher dose group (see Figure 5).

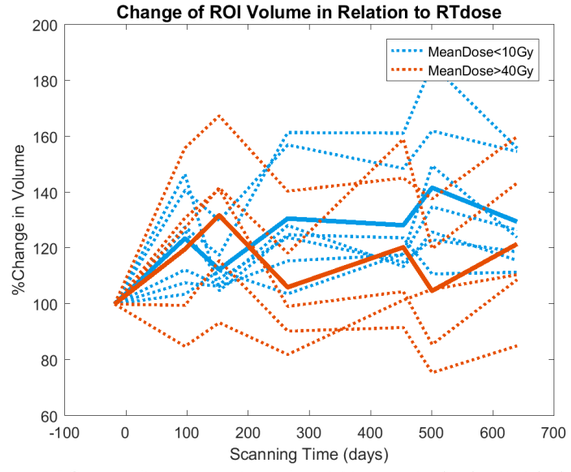


Figure 5. Comparing the change in volume between WM receiving low (blue) and high (red) mean RT dose. The dashed lines represent volume of individual ROIs, while the solid lines represent the average volume across ROIs, per group. mean RT dose. The dashed lines represent MD of individual ROIs, while the solid lines represent the average MD across ROIs, per group.

DISCUSSION

First, the fact that the volume of individual WM ROIs is changing in time due to RT has been neglected in the literature. To our interpretation, this is partly a biological phenomenon where the actual tissue volume is changing, and partly an imaging acquisition phenomenon where the heterogeneity of acquisition settings leads to a change in volume. Notably, the results show that even with homogeneous acquisition settings, the change in volume is relevant.

Volume change in brain structures has been of interest especially because of its potential link to cognitive functioning, such as a volume-memory relationship in the hippocampus (see Van Petten, 2004 for a review). RT-related changes in hippocampal volume seem to be consistent, show a dose-dependency, and correlate with follow-up memory assessment, making volume an important biomarker (Blomstrand, Kalm, Grandér, Björk-Eriksson, & Blomgren, 2014; Ma et al., 2017; Nolen et al., 2016; Seibert et al., 2017).

Moreover, as established in the last years, partial volume effects (PVE) have an influence on the diffusion metrics (Alexander, Hasan, Lazar, Tsuruda, & Parker, 2001). The principle behind PVE is that at tissue interfaces (e.g., CSF – WM), the limited imaging resolution of a voxel will include a mixture of tissue types, leading to measuring different anatomical structures which cannot be separated. Critically, Vos et al. (2011) demonstrated that the thickness, structure, orientation, and curvature of a WM bundle significantly modulates the PVE (for a representation of the phenomenon see Figure 6, with permission from Vos et al., 2011).

Therefore, besides the volume change per se being important functionally, the volume change of WM structures observed here modulates the respective PVE, in turn varying the DTI metrics. Analyses failing to acknowledge this effect are prone to error. To our knowledge, previous works on regional WM sensitivity to RT discuss volumes within ROIs receiving certain dose levels (e.g., Chapman et al., 2012; Connor et al., 2017; Zhu et al., 2016), or aim to avoid PVE by restricting the masking to connected voxels that only include WM structure (e.g., Connor et al., 2016), but none has accounted for the volume change effect described above.

Second, the effects of RT dosage on the DTI metrics seem to be consistent with most of the literature. Specifically, higher doses lead to a steeper decrease in FA and increase in MD, especially in the early-delayed timeframe. The persistent trend across time in MD is informative about the slow progressive WM change in the late-delayed timeframe. The similar trend, but amplitude shift observed in FA is also interesting, and could be interpreted as meaning that regions where higher doses are prescribed are already more affected by cancerous tissue infiltration. Biologically, the decrease in FA and increase in MD in this dose-tissue context most probably translate to higher dosage causing a loss of structure (i.e., higher percentage of liquid) in the affected WM structures. As expected, variability in the way RT dose influences DTI metrics can be observed both between patients and between WM structures. However, once again, these results need to be verified after accounting for the modulating effect of volume change.

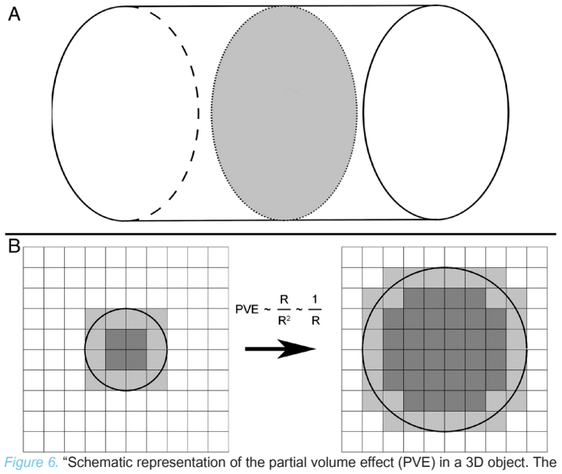


Figure 6. "Schematic representation of the partial volume effect (PVE) in a 3D object. The PVE is defined as the volume-to-surface ratio of an object. For a cylinder, which has a circular cross-section (indicated as the shaded area in A), this can be simplified to a circle, which has a surface-to-circumference ratio. Further simplification yield that the PVE scales with 1/R (with R the radius of the circle), showing that increasing volume means a reduction in PVE. This is shown in B, where a small and big circle have been plotted on a square grid. The relative number of PVE voxels (light gray) compared to voxels enclosed completely by the circle (dark gray) is larger for a small circle (left) than for a big circle (right)".

Third, we hope to have proven here that by using an adequate and robust pipeline tuned for typical clinical data, it is feasible to study detailed anatomical changes longitudinally following radiation treatment. The key of the analysis was having all relevant information in the same (native) space. To our surprise, using the CT space instead of the MRI space as registration target for the DWI data proved to be satisfactory. Still, we aim for a conservative approach in anatomical (e.g., demyelination, inflammation) and cognitive (e.g., causing dementia) interpretation due to the limitations in data acquisition and quality, and the exploratory nature of the results.

FUTURE DIRECTIONS

We are currently working on a similar analysis as presented above but including the whole spectrum of RT dosage across WM ROIs and patients. Then, we plan a large-scale statistical inference modelling the slopes (β) of change in DTI metrics per increase unit of dose (Gy; similar but simpler to the one described in Zhu et al., 2016). One of the main aims of the analysis will be to account for (i.e., regress out) the effect of volume change described earlier and so quantify its influence on the results. Generally, research in the field should aim for a robust analysis pipeline and proceed with care regarding strong anatomical and cognitive interpretation, giving the typical clinical data quality.

Clinically, the dMRI sequence has proven to be insightful in neuro-oncological settings (as described previously), including longitudinal dose-tissue interaction effects following RT (as described here). Therefore, we argue that it should be included in the standard acquisition protocol for treatment of brain cancer, and that WM should be included in the guidelines of radiation planning, complementing the established organs at risk. The acquisition protocol should be informed by the limitations pointed out in research, such as less heterogeneity in scanner settings, more isotropic voxel dimensions, and adequate b-value and gradient settings.

REFERENCES

Alexander, A. L., Hasan, K. M., Lazar, M., Tsuruda, J. S., & Parker, D. L. (2001). Analysis of partial volume effects in diffusion-tensor MRI. Magnetic Resonance in Medicine, 45(5), 770–780. https://doi.org/10.1002/mrm.1105

Baskar, R., Lee, K. A., Yeo, R., & Yeoh, K. W. (2012). Cancer and radiation therapy: Current advances and future directions. International Journal of Medical Sciences, 9(3), 193–199. https://doi.org/10.7150/ijms.3635

Basser, P. J., Basser, P. J., Jones, D. K., & Jones, D. K. (2002). Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. NMR in Biomedicine, 15, 456–67. https://doi.org/10.1002/nbm.783

Blomstrand, M., Kalm, M., Grandér, R., Björk-Eriksson, T., & Blomgren, K. (2014). Different reactions to irradiation in the juvenile and adult hippocampus. International Journal of Radiation Biology, 90(9), 807–815. https://doi.org/10.3109/09553002.2014.942015

Bucci, M. K., Bevan, A., & Roach, M. (2005). Advances in radiation therapy: conventional to 3D, to IMRT, to 4D, and beyond. CA: A Cancer Journal for Clinicians, 55(2), 117–134. https://doi.org/10.3322/canjclin.55.2.117

Chapman, C. H., Nagesh, V., Sundgren, P. C., Buchtel, H., Chenevert, T. L., Junck, L., Cao, Y. (2012). Diffusion tensor imaging of normal-appearing white matter as biomarker for radiation-induced late delayed cognitive decline. International Journal of Radiation Oncology Biology Physics, 82(5), 2033–2040. https://doi.org/10.1016/j.ijrobp.2011.01.068

Chenevert, T. L., Lauren, D., Taylor, J. M. G., Robertson, P. L., Harry, S., Rehemtulla, A., & Ross, B. D. (2018). Resonance Imaging : an Early Surrogate Marker of Brain Tumors Background : A surrogate marker for treatment response that can be ob- served earlier than comparison of se- quential magnetic resonance imaging (MRI) scans , which depends on rela- tively sl, 92(24), 2029–2036.

Connor, M., Karunamuni, R., McDonald, C., Seibert, T., White, N., Moiseenko, V., … Hattangadi-Gluth, J. A. (2017). Regional susceptibility to dose-dependent white matter damage after brain radiotherapy. Radiotherapy and Oncology, 123(2), 209–217. https://doi.org/10.1016/j.radonc.2017.04.006

Connor, M., Karunamuni, R., Mcdonald, C., White, N., Pettersson, N., Moiseenko, V., … Hattangadi-Gluth, J. (2016). White matter damage after RT Dose-dependent white matter damage after brain radiotherapy. Radiotherapy and Oncology, 121, 209–216. https://doi.org/10.1016/j.radonc.2016.10.003

Cox, J. D., Stetz, J., & Pajak, T. F. (1995). Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European organization for research and treatment of cancer (EORTC). International Journal of Radiation Oncology, Biology, Physics, 31(5), 1341–1346. https://doi.org/10.1016/0360-3016(95)00060-C

Dawson, L. A., & Jaffray, D. A. (2007). Advances in image-guided radiation therapy. Journal of Clinical Oncology, 25(8), 938–946. https://doi.org/10.1200/JCO.2006.09.9515

Douw, L., Klein, M., Fagel, S. S., van den Heuvel, J., Taphoorn, M. J., Aaronson, N. K., … Heimans, J. J. (2009). Cognitive and radiological effects of radiotherapy in patients with low-grade glioma: long-term follow-up. The Lancet Neurology, 8(9), 810–818. https://doi.org/10.1016/S1474-4422(09)70204-2

Gondi, V., Pugh, S. L., Tome, W. A., Caine, C., Corn, B., Kanner, A., … Mehta, M. P. (2014). Preservation of memory with conformal avoidance of the hippocampal neural stem-cell compartment during whole-brain radiotherapy for brain metastases (RTOG 0933): A phase II multi-institutional trial. Journal of Clinical Oncology, 32(34), 3810–3816. https://doi.org/10.1200/JCO.2014.57.2909

Greene-Schloesser, D., Robbins, M. E., Peiffer, A. M., Shaw, E. G., Wheeler, K. T., & Chan, M. D. (2012). Radiation-induced brain injury: A review. Frontiers in Oncology, 2(July), 1–18. https://doi.org/10.3389/fonc.2012.00073

Johannesen, T. B., Lien, H. H., Hole, K. H., & Lote, K. (2003). Radiological and clinical assessment of long-term brain tumour survivors after radiotherapy. Radiotherapy and Oncology, 69(2), 169–176. https://doi.org/10.1016/S0167-8140(03)00192-0

Jones, D. K., & Leemans, A. (2011). Diffusion Tensor Imaging. Methods in Molecular Biology, 711(2), 127–144. https://doi.org/10.1097/01.chi.0000246064.93200.e8

Kelley, K., Knisely, J., Symons, M., & Ruggieri, R. (2016). Radioresistance of brain tumors. Cancers, 8(4), 1–28. https://doi.org/10.3390/cancers8040042

Kerchner, G. A., Racine, C. A., Hale, S., Wilheim, R., Laluz, V., Miller, B. L., & Kramer, J. H. (2012). Cognitive Processing Speed in Older Adults: Relationship with White Matter Integrity. PLoS ONE, 7(11). https://doi.org/10.1371/journal.pone.0050425

Klein, S., Staring, M., Murphy, K., Viergever, M. A., & Pluim, J. P. W. (2010). Elastix: A toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging, 29(1), 196–205. https://doi.org/10.1109/TMI.2009.2035616

Kono, K., Inoue, Y., Nakayama, K., Shakudo, M., Morino, M., Ohata, K., … Yamada, R. (2001). The role of diffusion-weighted imaging in patients with brain tumors. AJNR American Journal of Neuroradiology, 22(6), 1081–1088. Retrieved from papers3://publication/uuid/C2BFF7E3-8CD1-4EA6-8958-0484D3DC298F

Krex, D., Klink, B., Hartmann, C., Von Deimling, A., Pietsch, T., Simon, M., … Schackert, G. (2007). Long-term survival with glioblastoma multiforme. Brain, 130(10), 2596–2606. https://doi.org/10.1093/brain/awm204

Kumar, A. J., Leeds, N. E., Fuller, G. N., Van Tassel, P., Maor, M. H., Sawaya, R. E., & Levin, V. A. (2000). Malignant Gliomas: MR Imaging Spectrum of Radiation Therapy- and Chemotherapy-induced Necrosis of the Brain after Treatment. Radiology, 217(2), 377–384. https://doi.org/10.1148/radiology.217.2.r00nv36377

Kumar, S., Haris, M., Purwar, A., Raj, M., Das, K., Behari, S., … Gupta, R. (2008). Serial Diffusion Tensor Imaging to Characterize Radiation Induced Changes in Normal Appearing White Matter following Radiotherapy in Patients with Adult Low Grade Gliomas (WHO grade II). Proceedings 16th Scientific Meeting, International Society for Magnetic Resonance in Medicine, Toronto, 3493. Retrieved from /MyPathway2008/3493

Leemans, A., Jeurissen, B., & Sijbers, J. (2009). ExploreDTI : a graphical toolbox for processing, (January 2016), 5–6.

Leemans, A., & Jones, D. K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. Magnetic Resonance in Medicine, 61(6), 1336–1349. https://doi.org/10.1002/mrm.21890

Ma, T. M., Grimm, J., McIntyre, R., Anderson-Keightly, H., Kleinberg, L. R., Hales, R. K., … Redmond, K. J. (2017). A prospective evaluation of hippocampal radiation dose volume effects and memory deficits following cranial irradiation. Radiotherapy and Oncology, 125(2), 234–240. https://doi.org/10.1016/j.radonc.2017.09.035

Mardor, Y., Roth, Y., Ocherashvilli, A., Spiegelmann, R., Tichler, T., Daniels, D., … Pfeffer, R. (2004). Pretreatment Prediction of Brain Tumors Response to Radiation Therapy Using High b-Value Diffusion-Weighted MRI. Neoplasia, 6(2), 136–142. https://doi.org/10.1593/neo.03349

Meyers, C. A., & Brown, P. D. (2006). Role and relevance of neurocognitive assessment in clinical trials of patients with CNS tumors. Journal of Clinical Oncology, 24(8), 1305–1309. https://doi.org/10.1200/JCO.2005.04.6086

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., … Mazziotta, J. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. NeuroImage, 40(2), 570–582. https://doi.org/10.1016/j.neuroimage.2007.12.035

Nagesh, V., Tsien, C. I., Chenevert, T. L., Ross, B. D., Lawrence, T. S., Junick, L., & Cao, Y. (2008). Radiation-Induced Changes in Normal-Appearing White Matter in Patients With Cerebral Tumors: A Diffusion Tensor Imaging Study. International Journal of Radiation Oncology Biology Physics, 70(4), 1002–1010. https://doi.org/10.1016/j.ijrobp.2007.08.020

Nimsky, C., Ganslandt, O., Hastreiter, P., Wang, R., Benner, T., Sorensen, A. G., & Fahlbusch, R. (2005a). Intraoperative Diffusion-Tensor MR Imaging: Shifting of White Matter Tracts during Neurosurgical Procedures—Initial Experience. Radiology, 234(1), 218–225. https://doi.org/10.1148/radiol.2341031984

Nimsky, C., Ganslandt, O., Hastreiter, P., Wang, R., Benner, T., Sorensen, A. G., & Fahlbusch, R. (2005b). Preoperative and intraoperative diffusion tensor imaging-based fiber tracking in glioma surgery. Neurosurgery, 56(1), 130–137. https://doi.org/10.1227/01.NEU.0000144842.18771.30

Nolen, S. C., Lee, B., Shantharam, S., Yu, H. J., Su, L., Billimek, J., & Bota, D. A. (2016). The effects of sequential treatments on hippocampal volumes in malignant glioma patients. Journal of Neuro-Oncology, 129(3), 433–441. https://doi.org/10.1007/s11060-016-2188-8

Price, S. J., Burnet, N. G., Donovan, T., Green, H. A. L., Pe A, A., Antoun, N. M., … Gillard, J. H. (n.d.). Diffusion Tensor Imaging of Brain Tumours at 3 T: A Potential Tool for Assessing White Matter Tract Invasion? https://doi.org/10.1016/S0009-9260(03)00115-6

Schultheiss, T. E., Kun, L. E., Ang, K. K., & Stephens, L. C. (1995). Radiation response of the central nervous system. International Journal of Radiation Oncology, Biology, Physics, 31(5), 1093–1112. https://doi.org/10.1016/0360-3016(94)00655-5

Scoccianti, S., Detti, B., Gadda, D., Greto, D., Furfaro, I., Meacci, F., … Livi, L. (2015). Organs at risk in the brain and their dose-constraints in adults and in children: A radiation oncologist's guide for delineation in everyday practice. Radiotherapy and Oncology. https://doi.org/10.1016/j.radonc.2015.01.016

Seibert, T. M., Karunamuni, R., Bartsch, H., Kaifi, S., Krishnan, A. P., Dalia, Y., … Hattangadi-Gluth, J. A. (2017). Radiation Dose–Dependent Hippocampal Atrophy Detected With Longitudinal Volumetric Magnetic Resonance Imaging. International Journal of Radiation Oncology Biology Physics, 97(2), 263–269. https://doi.org/10.1016/j.ijrobp.2016.10.035

Tax, C. M. W., Otte, W. M., Viergever, M. A., Dijkhuizen, R. M., & Leemans, A. (2015). REKINDLE: Robust Extraction of Kurtosis INDices with Linear Estimation. Magnetic Resonance in Medicine, 73(2), 794–808. https://doi.org/10.1002/mrm.25165

Tofilon, P. J., Fike, J. R., Tofilona, P. J., & Fikeb, J. R. (2016). The Radioresponse of the Central Nervous System : A Dynamic Process The Radioresponse of the Central Nervous System : A Dynamic Pr, 153(4), 357–370.

Tournier, J.-D., Mori, S., & Leemans, A. (2011). Diffusion tensor imaging and beyond. Magnetic Resonance in Medicine, 65(6), 1532–1556. https://doi.org/10.1002/mrm.22924

Van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: Review and meta-analysis. Neuropsychologia, 42(10), 1394–1413. https://doi.org/10.1016/j.neuropsychologia.2004.04.006

Vos, S. B., Jones, D. K., Viergever, M. A., & Leemans, A. (2011). Partial volume effect as a hidden covariate in DTI analyses. NeuroImage, 55(4), 1566–1576. https://doi.org/10.1016/j.neuroimage.2011.01.048

Vos, S. B., Tax, C. M. W., Luijten, P. R., Ourselin, S., Leemans, A., & Froeling, M. (2017). The importance of correcting for signal drift in diffusion MRI. Magnetic Resonance in Medicine, 77(1), 285–299. https://doi.org/10.1002/mrm.26124

Wang, S., Wu, E. X., Qiu, D., Leung, L. H. T., Lau, H.-F., & Khong, P.-L. (2009). Longitudinal Diffusion Tensor Magnetic Resonance Imaging Study of Radiation-Induced White Matter Damage in a Rat Model. Cancer Research, 69(3), 1190–1198. https://doi.org/10.1158/0008-5472.CAN-08-2661

Weller, M., Fisher, B., Taphoorn, M. J. B., Belanger, K., Brandes, A. A., Marosi, C., … Janzer, R. (2005). Radiotherapy plus Concomitant and Adjuvant Temozolomide for Newly Diagnosed Glioblastoma, 1, 19–28. https://doi.org/10.1056/NEJMoa043330

Zhu, T., Chapman, C. H., Tsien, C., Kim, M., Spratt, D. E., Lawrence, T. S., & Cao, Y. (2016). Effect of the Maximum Dose on White Matter Fiber Bundles Using Longitudinal Diffusion Tensor Imaging. International Journal of Radiation Oncology Biology Physics, 96(3), 696–705. https://doi.org/10.1016/j.ijrobp.2016.07.010

Supplementary material for this article can be found in the digital version of the issue.

Revolutionary neuroscientific methods through the ages

1801 Transcranial direct-current stimulation (tDCS)

1848 Phineas Gage

1875 Electroencephalography (EEG)

Microelectrodes 1940

Magnetoencephalography (MEG) 1968

Positron emission tomography (PET) 1970

1972 Computed tomography (CT)

1974 Genetic modification

1976 Patch-clamp

Magnetic resonance imaging (MRI) 1977

1980 Calcium imaging

Transcranial magnetic stimulation (TMS) 1985

1990 Two-photon excitation microscopy

1990 Functional MRI

Diffusion MRI 1992

Optogenetics 2005

Brainbow 2007

2007 Chemogenetics

Master Brain and Cognitive Sciences (MBCS) 2018

FINDING OUR WAY THROUGH THE BRAIN

by Rose Nasrawi & Iris Bosch

The evolution of neuroscientific methods

For thousands of years, humans have shown a tremendous amount of interest in the brain and its relation to the mind. Examples from Ancient Greece include the physician Hippocrates (460-370 BC), who discussed epilepsy as a disturbance of the brain, and philosopher Plato (428-348 BC), who believed the brain to be the seat of mental processes. Over the past 200 years, neuroscience has developed into a rich interdisciplinary research field, guided by the constant evolution of neuroscientific techniques. Let us take you back in time as we discuss the development of some of the most revolutionary neuroscientific techniques, and how they have contributed to our understanding of the brain.

MEASURING THE BRAIN

In order to understand how the brain functions, several methods have been developed that measure brain activity, in one way or another. Two of the most widely used methods are the electroencephalogram (EEG) and the magnetoencephalogram (MEG), first used on humans in 1929 and 1972 respectively. Synchronous post-synaptic potentials from spatially aligned pyramidal cells in the cortex create an electric and magnetic field at the skull. Changes in this electric field are measured with EEG, while changes in the magnetic field are measured with MEG. The high temporal resolution of EEG and MEG make both techniques very valuable in cognitive neuroscience: EEG/MEG signals can provide insights into dynamic brain processes responsible for specific cognitive functions on a millisecond time scale (Lopes da Silva, 2013). One of the main downsides of the techniques, especially regarding EEG, is the low spatial resolution. Electrical signals from neurons are altered as they travel through the brain tissue towards electrodes placed on the scalp for measuring electrical activity. This causes signals to spread out, making it complex to infer where each signal came from (Woodman, 2010).

Magnetic Resonance Imaging (MRI) overcomes the problem of spatial localisation. Structural MRI, performed on a human for the first time in 1977, enables neuroscientists to visualise the structure of a person's brain. MRI uses a strong magnetic field, magnetic field gradients, and radio waves to generate images of the human brain, making use of the magnetic properties of hydrogen protons. With the development of ultra-high field MRI (also referred to as 7-Tesla MRI), in vivo neuroimaging has flourished: it has enabled visualisation of the human brain in great detail, especially with regard to the sub-cortex. There are many varieties to MRI, including functional MRI (fMRI), diffusion MRI (dMRI), and quantitative MRI (qMRI). Due to its high spatial resolution, MRI has greatly contributed to our understanding of brain structure, and how this structure gives rise to cognition (Logothetis, 2009).

The development of microscopy is another great contribution to how scientists are able to image the brain and its structures. Ranging from the inventions of Antoni van Leeuwenhoek, the first light microscope to the more advanced (fluorescence) microscopes used today, microscopes have been essential for neuroscience disciplines. Of all of these inventions, the development of the two-photon excitation microscope in 1990 especially improved fluorescence microscopy. Fluorescence microscopy is widely used because of the contrast it can generate when imaging the brain, in combination with high specificity and sensitivity. As the name reveals, two-photon excitation microscopy uses two photons, instead of one, and only half the illumination energy (red instead of ultraviolet light) to excite a fluorophore. A fluorophore is a fluorescent chemical compound with the ability to re-emit light. The use of two photons has several advantages that are at the basis for the enduring popularity of the technique. First, the scatter of these relatively low energy rays causes less damage to brain tissue. Second, in order to excite the fluorophore, both photons need to reach it at the same time, meaning only that small fraction of tissue, where the photons are most concentrated, can be excited. This reduces out-of-focus background fluorescence and enables a whole new level of precision (Svoboda & Yasuda, 2006). Two-photon microscopy excels especially at three-dimensional optional sectioning and imaging live brain tissue, thereby flourishing fluorescence microscopy to a great extent.

CONTROLLING THE BRAIN

Apart from passively measuring brain activity, a lot of knowledge is gained by direct manipulation of the brain through stimulation or inhibition. The first neuroscientific method to directly stimulate the brain in a non-invasive way was transcranial direct-current stimulation (tDCS). In 1801, Giovanni Aldini successfully incorporated tDCS for the first time to improve the mood of melancholic patients. With tDCS, a constant direct current is applied to the skull via electrodes, leading to stimulation of neurons in the cortex. A similar, more recent technique is transcranial magnetic stimulation (TMS), first used by Anthony Barker around 1985. TMS acts by applying a magnetic field to the skull that causes an electric current to flow in the targeted brain region. Over the past 20 years, interest in transcranial stimulation techniques has increased, especially within clinical settings. tDCS and TMS have been shown to have beneficial effects in a wide range of diseases (e.g., epilepsy, stroke, chronic depression, and addiction). Both tDCS and TMS do not have the ability to achieve long-term influences on the brain (Stagg & Nitsche, 2011). Although this is an advantage in experimental settings, it becomes a strong disadvantage in clinical settings.

But then there was light! Since 2005, optogenetics has transformed the field of neuroscience by allowing researchers to control the signalling of specific neurons. Genetically modified viruses can be used to make ion channels of interest light-sensitive. The activity of these neurons can then be switched on and off with bursts of light (Deisseroth, 2011; Häusser, 2014). Optogenetics was soon followed by the development of chemogenetics. Chemogenetics can, through the use of chemically engineered receptors and exogenous ligands specific to that receptor, control neuronal signalling as well (Roth, 2016). Both methods enable researchers to control specific neurons within neuronal networks in vivo with a higher spatial and temporal specificity than ever. In animal experiments, flipping these neural switches has provided much knowledge on different brain networks and how they relate to behaviour and cognition. Researchers have for example been able to artificially generate a false fear-memory, using optogenetics (Ramirez, Liu, Lin, Suh, Pignatelli, Redondo et al., 2013).

Both methods have also already found their way outside of the lab. For example, optogenetics pioneer Karl Deisseroth started a company called Circuit Therapeutics in 2010, to pursue clinical trials using optogenetics to treat neurological diseases. It is only a matter of time until opto- and chemogenetics will follow techniques such as TMS to treat diseases.

FUTURE PERSPECTIVE

Over the past two decades, the field of neuroscience has evolved into a rich, broad, and interdisciplinary research field. In this brief review, we have tried to demonstrate how the development of neuroscientific techniques has contributed to this evolution. It is important to note, however, that the discussed methodologies are only the tip of the iceberg, with many neuroscientific methods remaining undiscussed. Nonetheless, it has become clear that the development of each of the discussed methods has contributed to our understanding of the brain, at different organisational levels: from single-cell recordings to connectivity estimates. It is our vision that future research in the field of neuroscience will strongly benefit from multi-modal research approaches. A clear example of this is the Human Connectome Project (HCP), which uses a multi-modal research approach to generate "a comprehensive structural description of the networks of elements and connections forming the human brain" (Sporns, 2011). The flaw of one research technique can often be the strength of another. By combining research approaches, we can obtain a more complete picture of the brain in its full glory.

REFERENCES

Deisseroth, K. (2011). Optogenetics. *Nature Methods, 8*(1), 26. DOI: 10.1038/nmeth.f.324

Häusser, M. (2014). Optogenetics: the age of light. *Nature methods, 11*(10), 1012. DOI: 10.1038/nmeth.3111

Logothetis, N. K. (2009). What we can do and what we cannot do with fMRI. *Nature, 453*, 869-878. DOI: 10.1038/nature06976

Lopes da Silva, F. (2013). EEG and MEG: Relevance to neuroscience. *Neuron, 80*(5), 1112-1128. DOI: 10.1016/j.neuron.2013.10.017

Ramirez, S., Liu, X., Lin, P. A., Suh, J., Pignatelli, M., Redondo, R. L., et al. (2013). Creating a false memory in the hippocampus. *Science, 341*(6144), 387-391. DOI: 10.1126/science.1239073

Roth, B. L. (2016). DREADDs for neuroscientists. *Neuron, 89*(4), 683-694. DOI: 10.1016/j.neuron.2016.01.040

Sporns, O. (2011). The human connectome: A complex network. *Annals of the New York Academy of Sciences, 1224*(1), 109-125. DOI: 10.1111/j.1749-6632.2010.05888.x

Stagg, C. J., & Nitsche, M. A. (2011). Physiological basis of transcranial direct current stimulation. *The Neuroscientists, 17*(1), 37-53. DOI: 10.1177/1073858410386614

Svoboda, K., & Yasuda, R. (2006). Principles of two-photon excitation microscopy and its applications to neuroscience. *Neuron, 50*(6), 823-839. DOI: 10.1016/j.neuron.2006.05.019

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials (ERPs) in studies of perception and attention. *Attention, Perception & Psychophysics, 72*(8), 2031-2046. DOI: 10.3758/APP.72.8.2031

Semantic Modulation of Visual Gamma Band Response and Auditory Steady-State Response: an EEG/MEG study

Marit Keemink
Supervisor: Dr. Olaf Hauk
MRC Cognition and Brain Sciences Unit
Cambridge, UK

ABSTRACT

The empirical literature on embodied semantics (i.e. whether sensory-motor systems contribute to semantic processing) is inconsistent when it comes to the timing and localization of effects; while fMRI, the method of choice in most positive findings in the literature, can offer high spatial resolution, it lacks the temporal precision required to distinguish between semantics and non-causal effects such as mental imagery or spreading activation. While EEG and MEG do offer high temporal resolution, physical constraints limit unambiguous localization of effects. To circumvent this issue, we investigated whether the Auditory Steady-State Response (ASSR) and Visual Gamma Band Response (VGBR), known to originate in primary auditory and visual cortex respectively, are modulated by semantics. Subjects were simultaneously presented with an annular grating inducing VGBR and modulated sine wave evoking ASSR, after which single words from different semantic categories (auditory, visual, action and abstract) were presented visually. We hypothesized an interaction between semantic category and response type (ASSR/VGBR) for auditory and visual words specifically, such that visual words would modulate VGBR differently than would auditory words, and vice versa for the modulation of ASSR.

Time-frequency analysis revealed that VGBR was strong and consistent across participants, whereas ASSR was less stable. We did not find evidence for the hypothesized cross-over effect between semantics and response type. Motivated by the disappointing ASSR power the effect of semantics on VGBR was investigated separately, which was also not significant. We did find a consistent dip in VGBR power after word onset. We conclude that auditory steady-state response and visual gamma band response are not modulated by semantics. Three possible explanations for this finding are discussed: sensory cortices are not involved in semantic processing, primary sensory cortices are not involved in semantic processing, or the neuronal populations generating steady-state and oscillatory responses within primary sensory cortices do not overlap with the neural populations involved in meaning representation.

KEYWORDS

EEG, MEG, gamma band, vision, hearing, semantics

INTRODUCTION

Embodied semantics is the theory that addresses the question of how meaning is represented in the human brain. According to this theory, knowledge representation is not amodal and abstract, but grounded in systems of perception and action (Barsalou, 2003, 2008; Fischer & Zwaan, 2008; Kiefer & Pulvermüller, 2012). The empirical literature on embodied semantics has mostly been focused around associations between sensory-motor brain areas and lexical semantic processing (Hauk & Tschentscher, 2013), with the majority of positive neuroimaging results coming from fMRI (Hauk, 2016). Using fMRI, researchers have found somatotopic activation of motor cortex during language comprehension (for a review, see e.g. Pulvermüller, 2013), as well as category-specific activation for visual (e.g. Goldberg, Perfetti, & Schneider, 2006; Simmons et al., 2007) and auditory domains (e.g. Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008). However, while fMRI can – in principle – localize effects unambiguously to specific sensory-motor areas (e.g. primary motor cortex or V1), it is not clear whether these effects reflect semantic processes, mental imagery or spreading activation (Machery, 2007; Mahon & Caramazza, 2008). The accurate timing information provided by EEG/MEG offers a potential distinction between imagery and semantics (see Hauk, 2016: 'the earlier a semantic effect occurs, the less likely it is to reflect mental imagery.'). Nonetheless, unambiguous localization of subtle effects linking semantics to primary sensory-motor areas using EEG/MEG is problematic due to the low spatial resolution of these methods: the inverse problem, calculating the magnetic or electric sources that generate the EEG/MEG signals that we measure outside the head, has no unique solution. This is because any number of distributions of current sources in the brain can result in the same measured signal. Although the problem can be constrained by making certain assumptions, such as in single-dipole fitting (Tuomisto, Hari, Katila, Poutanen, & Varpula, 1983), the spatial resolution of EEG and MEG remains limited.

The current study describes an attempt to 'work around' this trade-off between spatial and temporal resolution. We used a combined EEG/MEG approach, targeting the Visual Gamma Band Response (VGBR) (e.g. Perry, Randle, Koelewijn, Routley, & Singh, 2015) and Auditory Steady-State Response (ASSR) (e.g. Roß, Borgmann, Draganova, Roberts, & Pantev, 2000), known to originate in primary visual and auditory cortex respectively (See Box 1 for some background information about the VGBR and ASSR). Tracking these signals with EEG/MEG offers millisecond precision, whilst the nature of the response itself means that we do not have to rely on source estimation to localize it. To try to establish whether visual and auditory cortices are involved in the representation of meaning, we were interested to see whether these signals were modulated differentially by the semantic processing of auditory and visual concepts. To this end, we presented subjects with stimuli generating ASSR and VGBR simultaneously, during which single words from different semantic categories were presented. We then measured the effect of the presentation of these words on the power of the ASSR and VGBR in different time windows, reasoning that an early interaction between semantic category and brain response would provide novel evidence for the involvement of primary sensory cortices in semantics. The great temporal precision of EEG/MEG allowed to potentially distinguish between mental imagery and semantic processing, while targeting VGBR and ASSR eliminated the complication of EEG/MEG source estimation to pinpoint the effect to primary visual and auditory cortex.

Another point of discussion in the literature on embodied semantics is the degree to which the involvement of sensorimotor areas for semantics is dependent on task demands. For instance, Kiefer & Pulvermüller (2012) argued that semantic processing in the motor system happens 'early and automatically', implying that task demands should not significantly modulate these effects. Some studies, however, have found evidence for flexible and context-dependent semantic processing (Chen, Davis, Pulvermüller, & Hauk, 2015; Rogers, Hocking, Mechelli, Patterson, & Price, 2005; Van Dam, Van Dijk, Bekkering, & Rueschemeyer, 2012). This question could be addressed by comparing the modulation of brain responses in two different tasks, differing in the 'depth' of semantic processing required.

Preregistration

Some of the inconsistencies in previous literature may be explained by confirmation and publication bias (Hauk & Tschentscher, 2013). We therefore opted to pre-register our planned analysis strategy (see Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). Since the time frame of this project did not allow for peer-reviewed preregistration options, we chose to preregister with the Open Science Framework. The preregistration can be found at https://osf.io/e3djb/ (click 'View Registration Form').

METHODS

Inducing VGBR and ASSR

To induce visual gamma band responses (VGBR), participants were presented with an annular grating (concentric circles) with a spatial frequency of 3 cycles per degree of visual angle and maximal contrast, based on Perry et al. (2015) (see Figure 1, left), including a fixation dot in the centre. Based on Muthukumaraswamy & Singh (2013), we expected a maximal gamma band response between 50 and 80 Hz. To evoke an auditory steady-state response (ASSR), we presented a tone with a carrier frequency of 250 Hz modulated by a frequency of 35 Hz, with a short ramp at the onset (see Figure 1, right). This was close to the optimum modulation frequency of 40 Hz described by Roß, Borgmann, Draganova, Roberts, & Pantev (2000), but by evoking an ASSR with a slightly lower frequency we attempted to avoid overlap with the frequency range of the VGBR.

Word stimuli

The stimulus set included 60 words referring to visual concepts (e.g. 'gold') and 60 words referring to auditory concepts (e.g. 'whistle'), as well as 60 hand-related action words (e.g. 'throw') and 60 abstract words (e.g. 'law'), selected from a prior rating study employing a different set of participants. The stimuli were matched on word length, orthographic neighbourhood size, frequency of word form,

Box 1 : Background of ASSR and VGBR

Auditory Steady-State Response

The Auditory Steady-State Response is a cerebral evoked response to rapid sequences of periodic auditory stimuli, such that the successive evoked responses are overlapping. This creates a response whose frequency components stay stable over time in terms of amplitude and phase (Roß et al., 2000). In the present study this response is elicited by an amplitude-modulated tone, but it can also be evoked by tone pulses or clicks (Müller et al., 2009). Because ASSR is also present near hearing threshold levels, is easy to identify and has a relatively big amplitude, it can be used in clinical audiology to estimate hearing sensitivity (see Korczak, Smart, Delgado, Strobel, & Bradford (2012) for an overview of ASSR in the clinical practice). Because the oscillations are phase-locked to the frequency of the modulated tone, analysis can be based on the pre-determined modulation frequency. The biggest signal amplitude is found using modulation frequencies around 40 Hz (Roß et al., 2000). Multiple studies have pinpointed the ASSR to primary auditory cortex (Engelien, Schulz, Ross, Arolt, & Pantev, 2000; Pantev, Roberts, Elbert, Roß, & Wienbruch, 1996; Roß et al., 2003).

Visual Gamma-Band Response

There is considerable variability in the naming conventions of higher frequency brain oscillations; in the literature, gamma band can range from 30 to 600 Hz (Uhlhaas, Pipa, Neuenschwander, Wibral, & Singer, 2011). The visual gamma band responses we were interested in can be induced by presenting visual gratings, resulting in brain oscillations with a peak frequency between 50-80 Hz. (S.D. Muthukumaraswamy & Singh, 2013), where the amplitude of the response is highly dependent on stimulus properties, such as size (Perry et al., 2013), contrast (Hall et al., 2005), stimulus type, visual field coverage and motion ((S.D. Muthukumaraswamy & Singh, 2013). Evidence that these oscillations are generating in primary visual cortex comes from invasive animal studies (Eckhorn, Frien, Bauer, Woelbern, & Kehr, 1993; Gail, Brinksmeyer, & Eckhorn, 2000; Rols, Tallon-Baudry, Girard, Bertrand, & Bullier, 2001) and electrophysiological studies in humans (Hall et al., 2005; Hoogenboom, Schoffelen, Oostenveld, Parkes, & Fries, 2006) The VGBR has been found to be modulated by various cognitive processes, such as feature binding, attention and arousal, but relatively little is known about its functional role (Busch, Debener, Kranczioch, Engel, & Herrmann, 2004).

and unconstrained bi-/trigram frequencies using Match software (Van Casteren & Davis, 2007). This stimulus set was checked for unfamiliar and ambiguous words by two native speakers of English, after which further matching was done by hand. All words were rated on how vision-, sound-, action-related and concrete they were on a 7-point scale (see Figure 2). Paired t-tests revealed that the average ratings for words in their corresponding semantic categories (e.g. rating of vision-relatedness for words in the 'visual' category) were significantly higher than the other ratings in that category ($p < 0.01$ for all comparisons), as desired. Abstract words dropped in the matching process were used for filler trials that were excluded from any further analysis (see Trial outline).

We employed two experimental tasks, differing in the 'depth' of the semantic processing required: semantic target detection (TD) and lexical decision (LD), where we assumed the TD task to be more semantically demanding. In this task, 24 additional target words were included (10% of total), to which our participants had to respond by pressing a button with their left index finger. These were words referring to edible products containing flour and/or milk (e.g. 'cake'). In the LD task, 24 (10%) orthographically plausible but meaningless pseudowords were added, which also had to be responded to by button press. Target words and pseudowords were matched to the other stimuli on word length, word form frequency (only target words), orthographic neighbourhood size and unconstrained bi- and trigram frequency (see Table 1). A one-way ANOVA did not reveal any significant differences between the categories.

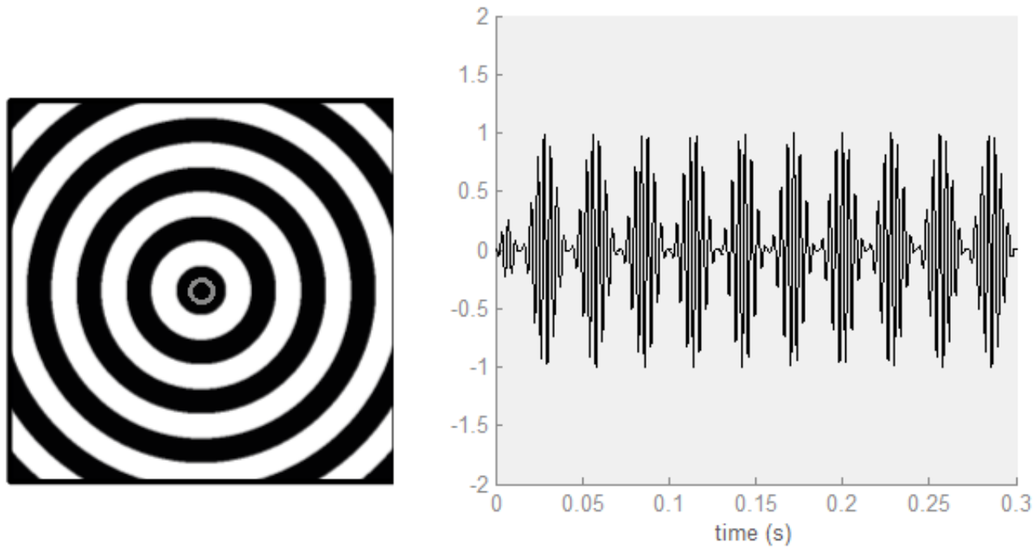


Fig. 1 **Left:** Illustration of the annular grating inducing VGBR. **Right:** Waveform of the first 300 ms of the ASSR-evoking tone.

Participants

Four pilot studies were run to test the set-up. After that, twenty healthy participants (age range: 19-40, mean age: 26.5, 12 females, 8 males) were recruited. All subjects were native speakers of English with no history of neurological, psychiatric or neurodevelopmental disorders. They all had normal or corrected-to-normal vision and hearing. Nineteen were right-handed, one ambidextrous (Oldfield, 1971). Participants were paid £30 for their time.

Procedure

After EEG setup and head digitisation (see below), participants were seated under the MEG helmet and fitted with earphones. The visual stimuli were presented through a projector outside the magnetically shielded room, the projected picture being approximately 37 by 49 cm at 129 cm distance from the helmet. The sound volume level was checked by playing the ASSR-evoking sound continuously, starting on the same volume level for each participant. Subjects were asked to indicate whether they heard the sound clearly, and whether the volume was the same in both ears. If they found the volume uncomfortably loud, it was reduced until they indicated it was at a comfortable level. Participants were instructed to use their left index or middle finger to press one button on a button box placed on their lap. Instructions were given verbally before each task, after which subjects did a practice run. They were given visual feedback after each button press during practice only. Before starting the actual tasks they were reminded of the instructions through text on the screen.

Trial outline

The annular gratings and the sounds were presented simultaneously, such that we could obtain VGBR and ASSR for identical trials. The duration of the fixation screen (grey with a fixation dot) varied slightly between trials, between 2.5 and 2.7 s, to prevent oscillatory entrainment with the rhythm of the trials. Participants were then presented with an annular grating (spatial frequency 3 Hz, contrast 100%) covering the whole screen, while simultaneously hearing the sound, for 1.7 s. After 700 ms, a word appeared in the centre of the grating (black letters in a white textbox, spanning a maximum of 1 degree of visual angle) for 150 ms. The duration of a single trial was thus between 4.2 and 4.4 s, making one task last approximately 21 minutes (excluding breaks). See Fig. 3 for a visual timeline of a single trial. Word order within the tasks was pseudorandomized. The order of the two different tasks within the session was counterbalanced.

Participants were supposed to only press a button in target trials, i.e. a target word in the target detection task and a pseudoword in the lexical decision task. Because button presses were quite rare, occurring in about 10% of trials, a filler trial was added after every button press. Participants could take a self-paced break approximately every five minutes. After each break, two filler trials were added. Filler stimuli were abstract words dropped in the matching process, which were not included in any further analysis.

After completing the two tasks (TD and LD), participants were subjected to a short localiser task in order to obtain ASSRs and VGBRs in separate trials and uncontaminated by superimposed words. In this localiser task no words were projected on top of the grating. Additionally, participants were either presented with the grating or with the sound, not simultaneously as in the previous tasks. We presented 100 trials for each condition (ASSR and VGBR). To maintain attention, 10% target trials were added in which the fixation dot slightly changed colour for 150 ms. Participants were instructed to press the button in response to the colour change.

The timing of these localiser trials was as follows: a fixation screen was presented for 1.5 to 1.7 seconds, then subjects either saw the grating or heard the sound for 1.25 s. During target trials, a slight colour change in the fixation dot would appear 600 ms after the onset of the grating/sound, lasting for 150 ms. SOA therefore varied between 2.75 and 2.95 s. Target trials and incorrect trials were excluded from further analysis.

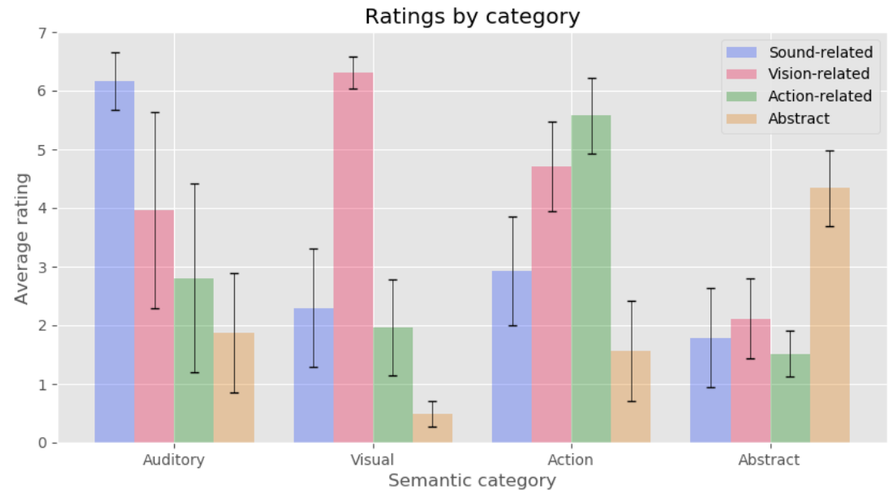


Fig. 2 Average ratings per semantic category. To make the graph easier to interpret, the 'concreteness' ratings were turned into a rating of 'abstractness' by subtracting the ratings from the maximum score of 7.

Table 1. Means and Standard Deviations per Category for Number of Letters, Word Form Frequency, Number of Orthographic Neighbours, Bi- and Trigram Frequency

Category		No. Letters	WF Frequency	Neighbours	Bigram Frequency	Trigram Frequency
Auditory	Mean	5.13	20.51	5.15	18927	1776.1
	Std. Deviation	1.142	54.46	4.943	9103.4	2213.1
Visual	Mean	4.98	23.30	4.92	17777	1425.8
	Std. Deviation	1.282	31.77	5.169	8209.3	1526.8
Action	Mean	4.92	21.73	5.70	17455	1555.1
	Std. Deviation	1.139	36.32	4.774	9469.4	2326.7
Abstract	Mean	5.42	27.77	4.45	17553	1934.6
	Std. Deviation	1.565	41.35	4.778	9030.7	1695.3
Pseudo	Mean	5.04		4.54	17317	1930.5
	Std. Deviation	.908		3.599	7131.2	1554.6
Target	Mean	5.38	11.36	4.25	16047	1604.4
	Std. Deviation	1.173	22.66	4.346	7787.8	1328.2
Total	Mean	5.13	22.24	4.94	17720	1688.6
	Std. Deviation	1.261	40.44	4.762	8694.8	1887.5

Data acquisition

Data was acquired on an Elekta Neuromag Vectorview system (Elekta AB, Stockholm, Sweden), containing 306 sensors (102 magnetometers and 204 gradiometers). EEG was acquired simultaneously from 70 electrodes mounted on an Easycap (EasyCap GmbH, Herrsching, Germany), with the recording reference electrode attached to the nose, and the ground electrode to the left cheek. The electrooculogram (EOG) was recorded from electrodes above and below the left eye (vertical EOG) and at the outer canthi (horizontal EOG). The sampling rate during data acquisition was 1000 Hz and an on-line band pass filter 0.03 to 330 Hz was applied. Prior to the EEG/MEG recording, the positions of 5 Head Position Indicator (HPI) coils attached to the EEG cap were digitised in order to monitor head position inside the MEG system. In addition, 3 anatomical landmark points (two preauricular points and nasion) as well as about 50-100 additional points that cover most of the scalp were digitised using a 3Space Isotrak II System (Polhemus, Colchester, Vermont, USA) for later co-registration with MRI data.

Data exclusion

The preregistration states that datasets with less than 50% target detection accuracy would be excluded. However, only looking at detection accuracy for target trials can be misleading when these only make up 10% of all trials: false alarms, responses when there is no target present, would be ignored. Therefore, we opted to express response accuracy in d' (d prime), which is the z-transform of the hit rate (the probability of a response when a target is present) minus the z-transform of the false alarm rate (the probability of a response when no target is present): d^'=z(H)-z(FA). High d' scores indicate high discrimination ability, whereas a d' near zero indicates performance at chance. Since d' > 1 for all participants in all three tasks, no datasets were excluded.

Pre-processing

First, data were subjected to spatio-temporal signal-space separation (SSS) implemented in the Maxfilter software (Version 2.2.12) of Elekta Neuromag to remove noise generated from sources distant to the sensor array (Taulu & Kajola, 2005; Taulu & Simola, 2006). The SSS procedure included movement compensation (locations recorded every 200 ms), bad channel interpolation, and temporal SSS extension (with default buffer length 10 s and sub-space correlation limit 0.98). The origin in the head frame was chosen as (0,0,45) mm.

The following steps of analysis were performed in the MNE-Python software package (Version 0.16) (Gramfort et al. 2014; Gramfort et al. 2013). Raw data were visually inspected, and consistently bad EEG channels were marked and interpolated (in 'accurate' mode). After interpolation, the average-reference operator was applied, as well as a notch filter at 50 and 100 Hz (filter length 6600 samples). Data were then FIR band-pass filtered between 0.1 and 100 Hz using default settings (filter length 66000 samples, low and high band widths 0.1 and 25 Hz, respectively). Independent Component Analysis (using the FastICA algorithm, Hyvärinen & Oja (2000)) was applied to the filtered data in order to remove eye movement artefacts, for those subjects where data were significantly contaminated by eye movements (judged by trial rejection rates due to EOG channels or frontal EEG channels during averaging and visual inspection), which was the case for all but one subject. The ICA procedure provided for the MNE-Python software uses the temporal correlation between ICA components and EOG channels as a rejection criterion. The success of the ICA procedure was judged by its effect on evoked responses averaged across all epochs.

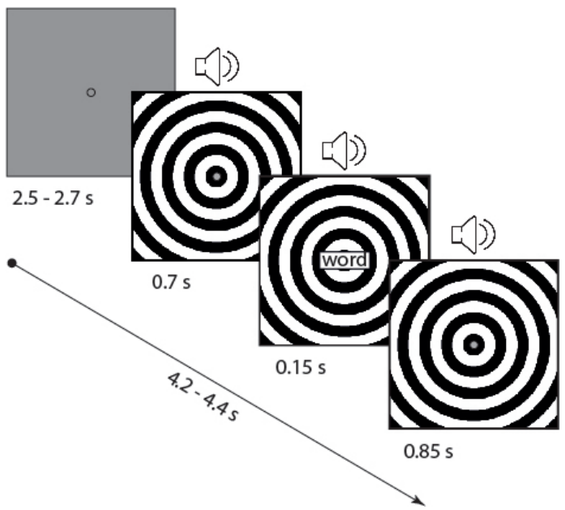


Fig. 3 Timeline of trials in the target detection (TD) and lexical decision (LD) tasks. Durations of stimuli are provided below the images.

Data were divided into epochs from -500 ms to 1200 ms around the onsets of the visual gratings and steady-state sound stimuli. As outlined in the preregistration, we were planning to apply the new automated artefact rejection algorithm "Autoreject" (as implemented in MNE-Python) (Jas, Engemann, Bekhti, Raimondo, & Gramfort, 2017). However, this proved to be too time-consuming to implement within the scope of this project. Therefore, epochs were rejected by using peak-to-peak amplitude thresholds (EEG: > 300 µV, gradiometers: > 100 pT, magnetometers: > 5 pT). For two subjects, strong ECG signals showing in the EEG/MEG data led to very high rejection rates. For these subjects, the thresholds were multiplied with a factor 1.5. The quality of the resulting data was judged on the evoked response averaged across epochs. Trials with incorrect behavioural responses were excluded from further analysis.

Time-Frequency Analysis

Time-frequency analysis was performed using Morlet wavelets between 30 and 100 Hz, with a frequency resolution of 1 Hz and number of cycles corresponding to half the wavelet frequency.

Peak channels for further analysis were determined from the localiser scan, separately for ASSR and VGBR as well as for different channel types (magnetometers, gradiometers and EEG). First, the ratio of power values from the wavelet analysis with respect to pre-stimulus baseline was computed. For ASSR, those channels that showed maximum power ratios for 35 Hz wavelets within the time window 350-1000 ms were selected. The first 350 ms were omitted in order to avoid contamination from initial evoked responses. For VGBR, we planned to first determine the maximum power ratio across all channels, frequencies between 30 and 100 Hz, and latencies from 350 to 1000 ms. However, a large portion of EEG data contained high-frequency artefacts (probably muscular activity) in frontal electrodes. Since EEG is more sensitive to high-frequency muscle artefacts compared to MEG (Suresh D. Muthukumaraswamy, 2013), we chose to limit the possible VGBR peak channels to posterior channels in EEG. Also, we narrowed the frequency band to 30-70 Hz, based on Perry, Hamandi, Brindley, Muthukumaraswamy, & Singh (2013). We then determined peak power across channels for the same latency window, averaged across frequencies +/- 5 Hz around the previously determined peak frequency. We used five peak channels per channel type.

Statistical Analysis

Average percent change in power with respect to 300 ms preword onset baseline was computed across time for peak channels and frequencies, as described in the section Time-Frequency Anal-

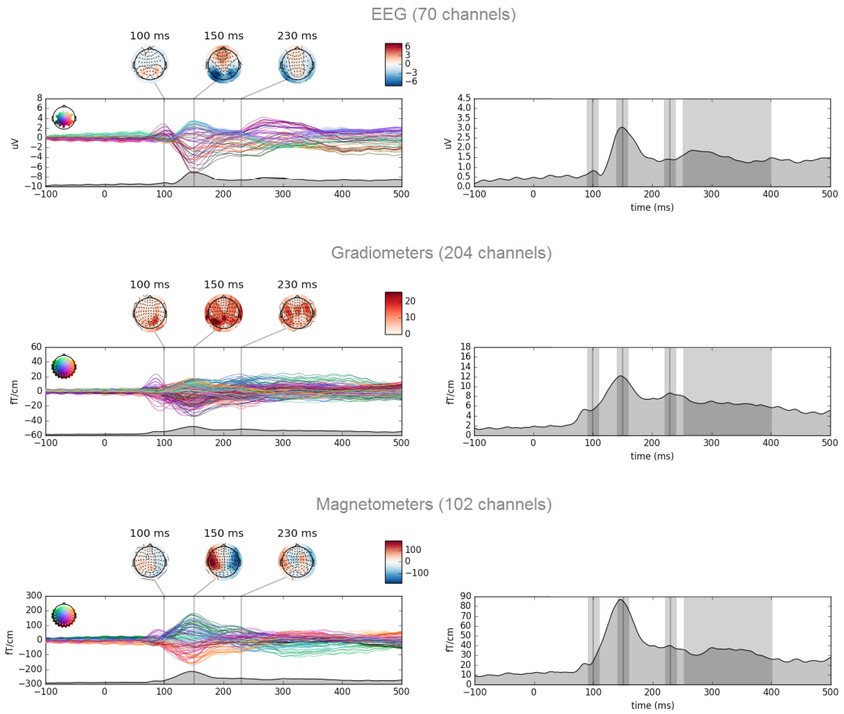


Fig. 4 **Left:** Evoked responses following word onset at 0 ms per channel type, with topographies at peak latencies. Global field power is plotted in grey at the bottom of each plot. **Right:** Global field power of the evoked responses per channel type (magnification of the GFP in the left plots). The time windows chosen for the statistical analysis are 20 ms around peak latencies, plus a longer window from 250 - 400 ms capturing the N400.

ysis. We ran an analysis focused on average amplitudes within pre-specified latency ranges. These latency ranges were defined similar to previous studies as 20 ms windows around peaks in the root-mean square (global field power) of the original word-evoked responses, at 100 ms (P1), 150 ms (N1), and 230 ms (N2) after word onset. The N1 peak occurred slightly earlier than the expected 170 ms mentioned in the preregistration, but within the typical peak latency range for this component (Callaway & Halliday, 1982). A longer time window capturing the N400 window was chosen between 250 and 400 ms. See Figure 4. These word-evoked responses are slightly different from conventional ERPs in that the presentation of the words overlapped with the visual and auditory stimulation evoking steady-state and visual gamma responses.

For every latency window, we obtained one value per word category (words related to visual and auditory concepts, respectively) and channel group (peak channels from ASSR and VGBR, respectively) per condition and subject. These values were subjected to a 2-by-2 ANOVA with factors Word Category and Channel Group. Based on two recent MEG studies (Mollo, Pulvermuller, & Hauk, 2016; Moseley, Pulvermuller, & Shtyrov, 2013), we hypothesised that the latency window around 170 ms should be the earliest sensitive to word semantics. These analyses were run for the three sensor types separately.

RESULTS

ASSR and VGBR

Average evoked responses to the sound or grating in the localizer scan show that the stimuli were presented and perceived correctly (see Figure 5). Note that, since the steady-state response is phase-locked to the stimulus, the 35 Hz response can be seen quite clearly in the average evoked response of the ASSR trials.

Time-frequency analysis of ASSR and VGBR in the localiser data (as described in Methods) revealed a relatively consistent VGBR across participants and channel types, with peak frequencies ranging from 47 to 61 Hz. The ASSR, however, proved more difficult to find. Although clearly present in time-frequency plots of the average power across subjects, (see Figure 6), it seemed unstable or even absent in a significant portion of subjects when plotted individually. Also, topographies showed a lateralized effect in some subjects.

This result was surprising, since pilot studies did reveal an ASSR for individual subjects within the current paradigm. Given the short timeframe of this project and the therefore limited options to explore the reasons and possible solutions for the disappointing ASSR strength and stability, the data presented here should be regarded as preliminary.

Comparing conditions

As outlined in Methods, we looked at the relative change in ASSR and VGBR power after word onset, comparing the average power change in the visual and auditory condition for peak channels in four time windows. Figure 7 shows the percent change in power over the entire trial window with respect to a pre-word onset baseline.

Between grating/sound and word onset, we would not expect to see any difference between conditions: all trials were identical up to that point. Where the VGBR shows very little variability between conditions, the ASSR power already shows relatively large differences between conditions before word onset. This confirms previous observations about the strength and stability of the ASSR in our data. Figure 8 shows the average response in the auditory (blue lines) and visual condition (red lines) for every participant. Whereas the VGBR looks very stable across participants, especially for gradiometers, the ASSR does not. Surprisingly, for some subjects ASSR power even seems to go down after the onset of the sound, which of course should not be the case.

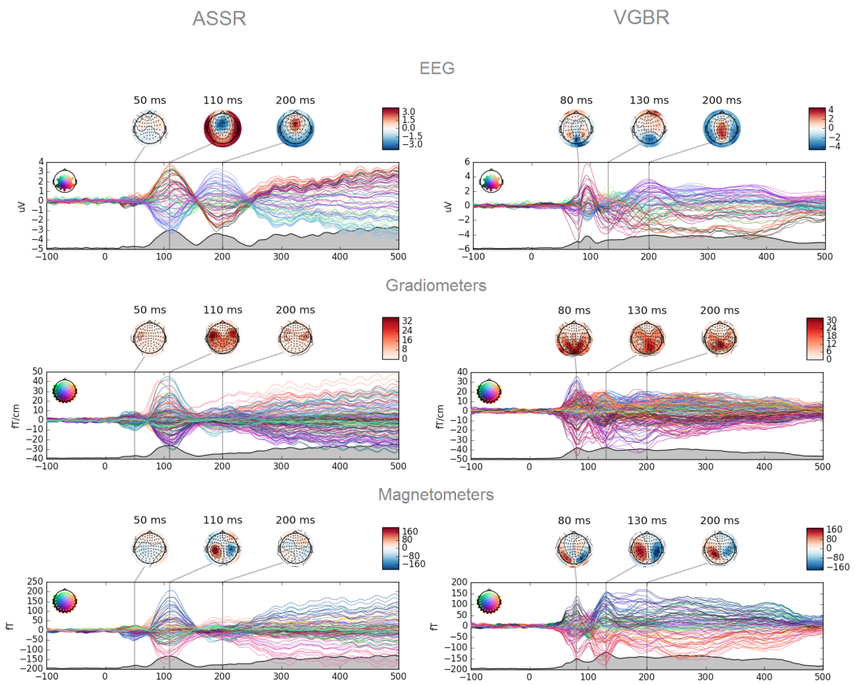


Fig. 5 Average evoked responses to the onset of the sound (ASSR) or grating (VGBR) in the localizer scan, plotted for each sensor type separately. Topographies are shown for peak latencies.

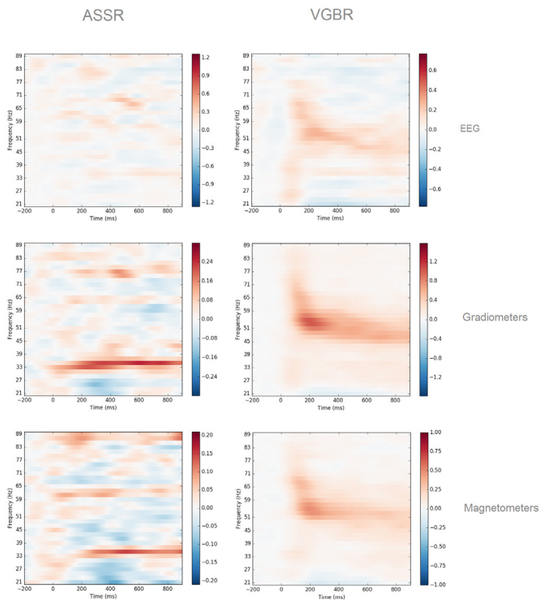


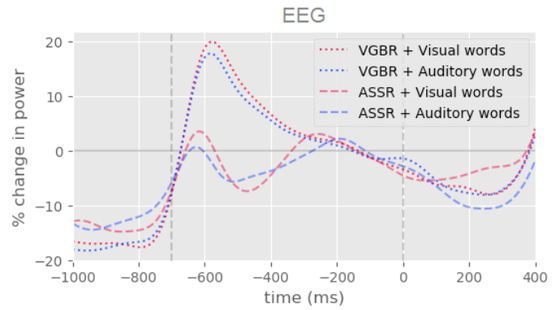
Fig. 6 Average evoked responses to the onset of the sound (ASSR) or grating (VGBR) in the localizer scan, plotted for each sensor type separately. Topographies are shown for peak latencies.

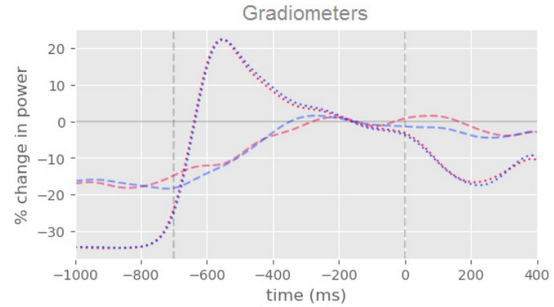
We hypothesized that there would be an interaction between Word Category and Response type (ASSR or VGBR) after word onset, where the strongest evidence for embodied semantics would be an effect in early time windows. From Figure 7 on the previous page, it already becomes apparent that there is no cross-over interaction in our data: the average change from baseline in one condition would have to be higher in one response type and lower in the other after word onset. In other words, the red line would have to dip below the blue line in one response type, and stay above the blue line in the other response type (we did not have a clear hypothesis about the direction of the effect, so whether the congruent conditions would lead to a smaller or larger change from baseline). This was confirmed by a series of 2 x 2 Anovas after checking for normality: there was no significant interaction between Word Category and Response Type in any time window for the three channel types (EEG, gradiometers or magnetometers).

Data from the lexical decision and target detection task were combined to test our primary hypothesis. We then investigated whether the absence of an effect was due to an interaction being present in one paradigm but not the other, by repeating the tests for both tasks separately. Again, there was no significant interaction between Word Category and Response Type in any time window for EEG, gradiometers and magnetometers.

Given the unreliability of the ASSR we also looked at the modulation of VGBR only, leaving ASSR out of the Anova. The hypothesis that visual cortex is involved in semantic processing of visual concepts could still be tested by investigating whether VGBR is modulated differently by visual words compared to other semantic categories. Figure 9 shows kernel density plots, showing the distribution of VGBR power change across participants for every word category. This is plotted for the four time windows and the different channel types separately. As the highly overlapping curves suggest, Welch's t-tests for every channel type and time window did not reveal any significant differences between visual words and auditory, action and abstract words (treated as one group).

These plots do show that there is a consistent dip in VGBR power after word onset across participants and channel types, meaning





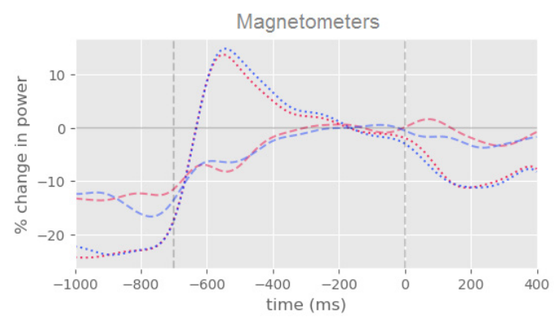


Fig. 7 Relative change in power for ASSR (dashed lines) and VGBR (dotted lines) in peak channels, normalized to a -300 ms to word onset baseline. The dotted line at -700 ms indicates grating/sound onset, word onset is at 0. Red and blue lines indicate the average response to visual and auditory words, respectively. Different panels correspond to different channel types.

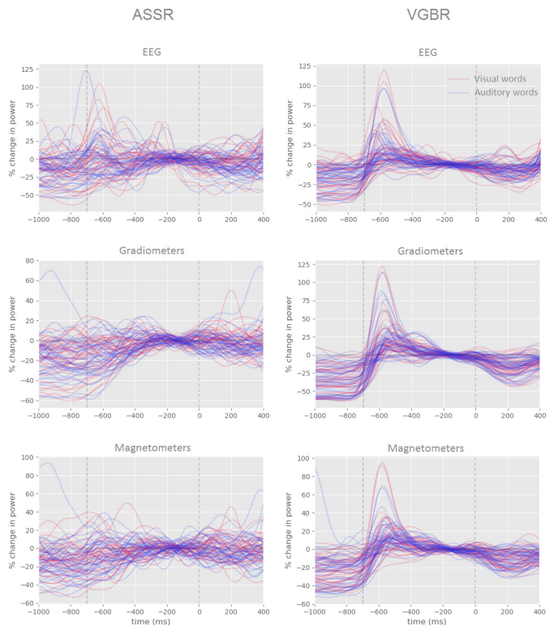


Fig. 8. Average responses per participant in the visual (in red) and auditory condition (in blue), baseline corrected from -300 ms to word onset.

that projecting a word on top of the annular grating decreases VGBR strength.

DISCUSSION

In this study we investigated whether auditory steady-state response (ASSR) and visual gamma band response (VGBR) are modulated by the semantic processing of single words. Employing a novel paradigm, we presented ASSR- and VGBR-evoking stimuli simultaneously before superimposing words from different semantic categories on the visual stimulus. We hypothesized that the processing of visual concepts would modulate VGBR power differently than would auditory concepts, and we expected the opposite effect for the ASSR. The data did not show this expected interaction, nor did they show an effect of semantic category on VGBR modulation. Although the visual gamma band response to our experimental stimuli was strong across participants, the ASSR did not show the same consistency. Since testing the hypothesis is reliant on a stable baseline response, results involving the modulation of ASSR should be regarded as preliminary.

Evoking ASSR

The absence of a stable ASSR across participants is interesting in itself. The robustness of the auditory steady-state response is what makes it useful for clinical purposes (see Box 1), so it is surprising that we failed to find it in all participants.

In many studies involving ASSR stimulus duration is longer than in our set-up (e.g. 200 seconds in Roß, Draganova, Picton, & Pantev (2003) and Roß et al., (2000)). Although other studies successfully investigated ASSR elicited by tones as short as 780 ms (Kuriki, Kobayashi, Kobayashi, Tanaka, & Uchikawa, 2013) and 800 ms (Müller, Schlee, Hartmann, Lorenz, & Weisz, 2009), these authors do not show whether the response was consistent across participants. Though, given the nature of a steady-state response being a series of overlapping evoked responses, the amplitude of the response remains stable over time (Roß et al., 2000), meaning that, in our study, presenting the sound for a longer time before presenting the word probably would not have made any difference in power pre-word-onset.

ASSR power can be modulated by selective attention (Bidet-Caulet et al., 2007; Müller et al., 2009), especially in the 40 Hz range (Skosnik, Krishnan, & O'Donnell, 2007). These studies showed an enhancement of the response for attended stimuli in one ear compared to unattended information in the other ear. Additionally, Müller et al., (2009) showed that responses to unattended stimuli were suppressed. However, both attended and unattended stimuli were in the same modality in these experiments. This is a big difference with our design: participants had to focus their attention to the visual stimuli in order to perform the tasks, while they could ignore the auditory domain completely. This might explain why we found a stronger VGBR than ASSR, although, to my knowledge, these two responses have not been compared directly in previous studies. Roß, Picton, Herdman, & Pantev (2004) do report largely enhanced ASSR responses in an auditory task compared to a visual task, but they do not specify power in the unattended condition. The previously mentioned studies also do not report effect size, so it remains unclear whether attentional effects can explain this result.

One issue we encountered during testing was a lack of control over how loudly participants heard the sound stimuli. The position of the earphones inside the ear proved to have a large effect on how loudly the sound was heard. Although measures were taken to ensure a good and steady fit in each participant (trying earphones in different sizes, sometimes taping them to the ear to prevent falling out), we had to rely on subjective reporting about the sound volume. Auditory evoked responses were present in all subjects, indicating that they at least heard something, but it is possible that there was still a significant difference in how loudly participants heard the tone, or that the earphones moved during the recording. This could explain the lateralized response we saw in some participants, and it potentially adds up to the attentional effects described above to explain the absence of an ASSR.

Semantic modulation

We did not find evidence for semantic modulation of VGBR and ASSR. We did find that briefly superimposing a word on the annular grating caused a decrease in VGBR power. The amplitude of the visual gamma response is highly dependent on stimulus properties (e.g. Perry et al. (2013)), but this effect of partially and briefly covering the stimulus has, to my knowledge, not been demonstrated before.

One of the motivations for this study was to potentially obtain a timeline of the involvement of sensory areas in semantics. As outlined in the introduction, this temporal information is essential to distinguish between semantics, imagery or spreading activation by association. However, we did not find an effect in any time

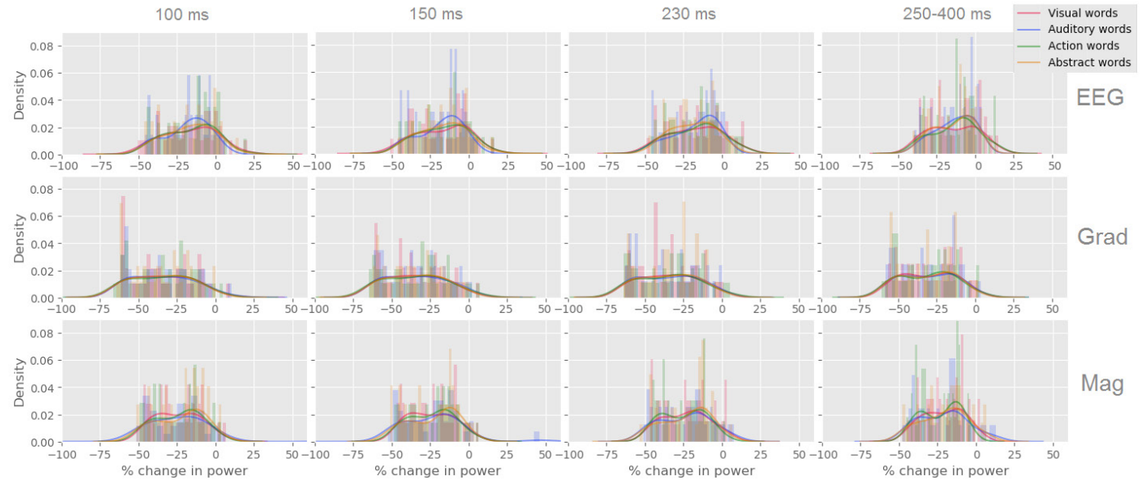


Fig. 9. Kernel density plots showing the distribution of VGBR power change across participants in the different time windows. Channel types (EEG, gradiometers and magnetometers) are plotted separately.

window. There are multiple possible explanations for this finding. Firstly, it could be that sensory areas are not involved in semantic representations or processes at all. As outlined in the introduction, it remains unclear whether previous findings linking sensory-motor areas to lexical semantics actually reflect the processing of meaning or something else, such as mental imagery or spreading activation. However, even if sensory cortices are not involved in semantic processing, we might still expect to find this non-causal activation in later windows, for instance the 250-400 ms time window.

The total absence of an effect leaves room for other interpretations: a second possibility is that meaning representation involves sensory areas, but not the primary visual and auditory cortices in which VGBR and ASSR originate in particular. This raises the question of what 'embodiment' really means, and down to what hierarchical level of processing one would expect sensory-motor systems to be involved in semantic processes; an extreme take on embodiment might suggest that the processing of visual concepts would involve activating the retina, as explained in Hauk & Tschentscher (2013). These authors point out that if one rejects this idea, there is no real theoretical reason to stop at primary motor- or sensory cortices either.

Lastly, semantic processing might involve primary sensory areas, but the neuronal populations within these areas generating steady-state responses and gamma oscillations might not overlap with neuronal populations that would be involved in semantics. Although the relation between frequency and amplitude of the VGBR and various stimulus properties has been studied extensively (see Box 1), any hypothesis about the functional role of gamma band oscillations in visual cortex remains relatively low-level: for instance, Perry et al. (2013) suggest that visual gamma activity reflects GABAergic inhibitory processes responsible for suppressing the surrounds of the receptive field. Again, this raises the question at what level we start talking about 'embodiment'; how specific is the visual information that one activates when processing the word 'pearl'? Is it specific enough to include information about visual field coverage, and if it does, does the mental simulation activate surround-suppression, regulated by inhibitory neurotransmitters? These dilemmas regarding the definition of 'embodiment' have not been resolved as of yet, and they will require solid, sophisticated empirical research to entangle.

Acknowledgements

I would like to thank Olaf Hauk for his excellent supervision and for tirelessly guiding me through the treacherous (python-filled) forest that is EEG/MEG analysis. Thanks to Lara Bridge and Alicia Smith for checking the stimulus set (and being wonderful office mates), and to Rezvan Farahibozorg for helping me assemble said stimulus set. Scripts to generate the VGBR- and ASSR-evoking stimuli were kindly provided by Gavin Perry and Srivas Chennu.

My stay at the MRC-CBU was funded by Erasmus+, Dr. Hendrik Mullerfonds, Stichting Jo Kolk Studiefonds and Fundatie van Renswoude, for which I'm very grateful.

REFERENCES

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. Language and Cognitive Processes, 18(516), 513–562. Retrieved from http://userwww.service.emory.edu/-harsalou/.

Barsalou, L. W. (2008). Grounded Cognition. Annu. Rev. Psychol, 59, 617–45. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P.-E., Giard, M.-H., & Bertrand, O. (2007). Effects of Selective Attention on the Electrophysiological Representation of Concurrent Sounds in the Human Auditory Cortex. Journal of Neuroscience, 27(35), 9252–9261. https://doi.org/10.1523/JNEUROSCI.1402-07.2007

Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., & Herrmann, C. S. (2004). Size matters: effects of stimulus size, duration and eccentricity on the visual gamma-band response. Clinical Neurophysiology, 115(8), 1810–1820. https://doi.org/10.1016/J.CLINPH.2004.03.015

Callaway, E., & Halliday, R. (1982). The Effect of Attentional Effort on Visual Evoked Potential N1 Latency. Psychiatry Research, 7, 299–308. Retrieved from http://www.psy-journal.com/article/0165-1781(82)90066-X/pdf

Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2015). Early Visual Word Processing Is Flexible: Evidence from Spatiotemporal Brain Dynamics. Journal of Cognitive Neuroscience, 27(9), 1738–1751. https://doi.org/10.1162/jocn_a_00815

Eckhorn, R., Frien, A., Bauer, R., Woelbern, T., & Kehr, H. (1993). High frequency (60–90 Hz) oscillations in primary visual cortex of awake monkey. NeuroReport, 4(3), 243–246. https://doi.org/10.1097/00001756-199303000-00004

Engelien, A., Schulz, M., Ross, B., Arolt, V., & Pantev, C. (2000). A combined functional in

vivo measure for primary and secondary auditory cortices. Hearing Research, 148(1–2), 153–160. https://doi.org/10.1016/S0378-5955(00)00148-9

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. The Quarterly Journal of Experimental Psychology, 61(6), 825–850. https://doi.org/10.1080/17470210701623605

Gail, A., Brinksmeyer, H. J., & Eckhorn, R. (2000). Contour Decouples Gamma Activity Across Texture Representation in Monkey Striate Cortex. Cerebral Cortex, 10(9), 840–850. https://doi.org/10.1093/cercor/10.9.840

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual Knowledge Retrieval Activates Sensory Brain Regions. The Journal of Neuroscience, 26(18), 4917–21. https://doi.org/10.1523/JNEUROSCI.5389-05.2006

Hall, S. D., Holliday, I. E., Hillebrand, A., Singh, K. D., Furlong, P. L., Hadjipapas, A., & Barnes, G. R. (2005). The missing link: analogous human and primate cortical gamma oscillations. https://doi.org/10.1016/j.neuroimage.2005.01.009

Hauk, O. (2016). Only time will tell – why temporal information is essential for our neuroscientific understanding of semantics. Psychonomic Bulletin & Review, 23(4), 1072–1079. https://doi.org/10.3758/s13423-015-0873-9

Hauk, O., & Tschentscher, N. (2013). The body of evidence: what can neuroscience tell us about embodied semantics? Frontiers in Psychology, 4(50). https://doi.org/10.3389/fpsyg.2013.00050

Hoogenboom, N., Schoffelen, J.-M., Oostenveld, R., Parkes, L. M., & Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. NeuroImage, 29(3), 764–773. https://doi.org/10.1016/J.NEUROIMAGE.2005.08.043

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. Neural Networks, 13(4–5), 411–430. https://doi.org/10.1016/S0893-6080(00)00026-5

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. Cortex, 48(7), 805–825. https://doi.org/10.1016/j.cortex.2011.04.006

Kiefer, M., Sim, E.-J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. Journal of Neuroscience, 28(47), 12224–30. https://doi.org/10.1523/JNEUROSCI.3579-08.2008

Korczak, P., Smart, J., Delgado, R., Strobel, T. M., & Bradford, C. (2012). Auditory Steady-State Responses. Journal of the American Academy of Audiology, 23(3), 146–170. https://doi.org/10.3766/jaaa.23.3.3

Kuriki, S., Kobayashi, Y., Kobayashi, T., Tanaka, K., & Uchikawa, Y. (2013). Steady-state MEG responses elicited by a sequence of amplitude-modulated short tones of different carrier frequencies. Hearing Research, 296, 25–35. https://doi.org/10.1016/J.HEARES.2012.11.002

Machery, E. (2007). Concept empiricism: A methodological critique. Cognition, 104, 19–46. https://doi.org/10.1016/j.cognition.2006.05.002

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. Journal of Physiology Paris, 102(1–3), 59–70. https://doi.org/10.1016/j.jphysparis.2008.03.004

Müller, N., Schlee, W., Hartmann, T., Lorenz, I., & Weisz, N. (2009). Top-down modulation of the auditory steady-state response in a task-switch paradigm. Frontiers in Human Neuroscience, 3, 1. https://doi.org/10.3389/neuro.09.001.2009

Muthukumaraswamy, S. D. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. Frontiers in Human Neuroscience, 7, 138. https://doi.org/10.3389/fnhum.2013.00138

Muthukumaraswamy, S. D., & Singh, K. D. (2013). Visual gamma oscillations: The effects of stimulus type, visual field coverage and stimulus motion on MEG and EEG recordings. NeuroImage, 69, 223–230. https://doi.org/10.1016/J.NEUROIMAGE.2012.12.038

Oldfield, R. C. (1971). The Assessment and Analysis of Handedness: The Edinburgh Inventory. Neuropsychologia, 9, 97–113. Retrieved from https://pdfs.semanticscholar.org/75c2/f465d59739dbc06b70fd34dc3c1b2336103e.pdf

Pantev, C., Roberts, L. E., Elbert, T., Roß, B., & Wienbruch, C. (1996). Tonotopic organization of the sources of human auditory steady-state responses. Hearing Research, 101(1–2), 62–74. https://doi.org/10.1016/S0378-5955(96)00133-5

Perry, G., Hamandi, K., Brindley, L. M., Muthukumaraswamy, S. D., & Singh, K. D. (2013). The properties of induced gamma oscillations in human visual cortex show individual variability in their dependence on stimulus size. NeuroImage, 68, 83–92. https://doi.org/10.1016/J.NEUROIMAGE.2012.11.043

Perry, G., Randle, J. M., Koelewijn, L., Routley, B. C., & Singh, K. D. (2015). Linear Tuning of Gamma Amplitude and Frequency to Luminance Contrast: Evidence from a Continuous Mapping Paradigm. PLOS ONE, 10(4), e0124798. https://doi.org/10.1371/journal.pone.0124798

Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends in Cognitive Sciences, 17(9), 458–470. https://doi.org/10.1016/J.TICS.2013.06.004

Rogers, T. T., Hocking, J., Mechelli, A., Patterson, K., & Price, C. (2005). Fusiform Activation to Animals is Driven by the Process, Not the Stimulus. Journal of Cognitive Neuroscience, 17(3), 434–445. https://doi.org/10.1162/0898929053279531

Rols, G., Tallon-Baudry, C., Girard, P., Bertrand, O., & Bullier, J. (2001). Cortical mapping of gamma oscillations in areas V1 and V4 of the macaque monkey. Visual Neuroscience, 18, 527–540. Retrieved from https://www.cambridge.org/core/services/aop-cambridge-core/content/view/5DDF635A8F0C5C01934E7B244CD873AB/S0952523801184038a.pdf/cortical-mapping-of-gamma-oscillations-in-areas-v1-and-v4-of-the-macaque-monkey.pdf

Roß, B., Borgmann, C., Draganova, R., Roberts, L. E., & Pantev, C. (2000). A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. The Journal of the Acoustical Society of America, 108(2), 679–691. Retrieved from http://hnplab.mcmaster.ca/publications/JASA_2000.pdf

Roß, B., Draganova, R., Picton, T. W., & Pantev, C. (2003). Frequency specificity of 40-Hz auditory steady-state responses. Hearing Research, 186(1–2), 57–68. https://doi.org/10.1016/S0378-5955(03)00299-5

Roß, B., Picton, T. W., Herdman, A. T., & Pantev, C. (2004). The effect of attention on the auditory steady-state response. Neurology & Clinical Neurophysiology, 2004, 22. Retrieved from http://europepmc.org/abstract/MED/16012602

Simmons, W. K., Ramjee, V., Beauchamp, M. S., Mcrae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. Neuropsychologia, 45, 2802–2810. https://doi.org/10.1016/j.neuropsychologia.2007.05.002

Skosnik, P. D., Krishnan, G. P., & O'Donnell, B. F. (2007). The effect of selective attention on the gamma-band auditory steady-state response. Neuroscience Letters, 420, 223–228. https://doi.org/10.1016/j.neulet.2007.04.072

Tuomisto, T., Hari, R., Katila, T., Poutanen, T., & Varpula, T. (1983). Studies of Auditory Evoked Magnetic and Electric Responses: Modality Specificity and Modelling. Il Nuovo Cimento, 2D(2), 471–494. Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF02455946.pdf

Uhlhaas, P. J., Pipa, G., Neuenschwander, S., Wibral, M., & Singer, W. (2011). A new look at gamma? High- (> 60 Hz) γ-band activity in cortical networks: Function, mechanisms and impairment. Progress in Biophysics and Molecular Biology, 105(1–2), 14–28. https://doi.org/10.1016/J.PBIOMOLBIO.2010.10.004

Van Dam, W. O., Van Dijk, M., Bekkering, H., & Rueschemeyer, S.-A. (2012). Flexibility in Embodied Lexical-Semantic Representations. Human Brain Mapping, 33, 2322–2333. https://doi.org/10.1002/hbm.21365

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. Perspectives on Psychological Science, 7(6), 632–638. https://doi.org/10.1177/1745691612463078

Unsupervised scene and place recognition based on features extracted from pretrained convolutional neural networks

Andreas Sebastian Wolters

Supervisor: Moritz Milde,

Neuromorphic Cognitive Systems Group,

Institute of Neuroinformatics, Zürich

ABSTRACT

It is a trivial task for humans to tell apart instances of distinct visual scenes (scene recognition) or to state whether one has been to a certain place before (place recognition). Previous research has shown that convolutional neural networks, or the feature activity thereof, allow to perform adequately on these tasks, under the condition that labelled data was available (Zhou et al., 2014); place recognition was shown to work well with convolutional neural network features even in absence of labelled training data (Chen et al., 2014). In this study an algorithmic approach is established that uses features extracted from pretrained convolutional neural networks and analyses these further, firstly with a matching algorithm to recognise places, secondly with a k-means clustering algorithm to recognise scenes. This approach neither requires retraining the network, nor the administration of labelled data. For the best-performing layers, scenes can be recognised with an accuracy of 93.7% and 88.71% for the AlexNet and VGG16 networks, respectively. Place allocation performance, again for the best-performing layers, reached 83.37% and 99.73% for these networks. In essence, this study shows that scenes can be distinguished, without any labels given, with accuracy levels that are well above what is expected to occur by chance. This is to the best of our knowledge – the first instance of unsupervised scene recognition and provides new insights on how to potentially achieve an efficient solution to simultaneous place and scene recognition in neuromorphic systems.

1. INTRODUCTION

When placed in a novel visual environment it is a trivial task for a human to make two high-level assessments. Firstly, is this a place that I have been to before (*place recognition*)? Secondly, what type of scene am I in – am I in an urban scene, an indoor environment or rather a countryside environment (*scene recognition*)? That place recognition is trivial to humans is based on anecdotal observations – it is a crucial component of everyday navigation – as human performance data on place recognition tasks is unavailable (Frampton & Calway, 2013); for scene recognition it has been shown that humans consistently outperform all available algorithms (Borji & Itti, 2014). The synthesis of place and scene recognition performance into capable algorithms comes with significant difficulties, partly because the mechanisms behind both place recognition (Lowry et al., 2016) and scene recognition (Sharma & Tripp, 2016) in the mammalian visual system are not yet understood.

The synthesis of visual scene and place recognition capabilities has usually been regarded as a task that falls within computer vision, a field that aims to build algorithms that extract relevant information from raw image data (Szeliski, 2011). Many of these algorithms have in common that they employ a first step of feature selection instead of using the whole input signal. The idea behind this process is that using such *"characteristic features of the signals – rather than the signals themselves – [...] improves performance"* (Wiatowski & Bölcskei, 2015, p. 2) and reduces computational demands (Hira & Gillies, 2015). Much of the earlier successful work in machine learning in general, and computer vision in specific, was achieved through approaches that rely on engineered features, e.g. features generated by the scale-invariant feature transform algorithm (Lowe, 1999). These features, describing local visual features, have been applied to a variety of tasks, such as gesture recognition (Wang & Wang, 2007), object recognition (Lowe, 1999) and robot navigation (Se et al., 2011).

Trainable artificial neural networks (ANNs), i.e. algorithms that

were fundamentally inspired by principles of neural computation, had been around since the 1950s (Rosenblatt, 1957). It was, however, only with the recent success of the convolutional neural network (CNN) on object recognition challenges like ImageNet (Jia Deng et al., 2009) that the promise of neural networks as highly-capable computational entities was confirmed (Krizhevsky et al., 2012). Interestingly, though, a recent study has shown that the performance of a CNN can be consistently, and significantly, improved if the layer that computes the classification output is replaced by a linear classifer such as a support vector machine (Tang, 2013). This has led to the statement that the impressive performance of these networks on complex tasks – even recently outperforming humans (He et al., 2016) – is likely to be the result of the superiority of the learned features rather than optimality in the inference process.

This finding has led to the suggestion that CNN features might be appropriate to be transferred to tasks that the network was not initially trained on (Athiwaratkun & Kang, 2015). Attempting to leverage ANNs for other tasks is called *transfer learning*. There are two different approaches to transfer learning, besides fully training a network from scratch: fine-tuning and feature extraction (Nogueira et al., 2016). Fine-tuning entails to retrain the task-specific layers, commonly the final fully-connected ones, without adjusting earlier layers (Yosinski et al., 2014). Treating a CNN as a feature extractor means that the feature activations of a CNN in response to a given input are extracted and further processed by different means (Wiatowski & B¨olcskei, 2015). Whilst fine-tuning requires labelled data, using a pretrained CNN as a feature extractor allows CNN features, which are trained with labelled data, to be used in contexts where labelled data is sparse, or unavailable (Nogueira et al., 2016).

The aim of this study is to assesses whether there exists an algorithm that is capable to simultaneously determine whether new visual information represents a known place and to which class of previously-presented scenes it belongs. The problem of recognising whether a place has been previously seen is integral to autonomous navigation; it is akin to the problem of loop closure in the simultaneous localisation and mapping approach to navigation, which describes the aim to recognise a previously-encountered place from a different perspective to then update the internal map representation, i.e. to close the loop of one's memorised path (Ho & Newman, 2006). A recent study has shown that CNN features extracted from Overfeat, a CNN trained on ImageNet, are appropriate for loop closure when further analysed by a matching algorithm; the results showed an immense improvement over features that were generated by a generative model (Chen et al., 2014). Data on human performance is – to the best of our knowledge – unavailable (Frampton & Calway, 2013). Hence the feasibility of place recognition with CNN features has been backed up empirically; this is, however, not the case for scene recognition. It has been shown that a CNN can be trained on labelled data stemming from scene recognition databases and reach appropriate performance (Zhou et al., 2014) and that this can be slightly improved if feature maps extracted from a CNN are used to train a linear classifier on the same data set (Wang & Wu, 2014); both these approaches, however, require large labelled data sets for training. Furthermore, human-level performance has not been achieved by either of these (Borji & Itti, 2014).

This study attempts to answer the question whether features extracted from a CNN trained on the object classification database ImageNet (Jia Deng et al., 2009) can be used to simultaneously recognise places and scenes in an unsupervised regime; no retraining of the network will be carried out. Whilst testing the feasibility of this approach is of scientific interest in its own right, it is also of interest for a potential application in autonomouslynavigating agents. Understanding the scene one is in is relevant to enable context-sensitive adaptation of one's driving behaviour (Seff & Xiao, 2016). Whilst localisation is an easilyachievable task with the help of the Global Positioning System, it is difficult to infer the general scene based on the location alone, due to a lack of relevant labelled data (Chu et al., 2006). Applying CNN features in an unsupervised context comes with a set of questions, most importantly about features of which layer depth allow for best performance. Previous studies have shown that the middle layers lead to the best results; it has been argued, rather intuitively, that nodes in early layers code only for basic shapes (and are hence *undertrained*) whereas nodes in later layers code mostly for task-relevant information and are hence *overtrained* for deployment in a different context (Yosinski et al., 2014). Equally, it is known that sparsity, i.e. the percentage of nodes that take on zero in a given layer, increases drastically with layer depth (Milde et al., 2017). This is of specific importance as newer accelerator architectures allow to disregard any zero value node, as in the NullHop architecture (Aimar et al., 2017), i.e. later layers come with additional computational benefits. In essence, it is tested whether unseen visual environments that are more similar to only one of a number of previously-seen visual scenes for a human observer will be classified as that respective scene with accuracies significantly above chance level and whether previously-presented images will be recognised as such.

2. METHODS

A novel task was created to test place and scene recognition in an unsupervised set up. The algorithm is first trained with no labels being presented. Input images representing a drive through either an inner city or countryside environment (see 2.1) were then given; for each image it was tested whether it was correctly defined as previously-observed or new (place recognition; see 2.4.1). It was furthermore tested whether the correct scene was recognised (scene recognition; see 2.4.2). The algorithm consisted of a CNN to extract features from the images (see 2.2), a place memory and matching algorithm for place recognition (see 2.4.1) and a scene recognition mechanism that was based on *k*-means clustering (see 2.4.2). Difference measures were also taken (see 2.3).

2.1 Input data

The image data was taken from the KITTI Vision Benchmark Suite, which consists of visual driving data recorded by Annieway, the autonomous driving platform of the Karlsruhe Institute of Technology (Geiger et al., 2012). Four data sets were chosen to represent data of one of two scenes, either inner *city driving* or *countryside driving*. The sequences 2011_09_26_drive_0039 and 2011_09_26_drive_0091 were chosen to represent inner city driving; the sequences 2011_09_26_drive_0014 and 2011_09_26_drive_0056 were chosen to represent countryside driving (see figures 1). The length of each of these sequences are shown in table 1. Unsynced and unrectified images were used. Two input sizes were analysed; firstly in original size, i.e **1392 x 512** pixels (width x height), or the shorter side of the image has been scaled down to the expect input size for each network, resulting dimension in **617 x 227** pixels for AlexNet and **609 x 224** pixels for VGG16 (see section 2.2 for further information about these networks).

To enhance the readability of this report, hereafter the two city data sets, i.e. 2011_09_26_drive_0039 and 2011_09_26_drive_0091, will be labelled as *city_1* and *city_2* respectively; equally, 2011_09_26_drive_0014 and 2011_09_26_drive_0056 will be labelled as *countryside_1* and *countryside_2*.

Table 1. **Length of data sets.** Data sets used in this study listed by their length, measured in number of images.

city_1	city_2	countryside_1	countryside_2
401	346	320	300

2.2 Convolutional neural networks

Feature extraction is carried out by a CNN; this describes a type of feedforward ANN with architectural parameters set to resemble characteristics reminiscent to those of the mammalian visual system. A CNN is a further development of the neocognitron (Fukushima et al., 1983). To give a simplified sketch of its workings, a CNN contains three types of layers: convolutional layers, pooling layers and fully-connected layers. Nodes are not connected to all nodes in the successive layers, but rather to a certain subset. In convolutional layers, each node represents a filter that is convolved over the section of the input volume that the selectively-connected subset is tuned to; abstractly, these layers detect features in the input image by applying filters to each image position. Pooling layers reduce the dimensionality; intuitively this entails that having identified a certain feature is deemed to be more important than retaining its exact location, i.e. feature representations become invariant to the location in which the original features occurred. Fully-connected layers then form the output of the network, which, in the context of object recognition, is a vector representing the likelihood that the image contains an instance of each class of objects that the network knows about; the highest likelihood value represents the object the network has recognised in the input (Krizhevsky et al., 2012).

Choice of network Two networks have been chosen for the purpose of this study; firstly, the 16-layer CNN of the Visual Geometry Group (VGG16; Simonyan & Zisserman 2014) of the University of Oxford, trained on the ImageNet database (Jia Deng et al., 2009), was chosen because of its relatively high number of convolutional layers. It was also shown that the VGG16 is capable to represent more complex features and shows higher levels of absolute sparsity (Yu et al., 2014) than the AlexNet, the other network that has been used in this study (Krizhevsky et al., 2012). The AlexNet has been chosen for its relatively small size, which might be important for a potential robotic application. An analysis of other networks, e.g. networks that are pretrained on driving data, is subject of future research (see 4.5).

Figure 1. **Input images.** Example images for all data sets, with 2011_09_26 drive_0039 (or *city_1*) and 2011_09_26 drive_0091 (or *city_2*) shown in the left column; 2011_09_26_drive_0014 (or *countryside_1*) and 2011_09_26_drive_0056 (or *countryside_2*) are shown in the right column. A comparison within a column would hence describe a *place* distinction whereas comparison within a row would represent a *scene* distinction.

Feature extraction For this study, the algorithm is run with feature activity stemming from a specific layer at a time. As a first step of pre-processing, each matrix of feature activity resulting from the analysis of one image, usually called a feature *map*, will be flattened into a one-dimensional vector consisting of *n* 32-bit floating point values. Vector lengths per data set and network choice are shown in tables 2 and 3; AlexNet was run with the full-sized data whereas VGG16 was run on the resized data set due to computational constraints. Implementations of both of these pretrained networks were taken from Caffe, a deep learning toolbox created at the UC Berkeley (Jia et al., 2014). The last, fully-connected layers of both networks have been removed so that images can be used that differ from the image size of the training set. Other networks can be easily tested with the code library that was produced; the library is available upon request.

Table 2. **Feature vector sizes for AlexNet.** Sizes of flattened feature vector per layer when AlexNet is used, across original and resized data sets

layer	AlexNet, resized data	AlexNet, full-sized data
conv1	802,560	4,185,216
conv2	525,312	2,790,144
conv3 & conv4	189,696	1,023,744
conv5	126,464	682,496

Table 3. **Feature vector sizes for VGG16.** Sizes of flattened feature vector per layer when VGG16 is used, across original and resized data sets

layer	VGG16, resized data	VGG16, full-sized data
conv1_1 & conv1_2	8,730,624	45,613,056
conv2_1 & conv2_2	4,372,480	22,806,528
conv3_1, conv3_2 & conv3_3	2,193,408	11,403,264
conv4_1, conv4_2 & conv4_3	1,103,872	5,701,632
conv5_1, conv5_2 & conv5_3	279,552	1,425,408

2.3 Analyses of difference

A node being active in a CNN can intuitively be described as denoting the presence of a specific feature in the input image; learning is thought to be a process that leads to each node developing their

tuning to one such specific feature (Schmidhuber, 2014). Due to the co-occurrence of similar objects (and, hence, the features that constitute these objects) in a scene, it has been hypothesised that significantly higher difference values will be observed when the feature activity resulting from images across the same scenes are compared, rather than when a comparison is carried out between feature activity resulting from images of the same scene. Three measures were carried out to examine difference: (1) root-squared difference (RSD), (2) Hamming distance and (3) correlation coefficients, which will be outlined below. All three analyses follow the same schema:

1. The comparison involves four data sets, two of each visual environment (see section 2.1 and figure 1).
2. Each data set will be compared with all others, resulting in two comparisons between the same scenes and four comparisons between different scenes.
3. Each given comparison occurs image-per-image, i.e. feature activity for image one of data set one will be compared with the feature activity for image one of data set two, and so on. If the sequences are of unequal length the analysis will be stopped once the end of the shorter sequence is reached.
4. The respective results will be compared by an independent *t*-test.

A principal component analysis has been carried out to understand if more variability can be explained with the same number of components with layer progression, indicating that the relevant information is found over less features; this analysis is presented in appendix B.

(1) Mean root-squared-difference To calculate the RSD, two respective feature vectors, *u* and *v*, are drawn; their difference is calculated node-wise and then squared. The sum of this node-wise squared difference is then taken and, in a final step, the square root of this value is drawn (see equation 1).

$$d^{RSD}(u,v) = \sqrt{(u_1 - v_1)^2 + \ldots + (u_n - v_n)^2} \quad (1)$$

(2) Mean Hamming distance Hamming distance, an information theoretic difference measure that goes back to Richard Hamming, quantifes the amount of positions that would need to be changed to turn one given vector into another given vector (Hamming, 1950). In this case, the respective feature maps, *u* and *v*, are treated in a binary format; the value of a node takes on a 1 if the value exceeds zero, or a 0 if the node's value is exactly zero. Intuitively, it will hence be compared whether similar visual scenes lead to a similar pattern of nodes being activated. Mathematically, the Hamming distance function of two feature maps, *d(u, v)*, is defined as the sum over all unequal values at the same position (see equation 2).

$$d^{HD}(u,v) = \sum_{k=0}^{n} [u_k \neq v_k] \quad (2)$$

(3) Mean Pearson's r To compare the correlation of feature vectors between different visual environments, the Pearson product-moment correlation coefficient, or – more commonly – Pearson's *r*, is drawn for each image comparison. Pearson's *r* results from the covariance of two given variables divided by the product of their standard deviations (see equation 3).

$$d^{CORR}(u,v) = \frac{cov(u,v)}{\sigma u * \sigma v} \quad (3)$$

where cov(u, v) describes the covariance measure, or the linear association between the two variables, which is defined as the expected value function of the given vector minus its mean (see equation 4).

$$cov(u,v) = \sum_{i=1}^{N} \frac{(u_i - \mu_u)(v_i - \mu_v)}{N} \quad (4)$$

2.4 The dual-stream algorithm

As a top-level description, the algorithm is built up of three major components: a CNN to extract features from the input images, a place memory combined with a matching mechanism that allows to compare new observations against previously-observed examples, and lastly a scene recognition algorithm, based on k-means clustering, that allows to classify an observation as belonging to a visual environment in an unsupervised fashion (see figure 2). Place recognition and scene classification occur in parallel, once the features are extracted by the CNN; hence this algorithm is reminiscent of a dual-stream system. These streams are further outlined in sections 2.4.1 and 2.4.2, respectively.

2.4.1 Place recognition

The first of the two processing streams entails a collection, or memory, of previously seen visual information. The main target of this algorithm is to define whether a new observation should be regarded as an example of a previously-seen place or should rather be defined as a new place. The algorithm contains a matrix *P* of such defined places as well as a comparison algorithm, namely *RSD* with an added term *t* that compares the result with a previously-defined threshold value. Firstly, a new observation *u* is compared against all stored places in the array *P* and the comparison with the lowest difierence value is stored. This value is then compared to the threshold (see below); if the value is above the threshold it is defined as a new place, otherwise it is defined as an instance of a previously-encountered place (see also algorithm 1).

Threshold definition The threshold is defined relative to the observed difference values between the first *n* observations of all training data sets; the mean RSD between all these observations was defined as the threshold for the purpose of this study.

2.4.2 Scene recognition and *k*-means clustering

Scene classification is based on *k*-means clustering, a class of algorithms that assign n observations to *k* groups over unlabelled data and can hence be described as an unsupervised machine learning algorithm (MacQueen, 1967). *k*, or the number of clusters, has to be defined a priori. Clustering is achieved by minimising the

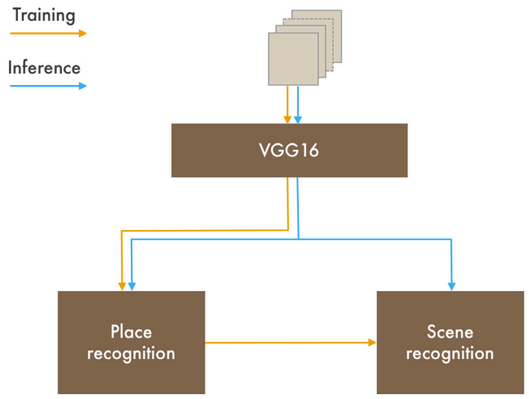


Figure 2. **Top-level algorithm schematics.** A set of input images are first analysed by the given CNN and the feature activity of the given layer is then fed to two algorithms. To achieve place recognition the feature activity is compared to previously-seen places; it is here that it is assessed whether the image represents a known place or should be defined as a new place. The scene recognition compartment analyses, through *k*-means clustering, which of the known general visual scenes the given input belongs to.

Algorithm 1. **Building up a place memory**

```
function RS_DISTANCE(u, v)                          ▷ Element-wise root-squared difference
    for element i in u do
        result += square_root(u_i + v_i)
for 0 to number of elements in P do                 ▷ Memory build up
    d_i = rs_distance(u, P_i)
if min(d) > threshold then                          ▷ Decision: known or new place?
    P_i+1 = u
```

sum of a squared distance measure, often Euclidean distance, between elements within a cluster and their centroids. After *k* centroids have been randomly initialised, the following iterative process is carried out until convergence (defined as no further change in cluster assignment):

1. Find the nearest centroid for each of the points, or features, in the data set.
2. Assign each point to the cluster that the centroid represents.
3. For the number of *k*, compute the centroid position of the data points that were assigned to this cluster; move the centroid to this position (MacQueen, 1967).

For this study, *k* was set to *k* = 2 in order to represent the two visual scenes to be distinguished from each other. The place memory, as outlined in 2.4.1, is being clustered. Clustering as such does not allow for classification; representative snippets of each visual environment are further analysed with respect to the closest cluster centroid of each of their observations. Two phases are distinguished, training (see 2.4.3) and inference (see 2.4.4). Cluster evaluation measures have also been taken and are presented in appendix C.

2.4.3 Training phase

The schematics of the training phase are outlined in figure 2. During training only the place memory processing stream is active; observations will not be directly analysed by the scene recognition algorithm. As a first step during training, all distinct places will be collected in the place memory (see 2.4.1 for the mechanism). Once all training data has been analysed by the place recognition algorithm, the resulting place matrix *P* is used to generate the *k*-means clustering. These two phases within training occurred sequentially in this study, i.e. all places were collected first, but an online updating set up is outlined for further research (see section 4.3).

2.4.4 Inference phase

Scenes are represented by vectors that result from the cluster centroid that is closest to each observation within a representative snippet of that given scene, called the *fingerprints*. This centroidal pattern is then also drawn for observations that are to be classified. To generate scene classifications, the overlap between the hash values of the representative scene snippets and the test sample is drawn. The test sample is defined as being part of the scene with which the overlap is higher; see algorithm 2. This is an instance of a nearest centroid classifier (Lange et al., 2004).

Algorithm 2. **Unsupervised classification**

```
for fingerprint in fingerprints do                  ▷ Iterate through all known scenes
    overlap_fingerprint = sum(fingerprint & allocation_observation)
scene_observation = argmax(overlap)                 ▷ Classification
```

2.5 Scenarios and accurate behaviours

Scenarios Three scenarios have been established and the respective desired behaviour defined for each of these; the resulting accuracy would be scored as a 1 if the algorithmic behaviour matches the desired behaviour over all test items. Testing is carried out in each of the domains, place recognition and scene recognition. Whilst the aim of this research effort was to carry out place recognition it was not possible to do so due to the lack of ground-truth information in the input data. It was, however, feasible to understand if the allocation of places was carried out in a sensible way; hence the task will hereafter be called *place allocation*. Three scenarios have been created to test place allocation; firstly, the training examples have been re-introduced into the system, with the desired behaviour being that each observation is allocated to a place memory that originally came from the same data set as that given observation (see figure 3a). Further images from the training data, which have been withheld during the training process, are presented in scenario two (see figure 3b); the desired behaviour is that all images are defined as not previously seen. The third scenario entails presenting separate data sets from the same scenes (see figure 3c); the desired behaviour, as before, is that all images are defined as not previously encountered. Scene recognition was based on that third scenario.

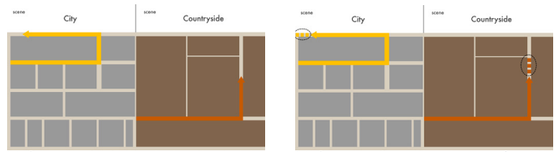
3. RESULTS

Analyses of difference were taken to test the hypothesis that the co-occurrence of similar objects in the same scene would lead to lower difference measures when members of the same scene are compared to each other (see 2.3). To test for place allocation and scene recognition performance, scenarios with their respective desired outcome were defined and the accuracy of the algorithm was measured (see 2.5). To briefly anticipate the results, a general trend can be observed in the analyses of difference, with less difference when images from different scenes are compared in earlier layers and the reverse picture occurring with later layers (see figures 4, 5

and 6). Accuracy ratings are around the expected chance level of 50% for features of many CNN layers but show high performance in select layers (see tables 4 and 5 as well as figures 7 and 8).

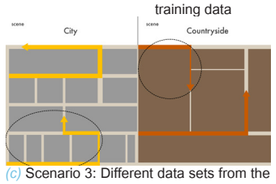
3.1 Analyses of difference

(1) Mean root-squared-difference analysis The results for the RSD analysis are presented in figure 4. For AlexNet, significant differences were found for *conv1* and *conv3*. The mean RSD when the same visual scenes are compared is lower in *conv3* and higher for *conv1*. For VGG16, significant differences were found for the first four convolutional layers and *conv3_2*; in all these the mean RSD value is higher for the same sequence comparison.



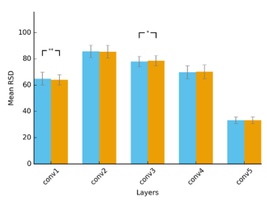
(a) Scenario 1: Reintroducing the training data

(b) Scenario 2: Further images from the training data

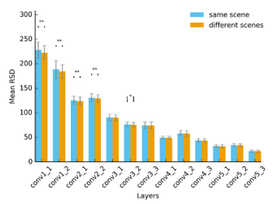


(c) Scenario 3: Different data sets from the same scenes

Figure 3. **Testing scenarios.** Testing scenarios visualised as a bird's-eye-view map; graph a) shows the training data. The dotted circles in graphs b) and c) denotes the data that has been used for testing in that respective scenario.



(a) AlexNet

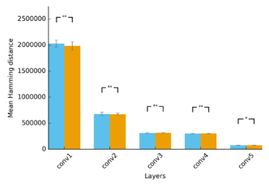


(b) VGG16

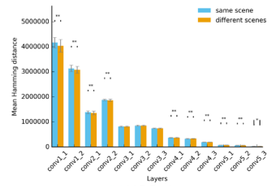
Figure 4. **Mean root-squared-difference analysis.** The mean RSD analysis was taken between feature vectors resulting from images of the same scene or different scenes across the two tested networks; significant differences can be observed in 44.4% of layers. No clear conclusions across layers can be drawn as same scene comparisons tend to show higher difference values in early layers with the reverse tendency being found in intermediate layers.

(2) Mean Hamming distance analysis The results for the Hamming distance analysis are presented in figure 5. For AlexNet, significant differences were found in all layers; lower mean Hamming distances were observed in *conv1*, *conv2* and *conv5* whereas higher mean Hamming distances were observed in *conv3* and *conv4*. For VGG16, significant difference were found in all layers but three intermediate ones; significantly higher differences, contrary to the hypothesis, were however found in six out of the thirteen convolutional layers.

(3) Mean Pearson's r analysis The results for the mean Pearson's *r* analysis are presented in figure 6. For AlexNet, significant differences were found for all layers; only for the last three layers the mean correlation coefficient was higher when the same scenes were compared. For VGG16, significant differences were found in all layers; for the first four layers the mean correlation coefficient is lower for the same sequence comparison; the following layers show a significantly stronger correlation between same sequence comparisons.

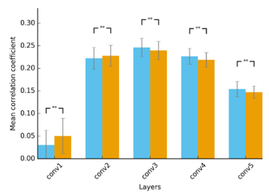


(a) AlexNet

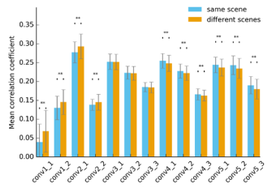


(b) VGG16

Figure 5. **Hamming distance analysis.** Hamming distance measures were taken and compared between feature vectors resulting from images of the same scene or different scenes across the two tested networks; significant differences can be observed in all layers but three intermediate ones in VGG16. No clear conclusions across layers can be drawn as same scene comparisons tend to show higher difference values in earlier layers with the reverse tendency being found in intermediate and later layers.



(a) AlexNet



(b) VGG16

Figure 6. **Mean Pearson's r analysis.** Correlation measures were taken and compared between feature vectors resulting from images of the same scene or different scenes across the two tested networks; significant differences can be observed in all but three layers. No clear conclusions across layers can be drawn as different scene comparisons tend to show higher correlation values in earlier layers with the reverse tendency being found in intermediate and later layers.

3.1 Algorithm accuracy

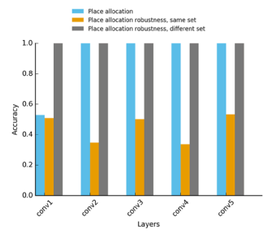
Accuracy results for each respective CNN are shown in tables 4 and 5; the place allocation and scene recognition results are visualised in figures 7 and 8, respectively. For the AlexNet layer, the highest performance for place allocation, across the three tasks, was **84.43%** in layer *conv5*; layer *conv3* showed the highest scene recognition performance with **93.7%**. In VGG16 the highest accuracy for place recognition was found in layer *conv3_1* with **99.73%**; scene recognition showed its best result in layer *conv4_3* with **87.97%**. These values result from inference over 1,376 (place recognition) and 721 (scene recognition) test images. Running the same algorithm with the raw image data as input led to significantly worse performance across all but one tasks; see figure 9.

Table 4. **Accuracy results for AlexNet.** Percentage of accurate behaviour when AlexNet was used for feature extraction; highest performance is highlighted in bold, 100%, 53.3%, 100% and 93.7% across the four tasks of place allocation with training data, similar testing, unrelated testing data and scene recognition.

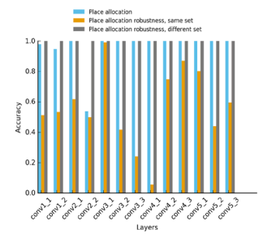
layer	Place allocation accuracy	Place allocation robustness, same set	Place allocation robustness, different set	Scene recognition
conv1	0.529	0.489	**1.0**	0.5
conv2	**1.0**	0.347	**1.0**	0.513
conv3	**1.0**	0.501	**1.0**	**0.937**
conv4	**1.0**	0.337	**1.0**	0.589
conv5	**1.0**	**0.533**	**1.0**	0.506

Table 5. **Accuracy results for VGG16.** Percentage of accurate behaviour across all tasks when VGG16 was used for feature extraction; highest performance is highlighted in bold, 100%, 99.1%, 100% and 87.97% across the four tasks of place allocation with training data, similar testing, unrelated testing data and scene recognition. - denotes that no meaningful classification could be made, in that one case due to a lack of hash value separability.

layer	Place allocation accuracy	Place allocation robustness, same set	Place allocation robustness, different set	Scene recognition
conv1_1	0.98	0.513	**1.0**	0.5
conv1_2	0.95	0.533	**1.0**	0.5
conv2_1	**1.0**	0.619	**1.0**	0.456
conv2_2	0.54	0.5	**1.0**	-
conv3_1	**1.0**	0.991	**1.0**	0.48
conv3_2	**1.0**	0.417	**1.0**	0.57
conv3_3	**1.0**	0.243	**1.0**	0.69
conv4_1	**1.0**	0.056	**1.0**	0.5
conv4_2	**1.0**	0.749	**1.0**	0.5
conv4_3	**1.0**	0.871	**1.0**	0.88
conv5_1	**1.0**	0.804	**1.0**	0.791
conv5_2	**1.0**	0.441	**1.0**	0.677
conv5_3	**1.0**	0.597	**1.0**	0.633

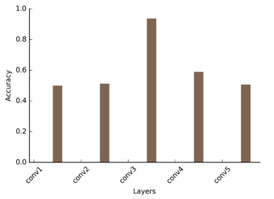


(a) AlexNet

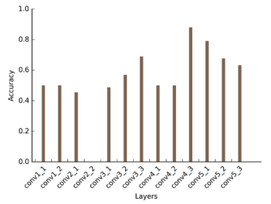


(b) VGG16

Figure 7. **Algorithmic accuracy in place allocation.** Results for the three place recognition tasks; graph a) shows the results for AlexNet features, graph b) for VGG16 features. For each test optimal behaviour was defined; this graph denotes the percentage of correct behaviour across 1,376 test images.

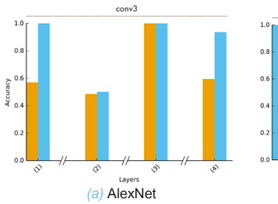


(a) AlexNet

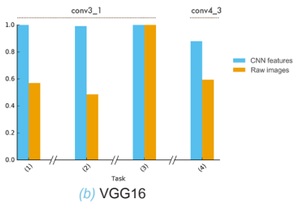


(b) VGG16

Figure 8. **Algorithmic accuracy in scene recognition.** Results for the scene recognition tasks; graph a) shows the results for AlexNet features, graph b) for VGG16 features. This graph denotes the percentage of correctly recognised scenes across 721 test images.



(a) AlexNet



(b) VGG16

Figure 9. **Comparison to raw image data from CNN features.** The same algorithm has been run with raw image data instead of CNN features. This graph shows the comparison with CNN features from the best-performing layer. The four tasks stand for (1) place allocation, (2) place allocation robustness, same set, (3) place allocation robustness, different set and (4) scene recognition. Graph a) shows the results for AlexNet features, graph b) for VGG16 features.

4. DISCUSSION

This study has shown that feature activity of specific CNN layers, if the distinctive places they represent are used for k-means clustering, can allow to classify a visual image as belonging to a general visual scene in a purely-unsupervised manner and significantly better than (1) what is expected by chance and (2) if the same algorithm is run with raw image data. As such, the results can be taken as a successful proof of concept; further testing with different data sets and scenarios is, however, needed. This study is – to the best of our knowledge the first approach to scene recognition in a fully unsupervised paradigm; previously it was shown only that scene recognition capabilities emerge when labelled data are used to either train a CNN (Zhou et al., 2014) or a support vector machine (Wang & Wu, 2014).

4.1 Layer choice

Accurate behaviour varies dramatically with layer depth; this becomes especially clear when place allocation robustness to images of the same set is taken into account, the accuracy of which ranges from 4% (in layer *conv4_1*), i.e. systematic mis-allocation, to 99% (in layer *conv3_1*) for the case of VGG16. This huge variability is, however, an intriguing finding; future research efforts should be devoted to gaining a systematic understanding of what layer depth generates features that are appropriate for a given new task. It had been argued before that medial layers are generally appropriate in the context of transfer learning (Yosinski et al., 2014). In this study it was found that – when VGG16 is used – the best accuracies in place recognition performance occurs earlier in the progression of layers (*conv3_1*) as it does for scene recognition (*conv4_3*; AlexNet shows adequate performance for both tasks in layer *conv3*). It must hence be inferred from this that feature appropriateness is dependent on the characteristics of the transferred-to task, as features extracted from different layers appear most appropriate to either place or scene recognition.

4.2 Handling the size of the place memory

If this system was to be used with more training examples, it is likely that the size of the place memory will become unmanageable rather quickly. Further research should assess ways of how the size of the place memory can be kept within manageable limits. A few approaches are conceivable; first, a consolidation mechanism could, whenever no updating occurs, iterate through all items in the place memory and remove those items that are most similar to each other. Equally, it has been shown that feature activity of a CNN is compressible due to the large percentage of zero activity which increases with layer progression (Aimar et al. 2017; see also appendix A). Further research should address whether the place memory could hold compressed representations rather than the actual input features. Lastly, recent research efforts have been looking at implementing the mechanics of a CNN whilst reducing the computational requirements (Tripathi et al., 2017), also through lowering the numerical precision (Milde et al., 2017); further research should aim to understand whether these approaches lead to the same perfor-

mance during transfer learning.

4.3 Supervision signals and continuous updating

It can be argued that some supervision was delivered in this study; firstly, the representative snippets for classification were defined *a priori*. Equally, it can be argued that the number of k, as it is set specifically to represent what is required for classification, represents a signal of supervision. It is, however, likely that both these signals will not be needed if the algorithm is run continuously. This study entailed clearly-separated phases of training and inference; it is, however, conceivable that the system can be used online, e.g. in application to a robotic system. Only minor changes would need to be made if such a system were to run in an environment with only two scenes. Firstly, a training phase needs to be run to generate sufficiently-long fingerprints for each visual scene, as well as a sufficiently large place memory. Secondly, for any new post-training observation, inference would need to be made first; after this, the place memory should be updated which, in turn, should lead to an update in the respective clustering space, e.g. through batched *k*-means approach (Bottou & Bengio, 1995).

Further research is required if such an online system should entail the capability to automatically detect and adapt to observing a new visual scene. Theoretically, an additional *k* can be spawned; the conditions under which this should be carried out require further research. It is conceivable that the distance between the centroids of such an observation representing a completely new scene – and the previously-defined cluster centroids is larger than what is observed when observations are presented that the cluster has been trained on; this hypothesis is yet to be tested. Hence a mechanism that incrementally adds clusters upon observation of completely unknown visual environments is conceivable, potentially alleviating the need for defining the value of *k*. Equally, once a new scene is detected, the next n observations could arbitrarily be defined as the fingerprint of this scene, alleviating the need for the other signal of supervision.

4.4 Further testing of place recognition capabilities

As noted before, further research is required to adequately test for place recognition, rather than just place allocation, as was done in this study. Adequate place *recognition* performance through leveraging CNN-derived features has previously been presented (Chen et al., 2014); further research should hence address the question whether the system used by Chen et al. can be combined with our approach to scene recognition in a meaningful way. More rigorous testing of place recognition performance would entail to test a number of different scenarios, e.g place recognition under vastly different lighting conditions (day and night) or when the environment has changed slightly (a certain item, e.g. a car, has been removed from a scene), to only name a few (Lowry et al., 2016).

4.5 Using other convolutional neural networks

We have chosen the VGG16 network due to its depth and the resulting slowly increasing complexity in the features it extracts (Yu et al., 2014) and the AlexNet network for its relatively low computational requirements (Krizhevsky et al., 2012). This analysis should be extended to other architectures for two reasons; firstly, it is largely unknown what the factors are that make the features of a given network appropriate for transfer to another domain. To illustrate this point, a recent study has developed a method that allows researchers to measure whether a certain pixel was used for or against a certain classification decision, in an attempt to probe the task solving strategies of different networks. Interestingly, vastly different strategies became apparent between the tested networks that were all trained on the same data set and showed similar performance (Zintgraf et al., 2017). Hence it is unknown which effects different training regimes and network architectures have on the underlying feature representations; a systematic analysis thereof might be an important next step. Secondly, large-scale data bases with labels coding which scene a given image belongs to have emerged recently (Zhou et al., 2014); it would be an interesting hypothesis for further research to test whether features extracted from network trained on such a specific data base would lead to superior performance.

4.6 Relevance to neuromorphic hardware

This study shows that k-means clustering can be seen as a capable scene distinction algorithm; it is, however, debatable whether such an algorithm is implementable in a biologically-plausible way (Pehlevan & Chklovskii, 2015). First and foremost, the data that was used in this study was comprised of full frames of visual information which goes contrary to vision in biological organisms. A retinal cell does not encode an absolute value of the input it receives at any given time point, but rather changes in contrast (Posch et al., 2014). Frame-based approaches are also computationally inefficient; information is transmitted and processed across time steps even if the input does not change, leading to redundant data and processing thereof. Event-based vision sensors have been established that only transmit information in form of events – if the intensity of visual input changes from one time step to the next, i.e. they transmit a sparse representation of the input image (Brandli et al., 2014). These sensors alleviate biological implausibilities and have been shown to speed up computation in tasks like optic flow estimation (Rueckauer & Delbruck, 2016) in comparison to frame-based approaches. The data that is generated by event-based vision sensors is best processed in an asynchronous manner as this alleviates the need for external encoding of timing; asynchronous processing naturally preserves temporal information (Chicca et al., 2014).

Such fast asynchronous parallel computation is achieved by neuromorphic devices (Schuman et al., 2017). These chips were originally motivated as a means to simulate the behaviour of neurons directly in hardware implementations (Mead, 1990). Neuromorphic engineering is a term that encompasses a variety of such approaches, e.g. through analogue means, digital means, or a mixture thereof (Schuman et al., 2017). Low power consumption and fast processing times are some of the advantages that make neuromorphic chips well-suited in the context of autonomous agents. This leads to the question – largely for further research whether the algorithm that was used in this study is potentially implementable in neuromorphic hardware. Three components would need to be implemented, (1) a place memory, (2) a matching operation and (3) a *k*-means clustering mechanism.

Memories (1) have been previously implemented in neuromorphic chips through spikebased learning rules or simulations of plasticity rules (Indiveri & Liu, 2015). The matching operation in this study (2) was carried out through a difference operation which was shown to be implementable in neuromorphic hardware in previous studies (Temam & Heliot, 2011); further research would be required to examine how the comparably large vectors could be compared within a reasonably time frame. *k*-means clustering (3) can principally be carried out based on the winner-take-all principle (Meila & Heckerman, 2013). This computational principle describes a particular set up of ANNs in which neurons within a given network compete with each other for activation; this is achieved through an organisation in which self-excitation of nodes is combined with mutual inhibition between nodes. This process, sometimes referred to as competitive learning, results in the node, or cluster of nodes, that most closely resembles the input to remain active whilst the activity of all other nodes are suppressed (Oster et al., 2009). One such model is the self-organising map algorithm, which consequently has been shown to be able to substitute a *k*-means clustering algorithm (Ba¸c ̃ao et al., 2005) whilst adding a topographic arrangement and potentially being biologicallyplausible. To briefly introduce these, self-organising maps are a class of ANN algorithms that carry out unsupervised learning. Conceptually, the node that shows the closest match with the input data is selected and, subsequently, its neighbouring nodes are strengthened to a lesser degree. As a result the dimensionality of the input data is reduced and the map structures result in topographic clusters after learning, with elements within a cluster theoretically sharing one or more characteristics with each other (Kohonen, 1990). Such self-organising maps have been shown to be implementable in biologically-inspired hardware (Rodriguez et al., 2015), though plasticity in the hardware is required, which is a topic of ongoing research (Maldonado Huayaney et al., 2016). It should hence theoretically be feasible to implement the algorithm in neuromorphic hardware; this would be the first instance of simultaneous scene and place recognition in neuromorphic chips.

5. CONCLUSIONS

In summary, it is a trivial task for humans to tell apart instances of distinct visual scenes or to remember whether we have been to a certain place before. Previous research has shown that a CNN, or the feature activity thereof, allow to perform adequately on tests measuring both these tasks when labelled training data was given (Zhou et al., 2014); place recognition was shown to work well with CNN features even in absence of labelled training data (Chen et al., 2014). The results of this study, whilst little more than a proof of concept, show that, firstly, there exists the possibility to use features extracted from a CNN trained on ImageNet to achieve adequate performance on the two tasks – scene and place recognition simultaneously. This study furthermore shows that scenes can be distinguished without any labels given to accuracies that are well above what is expected by chance level. This is the first instance of unsupervised scene recognition, to the best of our knowledge. In essence, whilst a lot of open questions remain, this study represents a successful proof of concept that warrants further research.

ACKNOWLEDGEMENTS

First and foremost, I much appreciate the superb supervision by Moritz Milde – I have learnt an immense amount during this project. I would equally like to thank Prof. Giacomo Indiveri for making this project possible in the first place. I would also like to acknowledge the valuable inputs and interesting discussion with Dr. Lorenz Müller and Enea Ceolini. Furthermore, I much appreciate the co-assessment by Prof. Jaap Murre of the University of Amsterdam.

REFERENCES

Aimar, A., Mostafa, H., Calabrese, E., Rios-Navarro, A., Tapiador-Morales, R., Lungu, I.-A., Milde, M. B., Corradi, F., Linares-Barranco, A., Liu, S.-C., Delbruck, T. (2017). NullHop: A Flexible Convolutional Neural Network Accelerator Based on Sparse Representations of Feature Maps. *arXiv*.

Athiwaratkun, B. Kang, K. (2015). Feature Representation in Convolutional Neural Networks. *arXiv*.

Bação, F., Lobo, V., Painho, M. (2005). Self-organizing Maps as Substitutes for K-Means Clustering. In *International Conference on Computational Science*, pages 476–483. Springer, Berlin, Heidelberg.

Borji, A. Itti, L. (2014). Human vs. computer in scene and object recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 113–120.

Bottou, L. Bengio, Y. (1995). Convergence Properties of the K-Means Algorithms. Advances In *Neural Information Processing Systems*, 7:585—592.

Brandli, C., Berner, R., Yang, M., Liu, S.-C., Delbruck, T. (2014). A 240x180 130dB 3us Latency Global Shutter Spatiotemporal Vision Sensor. *IEEE Journal of Solid-State Circuits*.

Chen, Z., Lam, O., Jacobson, A., Milford, M. (2014). Convolutional Neural Network-based Place Recognition. *arXiv*.

Chicca, E., Stefanini, F., Bartolozzi, C., Indiveri, G. (2014). Neuromorphic Electronic Circuits for Building Autonomous Cognitive Systems. *Proceedings of the IEEE*, 102(9):1367–1388.

Chu, S., Narayanan, S., Kuo, C.-c., Mataric, M. (2006). Where am I? Scene Recognition for Mobile Robots using Audio Features. In *2006 IEEE International Conference on Multimedia and Expo*, pages 885–888. IEEE.

Frampton, R. Calway, A. (2013). Place recognition from disparate views. *Proceedings of the British Machine Vision Conference*.

Fukushima, K., Miyake, S., Ito, T. (1983). Neocognitron: A Neural Network Model for a Mechanism of Visual Pattern Recognition. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13(5):826–834.

Geiger, A., Lenz, P., Urtasun, R. (2012). Are we ready for autonomous driving? the KITTI vision benchmark suite. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 3354–3361. IEEE.

Hamming, R. W. (1950). Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 29(2):147–160.

He, K., Zhang, X., Ren, S., Sun, J. (2016). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 11-18-Dece, pages 1026–1034.

Hira, Z. M. Gillies, D. F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in Bioinformatics*.

Ho, K. L. Newman, P. (2006). Loop closure detection in SLAM by combining visual and spatial appearance. *Robotics and Autonomous Systems*, 54:740–749.

Indiveri, G. Liu, S.-C. (2015). Memory and Information Processing in Neuromorphic Systems. *Proceedings of the IEEE*, 103(8):1379–1397.

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T. (2014). Caffe: Convolutional Architecture for Fast Feature Embedding. arXiv. Jia Deng, Wei Dong, Socher, R., Li-Jia Li, Kai Li, Li Fei-Fei (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. IEEE.

Jolliffe, I. (2014). Principal Component Analysis. *Wiley StatsRef: Statistics Reference On-*

line, pages 1–5.
Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78(9):1464–1480. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*, pages 1–9.
Lange, T., Roth, V., Braun, M. L., Buhmann, J. M. (2004). Stability-Based Validation of Clustering Solutions. *Neural Computation*, 16(6):1299–1323.
Lowe, D. (1999). Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pages 1150–1157. IEEE. Lowry, S., Sünderhauf, N., Newman, P., Leonard, J. J., Cox, D., Corke, P., Milford, M. J.
(2016). Visual Place Recognition: A Survey. IEEE Transactions on Robotics, 32(1). MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *5th Berkley symposium on mathematical statistics and probablity*, pages 281—-297.
Maldonado Huayaney, F. L., Nease, S., Chicca, E. (2016). Learning in Silicon Beyond STDP: A Neuromorphic Implementation of Multi-Factor Synaptic Plasticity With Calcium-Based Dynamics. *IEEE Transactions on Circuits and Systems I: Regular Papers, 63*(12):21892199.
Mead, C. (1990). Neuromorphic electronic systems. *Proceedings of the IEEE, 78*(10):16291636.
Meila, M. Heckerman, D. (2013). An Experimental Comparison of Several Clustering and Initialization Methods. *arXiv*.
Milde, M. B., Neil, D., Aimar, A., Delbruck, T., Indiveri, G. (2017). ADaPTION: Toolbox and Benchmark for Training Convolutional Neural Networks with Reduced Numerical Precision Weights and Activation. *arXiv*.
Nogueira, K., Penatti, O. A. B., dos Santos, J. A. (2016). Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification. *arXiv*.
Oster, M., Douglas, R., Liu, S.-C. (2009). Computation with Spikes in a Winner-Take-All Network. *Neural Computation*, 21(9):2437–2465.
Pehlevan, C. Chklovskii, D. B. (2015). A Hebbian/Anti-Hebbian Network Derived from Online Non-Negative Matrix Factorization Can Cluster and Discover Sparse Features. *arXiv*.
Posch, C., Serrano-Gotarredona, T., Linares-Barranco, B., Delbruck, T. (2014). Retinomorphic Event-Based Vision Sensors: Bioinspired Cameras With Spiking Output. *Proceedings of the IEEE, 102*(10):1470–1484.
Rand, W. M. (1971). Objective Criteria for the Evaluation of Clustering Methods. *Journal of the American Statistical Association, 66*(336):846–850.
Ringnér, M. (2008). What is principal component analysis? Nat Biotechnol, 26(3):303–304.
Rodriguez, L., Miramond, B., Granado, B. (2015). Toward a Sparse Self-Organizing Map for Neuromorphic Architectures. *ACM Journal on Emerging Technologies in Computing Systems, 11*(4):1–25.
Rosenberg, A. Hirschberg, J. (2007). V-measure: A conditional entropy-based external cluster evaluation measure. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language (EMNLP-CoNLL'07), volume 1*, pages 410–420.
Rosenblatt, F. (1957). The perceptron, a perceiving and recognizing automaton. Technical report.
Rueckauer, B. Delbruck, T. (2016). Evaluation of Event-Based Algorithms for Optical Flow with Ground-Truth from Inertial Measurement Sensor. *Frontiers in neuroscience,* 10:176.
Schmidhuber, J. (2014). Deep Learning in Neural Networks: An Overview. pages 1–88.
Schuman, C. D., Potok, T. E., Patton, R. M., Birdwell, J. D., Dean, M. E., Rose, G. S., Plank, J. S. (2017). A Survey of Neuromorphic Computing and Neural Networks in Hardware. *arXiv*.
Se, S., Lowe, D., Little, J. (2011). Vision-based mobile robot localization and mapping using scale-invariant features. In *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164), volume 2*, pages 2051–2058. IEEE.
Seff, A. Xiao, J. (2016). Learning from Maps: Visual Common Sense for Autonomous Driving. *arXiv*.
Sharma, S. Tripp, B. (2016). How Is Scene Recognition in a Convolutional Network Related to that in the Human Visual System? In *ICANN 2016: Artificial Neural Networks and Machine Learning ICANN 2016*, pages 280–287. Springer, Cham.
Simonyan, K. Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*.
Szeliski, R. (2011). *Computer vision: algorithms and applications.* Springer.
Tang, Y. (2013). Deep Learning using Linear Support Vector Machines. *arXiv*.
Temam, O. Heliot, R. (2011). Implementation of signal processing tasks on neuromorphic hardware. In *The 2011 International Joint Conference on Neural Networks*, pages 11201125. IEEE.
Tripathi, S., Dane, G., Kang, B., Bhaskaran, V., Nguyen, T. (2017). LCDet: Low-Complexity Fully-Convolutional Neural Networks for Object Detection in Embedded Systems. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, volume 2017-July*, pages 411–420.
Wang, C.-C. Wang, K.-C. (2007). Hand Posture Recognition Using Adaboost with SIFT for Human Robot Interaction. In *Recent Progress in Robotics: Viable Robotic Service to Human*, pages 317–329. Springer Berlin Heidelberg, Berlin, Heidelberg.
Wang, Y. Wu, Y. (2014). Scene Classification with Deep Convolutional Neural Networks. *arXiv*.
Wiatowski, T. Bölcskei, H. (2015). A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction. *arXiv*, pages 1–48.
Yosinski, J., Clune, J., Bengio, Y., Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems 27* (Proceedings of NIPS), 27:1–9.
Yu, W., Yang, K., Bai, Y., Yao, H., Rui, Y. (2014). Visualizing and Comparing Convolutional Neural Networks. *arXiv preprint arXiv:1412.6631*.
Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A. (2014). Learning Deep Features for Scene Recognition using Places Database. *Advances in Neural Information Processing Systems 27*, pages 487–495.
Zintgraf, L., Cohen, T., Adel, T., Welling, M. (2017). Visualizing Deep Neural Network Decisions: Prediction Difference Analysis. *arXiv*, pages 1–12.

Supplementary material for this article can be found in the digital version of the issue.

A Very Short Introduction to Cognitive Neuroscience

Review by Sammy Millard

abc BOOK REVIEW

Taking shape in the late 1970s, the relatively new discipline of cognitive neuroscience sits on the border between psychology, neuroscience, and computer science, while still retaining strong links with philosophy of mind. Due to its interdisciplinary nature, you would think that such a broad and complex topic could not be covered adequately in such a small book (less than 7 inches tall and only 110 pages long).

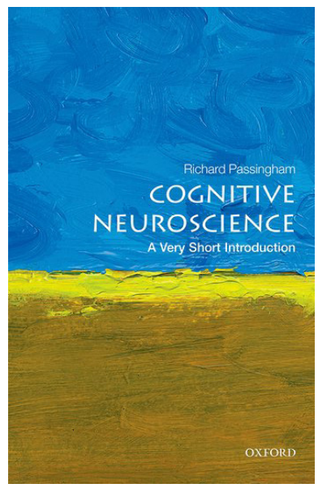
However, in this Very Short Introduction, Richard Passingham, an eminent cognitive neuroscientist studying frontal lobe mechanisms for decision-making and executive control, successfully covers the main concepts, history, and misconceptions of this diverse discipline. Moreover, he achieves this with a conversational and engaging style that is easy to follow; a breath of fresh air in comparison to the heavy textbooks you frequently see in this discipline!

The book begins with a chapter on how cognitive neuroscience developed as a way of studying what goes on in the head, something that earlier behaviourism ignored. Diagrams were used to illustrate information flow in the brain, and brain damaged patients were studied to enable suggestions of the function of missing components. By the late 1970s the development of the discipline *cognitive neuroscience* began. Passingham then gives an overview of which approaches, paradigms, and neuroscientific techniques are often used within cognitive neuroscience to bring those unacquainted with these concepts up to speed in an engaging way.

Primarily focusing on the human mind, in the main body of this book Passingham guides the reader through different branches of cognition: perceiving, attending, remembering, reasoning, deciding, checking, and acting. Each of these chapters begins by presenting three questions (e.g. in the reasoning chapter: do we think in languages?) and ends by answering them. In between, the history, misconceptions, and current state of each topic are discussed with obvious enthusiasm.

In the final chapter, Passingham goes on to contemplate what the future may hold for this exciting field. For example, the development of portable Magnetoencephalography (more commonly known as MEG) devices that image the brain using magnetic fields produced by electric currents in the brain, which will increase the range of behaviours available for scientific study. The use of computational models is also discussed, in that these will become increasingly more flexible, biologically plausible, and therefore useful to us in the future. The challenges of these future developments are also mentioned, such as the difficulty of creating computational models able to account for both fine and gross neuronal architecture.

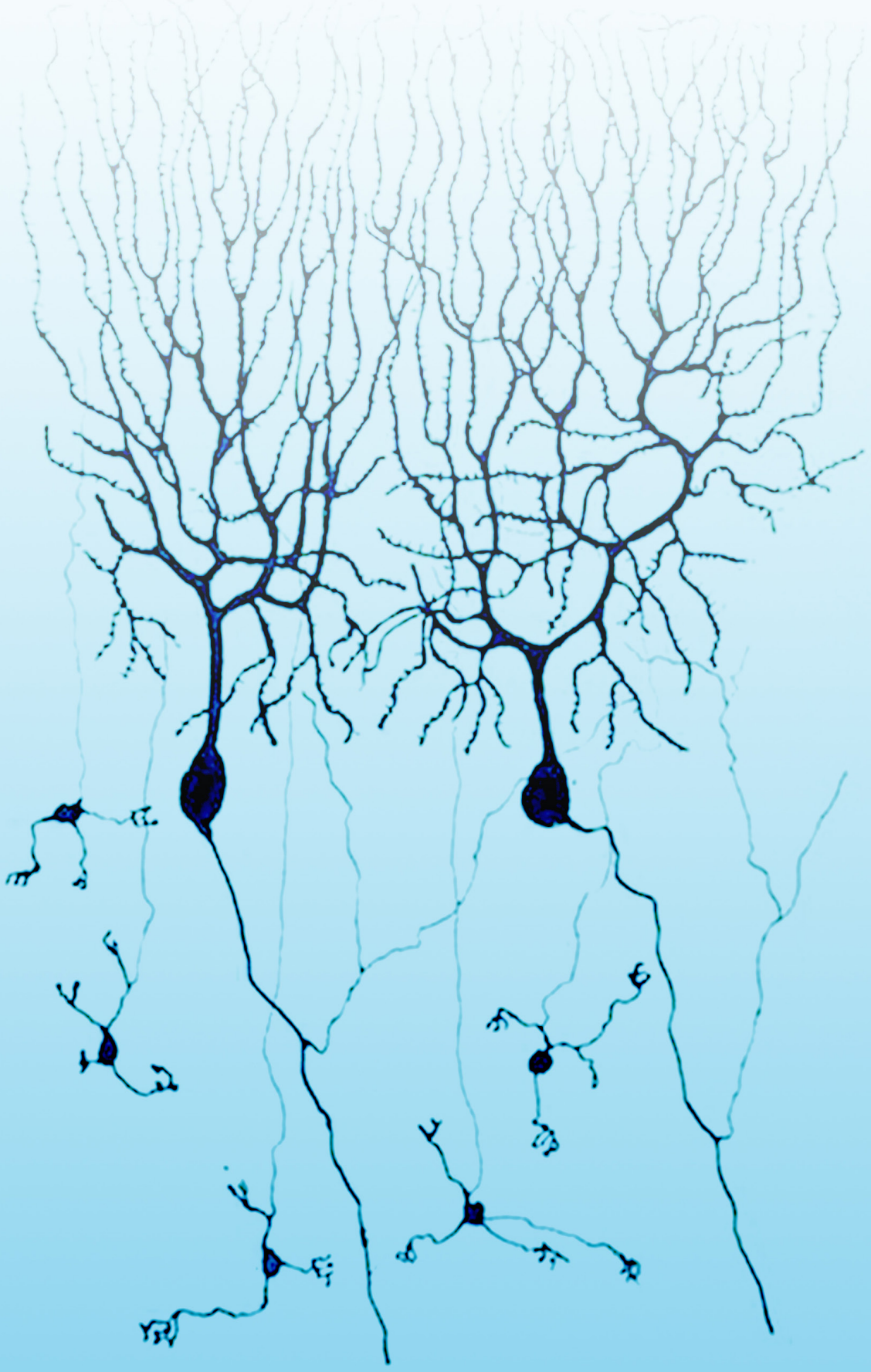
At the end of this Very Short Introduction, I was left feeling excited about what the future holds for this diverse discipline, as well as impressed with Passingham's ability to make these complex topics accessible. I would highly recommend this book to those who are unfamiliar with this discipline, but also those beginning to study it themselves. For those in the latter category, the book will enable you to learn what topics are possible within this diverse field, and will also equip you with a broader and simpler understanding of what you study. This is always helpful when you are inevitably asked that question, "but what do you *actually* study?" from your aunt at the next family gathering. Of course, because this was only an introduction, you may be left wanting more! Helpfully, the further reading section, containing suggestions for textbooks and reviews on particular topics, acts as a pleasant final parting note to this Oxford Press *Very Short Introduction to Cognitive Neuroscience*.



UNIVERSITY OF AMSTERDAM

IIS

abc amsterdam brain & cognition



Download the digital version of the Journal

student.uva.nl/bcs/content/az/abc-journal/abc-journal.html

Visit our Facebook page

facebook.com/ABCJournal