Table of Contents

Editorials **2**

Inhibition in Language Switching: fMRI Evidence From Dutch-English-German Trilinguals **5**
Angela de Bruin

Unsupervised Feature Learning Improves Prediction of Human Brain Activity in Response to Natural Images **22**
Umut Güçlü

Social and Monetary Reward Processing in Autism Spectrum Disorders With and Without Attention-Deficit Hyperactivity Disorder **34**
Jana Kruppa

Representations During Visual Mental Imagery: Effects of Low- and Higher-Level Features **50**
Brónagh McCoy

Exploring the Automaticity of Language-Perception Interactions **69**
Erik Meijs

Pointing for me or for you? Kinematics and Neural Correlates of Communicative Pointing **92**
Anke Murillo Oosterwijk

Abstracts **115**

Institutes associated with the Master's Programme Cognitive Neuroscience **119**

From the Editors-in-Chief of the CNS Journal

Dear Reader,

We hereby proudly present the first issue of the 9th volume of the *Proceedings of the Master's Programme Cognitive Neuroscience*. We are pleased to launch this volume with a special issue that highlights the many facets of functional magnetic resonance imaging (fMRI) research.

During the difficult selection process for this issue, it quickly became apparent that many of the highest-quality theses we received employed fMRI techniques. To us this shows that the fMRI research at the Donders Institute is at the forefront of investigating the complex functioning of the brain. The diversity of the studies that were selected for this issue emphasizes that fMRI goes beyond looking for "blobs", and that this technique is still evolving. In other words, fMRI methods still hold great promise for future neuroscientific research.

Although there is no doubt that fMRI is a very valuable technique in itself, we believe that we will soon see an upsurge of studies where it is combined with other neuroscience methods, like electrophysiology and transcranial magnetic stimulation. Indeed, several people at the Donders Institute are now experimenting with such combinations, and we hope that these methods will also become a part of the training that the CNS Master students will receive in the future. Overcoming the limitations of using just one technique for measuring brain activity is an important goal, as this will allow us to paint a more elaborate and complete picture of the neural underpinnings of cognitive functions. We are highly enthusiastic about reflecting these exciting developments in future issues of the journal.

On a different note, our Master's programme has seen the appointment of a new Programme Director, Prof. Ardi Roelofs, since the last issue was published. We are very grateful for the extensive support provided by the previous Programme Director, Prof. Ruud Meulenbroek, and we would like to congratulate Prof. Roelofs on his appointment. You would not be holding this journal in your hands if it were not for the support of the CNS Master management and the Donders Institute.

On behalf of our dedicated journal team, we wish you a very pleasant read!

Heidi Solberg Økland & Lorijn Zaadnoordijk
Editors-in-Chief

From the Speaker of the Donders Graduate School for Cognitive Neuroscience

Dear Readers,

It is my sincere pleasure to introduce the new issue of the *Proceedings of the Master's Programme Cognitive Neuroscience* (CNS Proceedings) from the Donders Institute at Radboud University Nijmegen. The well-known CNS Master's program has won numerous awards for its outstanding quality as the educational wing of the Donders Institute. The CNS Proceedings constitutes a key spin-off of the program. This marvellous endeavour enables Master's students to learn how to edit a scientific journal. To do so, they embrace all aspects of the editorial portfolio from overseeing the peer-review process, communicating with authors, nudging busy referees and developing the final layout of the journal.

The current issue covers a variety of topics along the four Donders research themes. It features encoding/decoding models (Güçlü), trilingualism (de Bruin), reward processing in autism (Kruppa), communicative pointing (Murillo Oosterwijk), visual mental imagery (McCoy) and language-perception interactions (Meijs).

The prominent method employed in this issue's papers is functional magnetic resonance imaging (fMRI). Since its advent more than 20 years ago, it has become the most predominant and ubiquitous research tool in cognitive neuroscience. fMRI has been crucial to the explosion of knowledge in neuroscience, advancing our understanding of the pivotal neural systems that underpin higher cognitive functions such as attention, memory and language, to name just a few. Ground-breaking technological advancements, such as machine-learning based pattern analyses of brain activity and layer-resolved neuroimaging at high magnetic fields will tackle the future burning questions in cognitive neuroscience. However, fMRI also faces harsh challenges, from irreproducibility of findings due to lack of power to 'Voodoo correlations' and popular science fMRI applications that fall short of grounding themselves on genuine scientific foundations - the 'Neuro-Disney' of today.

Skimming through past issues of the CNS Proceedings, it becomes apparent that many of the dedicated students who published in this journal continue their work here at the Donders Institute as PhD students. The Master's program is complemented by a second educational pillar within the Donders Graduate School for Cognitive Neuroscience - the PhD track. Like the MSc track, it is aligned with the overarching four Donders research themes, which ensures optimal crosstalk between training and research. It is encouraging that publishing in the CNS proceedings lays the foundation for embarking on a successful career in science, a sign of the thriving and vibrant Donders neuroscience community.

My warmest congratulations to the authors and editors of the current issue of the CNS Proceedings. I wish you all a pleasant read!

Best wishes,

Christian Doeller
Principal Investigator at the Donders Institute (Theme: 'Plasticity and Memory'),
Donders Centre for Cognitive Neuroimaging and the Faculty of Science
Speaker of the Donders Graduate School for Cognitive Neuroscience

Proceedings of the Master's Programme Cognitive Neuroscience of the Radboud University Nijmegen

Editors-in-Chief
Heidi Solberg Økland
Lorijn Zaadnoordijk

Editor Language & Communication
Friederike (Rike) Seyfried

Editor Perception, Action & Control
Christian Utzerath

Editor Learning, Memory & Plasticity
Ramona Barte

Editor Brain Networks & Neuronal Communication
Sophie Esterer

Layout Team
Jessica Classen
Michal Czaplinski
Suhas Hassan Vijayakumar

Public Relations Chief
Lorijn Zaadnoordijk

Senior Subeditors
Imme Lammertink
Heidi Solberg Økland

Webmaster
Jana Krutwig

Assistant Editors Language & Communication
Rosanne Rietveld
Ksenija Slivac

Assistant Editors Perception, Action and Control
Andreea Loredana Cretu
Martin Sjøgård

Assistant Editors Learning, Memory & Plasticity
Dorien Maas
Constanze von Randow

Assistant Editors Brain Networks & Neuronal Communication
Jordy Thielen
Kelly Woudsma

Public Relations Assistants
Caitlin Coughler
Nadine Dijkstra
Ksenija Slivac
Renske van der Cruijsen

Assistant Subeditors
Anne Mickan
Marise van de Molengraft

Programme Director: **Ardi Roelofs**
Senior Advisor: **Roshan Cools**
Cover Image: **Layout Team**
Cover Acknowledgement: **MRI slices from University of North Carolina**
Journal Logo: **Claudia Lüttke**
Photo Editors-in-Chief: **Tim van Mourik**
Photo Christian Doeller: **supplied by Christian Doeller**

Contact Information: **Journal CNS**
Radboud University
Postbus 9104
6500 HE Nijmegen
The Netherlands

nijmegencns@gmail.com

Inhibition in Language Switching: fMRI Evidence From Dutch-English-German Trilinguals

Angela de Bruin[1]
Supervisors: Ton Dijkstra[1], Ardi Roelofs[1], Ian FitzPatrick[1,2]

[1] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2] Heinrich Heine Universität Düsseldorf, Germany

The prevailing theory of language switching states that unbalanced bilinguals use inhibition to switch between their languages (Inhibitory Control or IC model; Green, 1998). Using functional Magnetic Resonance Imaging (fMRI), we examined the brain mechanisms underlying language switching and investigated the role of domain-general inhibition areas such as the right inferior frontal gyrus (rIFG) and the pre-supplementary motor area (pre-SMA). Dutch-English-German trilinguals performed a picture naming task in the MRI scanner in a blocked-language and a mixed-language context. The rIFG and pre-SMA showed more activation for switches to the second and third language (L2 and L3) compared to non-switch trials and blocked trials. No such difference was found for switches to the first language (L1). Our results indicate that language switching recruits brain areas related to domain-general inhibition. In this way, our study supports the claim that multilinguals use inhibition to switch between their languages.

Keywords: inhibition, executive control, language switching, trilingual, fMRI

Corresponding author: Angela de Bruin; E-mail: angela.de.bruin@hetnet.nl

1. Introduction

Bilinguals are able to switch seemingly effortlessly between the languages that they speak. An important question is how they manage to do this. Green's Inhibitory Control (IC) model (Green, 1998) proposed that language switching is controlled by language-external inhibitory control networks. When a bilingual names an object, multiple lexical items in both languages become active and compete for selection. Only the item with the highest level of activation, however, will ultimately be selected. This can be achieved by inhibiting the lexical items in the non-target language. A first assumption of the IC model is that the amount of inhibition depends on the speaker's relative proficiency in a language. In unbalanced bilinguals, the first language (L1) is usually more dominant than the second language (L2). The IC model predicts more inhibition of the stronger L1 when speaking in L2, compared to less inhibition of the weaker L2 when speaking in L1. A second assumption holds that it takes time to overcome this inhibition. On the one hand, naming in L2 requires more inhibition of the stronger L1. As a consequence, it takes more time to switch back to L1, prolonging naming latencies. On the other hand, naming in L1 requires less inhibition of the weaker L2 and therefore it takes less time to switch to L2.

Abutalebi and Green (2008) proposed a related model by specifying the brain networks involved in language switching. According to this model, language switching is instantiated by brain regions also related to executive control, such as the anterior cingulate cortex (ACC) and caudate nucleus. These brain regions are involved in various aspects of executive control, such as conflict monitoring, but they aren't involved in inhibition in particular (Barbey et al., 2012; Collette et al., 2005; Niendam et al., 2012). Two brain regions that are specifically associated with inhibition, the right inferior frontal gyrus (rIFG) and pre-supplementary motor area (pre-SMA) (Jahfari et al., 2011), are not included in this model and also often missing in functional magnetic resonance imaging (fMRI) studies on language switching. In our study, we specifically wanted to address these inhibition-related brain regions to test the role of inhibition in language switching as predicted by Green's IC model (Green, 1998). We first discuss relevant behavioural and electroencephalography (EEG) studies on language switching that have considered the involvement of inhibition in language switching. In our study, we specifically wanted to address these inhibition-related brain regions to test the role of inhibition in language switching as predicted by Green's IC model (Green, 1998). We first discuss relevant behavioural and electroencephalography (EEG) studies on language switching that have considered the involvement of inhibition in language switching. Whereas some of these studies claim to have found evidence supporting the role of inhibition in language switching, others have challenged the necessity of inhibition.

1.1 Behavioural and EEG studies on language switching

Language switching studies have often used picture or digit naming experiments, in which participants name two consecutive trials in the same language (non-switch trials) or in different languages (switch trials). The difference in naming latencies between switch and non-switch trials is defined as the switch cost. Meuter and Allport (1999) were among the first researchers to conduct a switching study in bilingual language production. In their experiment, unbalanced bilinguals were asked to name digits according to a colour cue in either their L1 or L2. Naming latencies were longer in switch trials than in non-switch trials, but this effect was asymmetrical for the two languages: Switching to L1 required more time than switching to L2. This larger L1 switch cost is often taken as evidence supporting the IC model, because more time is needed to overcome L1 inhibition. Further evidence for the role of inhibition in language switching was collected in EEG experiments focussing on the N2 component of the event-related brain potential (ERP). The N2 is a frontal ERP component that is said to reflect inhibition. Jackson, Swainson, Cunnington, and Jackson (2001) found an asymmetrical behavioural effect: Switching to L1 was slower than switching to L2. The ERP data showed an increased N2 ERP component for switch compared to non-switch trials. Interestingly, this N2 effect was only significant for switches to L2 (requires more inhibition of L1), but not when switching to L1 (requires less inhibition of L2). These findings are in line with the IC model. Although most language switching studies have involved bilingual speakers, some studies have also tested multilinguals with three or four languages. Multilingual speakers can inform us about the role of language strength in language switching, as the same speaker has to control multiple non-native

languages that differ in strength. L1 switch costs in naming latencies were found to be larger than L2 or L3 switch costs. Furthermore, L2 switch costs were larger than L3 costs (Schwieter & Sunderman, 2011; Schwieter, 2013).

There are, however, also studies that argue against the necessity for inhibition in (certain instances of) language switching. For instance, Costa and Santesteban (2004) replicated the finding of asymmetrical switch costs in unbalanced Catalan-Spanish bilinguals. In balanced trilinguals with equal proficiency in Spanish and Catalan, however, switch costs were not only symmetrical for L1 and L2, but also for a weaker L3. Costa and Santesteban concluded that unbalanced bilinguals might rely on inhibition, whereas balanced bilinguals might use other mechanisms to select the target language. Other studies have challenged the necessity of inhibition even in unbalanced bilinguals. When unbalanced bilinguals could voluntarily switch between languages, switch costs to L1 and L2 were symmetrical (Gollan & Ferreira, 2009). Verhoef, Roelofs, and Chwilla (2009) found symmetrical and asymmetrical switch costs within one group of unbalanced bilinguals, depending on the preparation time for a picture stimulus. Whereas a short preparation time resulted in asymmetrical switch costs, a longer preparation yielded symmetrical costs and a larger N2 amplitude. Christoffels, Firk, and Schiller (2007) also conducted an EEG experiment on bilingual picture naming in blocked- and mixed-language contexts. In the blocked context, when all pictures were named in either German or Dutch, L1 naming was faster than L2 naming. In contrast, when participants had to switch between L1 and L2 in the mixed context, L2 naming was slightly faster than L1 naming. Switch costs, however, were symmetrical for L1 and L2. The ERP data showed an increased amplitude for non-switch compared to switch trials for L1 naming, but no such difference for L2 naming. Taken together, these results are difficult to reconcile with the IC model and suggest that language switching, even in unbalanced bilinguals, might be instantiated by mechanisms other than inhibition.

In summary, some but not all studies have found asymmetrical switch costs in naming latencies and clear evidence supporting inhibition. It is therefore questionable whether an (a)symmetry in switch costs alone could inform us whether inhibition is involved in language switching. In all, behavioural and EEG studies alone do not seem to provide unequivocal evidence in favour of the IC model. Next, we consider whether fMRI studies shed more light on the brain mechanisms involved in language switching. We first discuss the literature on brain areas involved in inhibition in general and then provide an overview of fMRI studies on language switching in particular.

1.2 FMRI studies on inhibition and language switching

Several studies have found that the rIFG, pre-SMA, and subthalamic nucleus (STN) are important brain areas for domain-general inhibition (cf., Aron, Monsell, Sahakian, & Robbins, 2004a; Aron, Robbins, & Poldrack, 2004b; Aron, 2007; Forstmann et al., 2008; Jahfari et al., 2011; Van den Wildenberg et al., 2010). Aron et al. (2004a) compared patients with a lesion in either the right or left IFG in a stop-signal task requiring response inhibition. Patients with a rIFG (but not lIFG) lesion showed disrupted inhibition of inappropriate responses. Furthermore, the damage to the rIFG correlated with the time needed to inhibit these responses. Jahfari et al. (2011) looked at effective connectivity patterns during a combined Simon and stop-signal task. Participants had to respond to a coloured shape by pressing either a left or right button. The response button could match or mismatch the spatial location of the stimulus. On one-third of the trials, participants heard a stop signal after the stimulus presentation, indicating that they had to inhibit their response. Brain activity during stop signals was explained best by a right-lateralized network including a fast and direct pathway between the rIFG, pre-SMA, and STN, showing that these areas are involved in inhibition. The rIFG and pre-SMA are also active during speech inhibition (Xue, Aron, & Poldrack, 2008). Both the inhibition of manual responses as well as of speech production (letter and pseudoword naming) elicited activation in the rIFG and pre-SMA. Focussing on the STN, there was significant activation in manual inhibition, but not during speech inhibition. The common activation of the rIFG and pre-SMA in inhibiting both speech and manual responses suggests that these areas are part of a domain-general response inhibition mechanism.

Surprisingly, although inhibition might play an important role in language switching, fMRI studies on language switching have not directly investigated brain regions related to domain-general inhibition. Rather, they often report activation in areas related to other aspects of executive control. In their review of language switching studies, Abutalebi and Green (2008) proposed a brain network for

language switching that is also involved in executive control outside the language system. According to their model, language switching requires activation of a network including the prefrontal cortex, ACC, caudate nucleus, and supramarginal gyrus. The recruitment of these areas is not specific to language switching. The ACC is involved in error detection, conflict monitoring, and conflict resolution (Aarts et al., 2008; Kerns et al., 2004). Similarly, activation in the basal ganglia does not seem to be language specific, as they are activated in motor control and planning in general as well as in executive control (Cools, 2011; Frank, 2011; Graybiel, 2000). Because these areas also play a role in general executive control, it is suggested that language switching does not require a 'special' language system, but rather shares features with other executive control functions (Abutalebi & Green, 2008). In a meta-analysis of fMRI studies on language switching, Luk, Green, Abutalebi, and Grady (2012) also reported eight areas related to executive control or language processing: Left IFG, left middle temporal gyrus (MTG), left middle frontal gyrus (MFG), right precentral gyrus, right superior temporal gyrus (STG), pre-SMA, and bilateral caudate nuclei. This analysis is largely compatible with the model proposed by Abutalebi and Green (2008), although the ACC is missing here. Nevertheless, whereas both Abutalebi and Green (2008) and Luk et al. (2012) discuss several areas related to executive control, brain regions related to inhibition in particular, the rIFG and pre-SMA, are missing or only briefly mentioned.

Several fMRI studies on language switching have found activation in the brain regions discussed by Abutalebi and Green (2008) that are related to language or executive control. The dorsolateral prefrontal cortex (DLPFC) showed increased activation when early Spanish-English bilinguals named pictures in a mixed compared to blocked manner (Hernandez, Martinez, & Kohnert, 2000). Other studies have shown that these brain regions related to executive control are especially needed when switching to L2. Wang, Xue, Chen, Xue, & Dong (2007) reported that L2 switching compared to non-switching activated the ACC, left frontal gyrus, SMA, and left temporal gyrus. Similarly, Hosoda, Hanakawa, Nariai, Ohno, & Honda (2012) reported that switching to L2 compared to switching to L1 yielded greater activation in the right DLPFC, left STG, ACC, left IFG, and left caudate nucleus. Abutalebi et al. (2013a) also found activation in the left caudate nucleus and a cluster of the pre-SMA/ACC in trilingual language switching compared to non-switching. Only activation in the caudate nucleus, however, was influenced by language proficiency.

Although most language switching studies have focussed on local switch costs in a mixed context only, Guo, Liu, Misra, & Kroll (2011) compared the brain mechanisms of local versus global inhibition. Local inhibition was defined as the control over a restricted set of memory items (e.g., specific lexical items) and was tested by comparing mixed- to blocked-language naming. In contrast, global inhibition was described as the activation or inhibition of the entire language system. This was examined by comparing the order of languages within the blocked naming context: L1 naming after L2 naming, versus L1 naming before L2 naming. Different brain regions were said to be active during global and local inhibition. During global inhibition, activation in the DLPFC and the parietal cortex was found. Local inhibition was reflected in increased activation in the ACC and SMA. This study, however, labelled these brain mechanisms a priori as 'inhibitory'. It is questionable whether the ACC and SMA can truly be argued to reflect inhibition, as they are not linked to inhibition in particular, but rather to other aspects of executive control. Taken together, these studies showed that language switching activated executive control areas, while no evidence for inhibition-specific activation was obtained.

Summarizing, multiple brain regions have been found to play a role in language switching: Left IFG, MTG, STG, precentral gyrus, DLPFC, ACC, SMA, and striatum. These brain areas are involved in language processing or belong to the frontoparietal network that has been linked to various aspects of executive control (e.g., Niendam et al., 2012). None of the available fMRI studies, however, have directly focussed on the role of domain-general inhibition networks in language switching. Two brain regions that have been linked to inhibition in particular, the rIFG and pre-SMA (Jahfari et al., 2011), are missing in most fMRI studies on language switching. Rather, the claim of inhibition is often based on activation in brain areas such as the ACC that are related to, amongst others, conflict resolution instead of inhibition.

1.3 Current study

The current study therefore aimed to examine the role of domain-general inhibition areas in language switching. We used fMRI to investigate the brain mechanisms involved in language switching and we specifically wanted to compare the role of inhibition networks in switching to non-native languages (L2,

L3) versus switching to the native language (L1). To investigate this question, participants performed an overt picture naming experiment in the MRI scanner. Pictures were named in a blocked context (all pictures named per language) and in a mixed context (pictures named interchangeably in three languages). The mixed context consisted of both switch trials and non-switch trials. We focussed on local inhibition only and tested the presence of such inhibition in two ways: As the difference between switch and non-switch trials, and as the difference between mixed and blocked naming. Following the IC model (Green, 1998), we predicted to find (local) inhibition during language switching, reflected by more activation in the rIFG and pre-SMA (Jahfari et al., 2011). We expected more inhibition during switch than non-switch trials, specifically for L2 and L3, because switching to these weaker languages requires inhibition of L1 (Green, 1998). Comparing blocked to mixed naming, we predicted to find more inhibition in the mixed context. Again, we specifically expected this difference between blocked and mixed naming to be larger for L2 and L3 than for L1.

Besides this main issue, we also wanted to examine the link between language switching and performance on non-linguistic inhibition tasks. Many researchers have claimed that language switching is related to non-linguistic inhibition tasks and task switching (e.g., Bialystok, Craik, Klein, & Viswanathan, 2004). To further investigate the relationship between language switching and domain-general inhibition, our participants performed two tasks that have been argued to reflect inhibition skills (Simon task and stop-signal task; Van den Wildenberg et al., 2010). We expected to find a correlation between language switch costs and performance on the Simon and stop-signal task.

As a third aim, we intended to address the influence of language proficiency on inhibition by comparing two non-native languages of different proficiency levels. We therefore tested trilingual participants in Dutch (L1), English (L2), and German (L3). In this way, the influence of non-native language proficiency on language switching could be tested directly within a participant.

2. Methods

All participants first took part in a behavioural picture naming experiment, which was followed by a Simon task (cf., Bialystok et al., 2004), stop-signal task (Verbruggen, Logan, & Stevens, 2008), operation span task (cf., Conway, Cowan, Bunting, Therriault, & Minkoff, 2002), LexTALE (Lemhöfer & Broersma, 2011), and Boston Naming Test (BNT, Kaplan, Goodglass, & Weintraub,1983). Approximately three months later, the same participants performed the picture naming task in the MRI scanner.

2.1 Participants

During the behavioural experiment, 27 students of the Radboud University Nijmegen participated in return for either course credits or payment. All participants were right-handed Dutch (L1) native speakers with good proficiency in English (L2) and intermediate proficiency in German (L3). All participants had normal or corrected-to-normal vision and none had any neurological, reading or hearing impairment. All participants gave informed consent. The data of three participants were excluded because they either had an insufficient level of proficiency in German (two participants) or high error rates (> 33.3%) in the picture naming task (one participant). The final sample of the behavioural experiment consisted of 24 subjects, 19 female, aged from 18 to 27 years (M = 22.22 years; SD = 3.89). A subset of 18 participants took part in the fMRI experiment. One participant was excluded from further analyses as she did not finish the experiment. The final sample consisted of 17 participants, 12 female, aged from 18 to 25 years (M = 21.82; SD = 2.30). Table 1 shows the L2 and L3 Age of Acquisition (AoA) according to an online self-rating questionnaire, the self-rated number of hours of language use (5-point scale: 1 = 'less than 1 hour per week'; 5 = 'more than 10 hours per week'), and self-rated language skills (7-point scale: 1 = 'very poor'; 7 = 'very good'). Furthermore, Table 1 shows the average scores per language in the Boston Naming Test (maximum score 60 points) and LexTALE (maximum score 100 points). All languages differed significantly from each other, in the expected order (L1 Dutch, L2 English, L3 German), in terms of AoA, self-ratings, and language proficiency tests.

2.2 Picture naming task

2.2.1 Materials and methods

The same stimuli were used during the behavioural and the fMRI experiment. Eighteen line drawings were selected from the picture database from the Max Planck Institute. The pictures depicted high frequent, concrete objects, and the

Table 1. Mean and standard deviation (between brackets) of the measures of the subjects' language backgrounds, self-ratings, and proficiency tests.

	L1 (Dutch)	**L2 (English)**	**L3 (German)**
Self-rating			
AoA		10.00 (1.94)	12.59 (0.94)
Hours per week	5.00 (0.00)	3.30 (1.34)	2.30 (1.73)
Listening	6.41 (.87)	5.71 (.77)	5.00 (.94)
Reading	6.76 (.44)	6.00 (.71)	4.82 (1.07)
Writing	6.47 (.62)	5.00 (.71)	3.65 (1.06)
Speaking	6.66 (.62)	5.35 (.79)	4.41 (.94)
Vocabulary	6.53 (.72)	5.35 (.61)	4.00 (1.27)
Proficiency test			
LexTALE	88.59 (8.49)	77.59 (12.80)	64.12 (6.65)
Boston Naming Test	53.43 (3.61)	38.57 (5.92)	20.50 (6.22)

picture names were matched on number of syllables, phonemes, and L1 frequency (based on the CELEX database). None of the picture names were identical cognates or false friends across the three languages. Each participant received a different pseudorandom list of stimuli. Stimuli were presented at the centre of a white background and had a size of 80 mm wide and 80 mm long. The presentation of the stimuli was controlled by Presentation Software (Neurobehavioral Systems, Albany, CA, USA).

2.2.2 Design and procedure

The experiment consisted of a blocked and a mixed naming part. In the blocked part, participants named the pictures in their L1, L2, and L3 separately, with the order of the languages counterbalanced across participants. During the mixed context, participants named successive pictures in their L1, L2, or L3 according to a colour cue. In this context, one-third of the trials were non-switch trials and two-third switch trials. Switch and non-switch trials were presented in a random order. In total, participants named 90 trials in the blocked context (30 per language) and 270 in the mixed context (90 per language).

A trial started with the presentation of a fixation cross for 500 ms (see Fig. 1), followed by a picture stimulus. This stimulus was presented for 2750 ms and then followed by a jittered interval of 4 to 6 seconds to improve the sampling of the slow BOLD-response. Voice onset times were measured with the inbuilt voicekey function of Presentation and responses were recorded with a noise cancelling microphone. Voice onset times were checked manually to ensure that they reflected the onset of speech rather than the RF pulse of the MRI scanner. Voice recording started at the presentation of the picture and naming times were only recorded during the 2750 ms the picture was on the screen. Each picture was surrounded by a colour frame that indicated the language in which participants had to name the picture (e.g., red – Dutch). Each language was combined with two different colours to avoid confounds between cue and language switching. Thus, the colour cue always changed between two subsequent trials, even if participants did not have to switch between languages. Although participants did not need the colour cue to select the language in the blocked context, the colour frame was present in both blocked and mixed naming to minimize differences between the two contexts.

Participants received written instructions before the start of the experiment, asking them to name each picture as fast and as accurately as possible, while minimizing head movements. They also received a booklet with the 18 pictures and corresponding Dutch, English, and German names to familiarize themselves with the stimulus set. After the instruction phase, participants were positioned in the scanner and named the pictures in the blocked context (90 trials). The blocked context (lasting approximately 15 minutes) was followed by the T1-scan (lasting 10 minutes). Next, the mixed picture naming task started, lasting approximately 45 minutes. The order of blocked versus mixed context was not counterbalanced to avoid local inhibition from the mixed context interfering in the blocked context (cf., Guo et al., 2011). The duration of the entire fMRI experiment was 105 minutes (90 minutes scanning time). The entire behavioural experiment lasted 90 minutes: 45 minutes for the picture naming

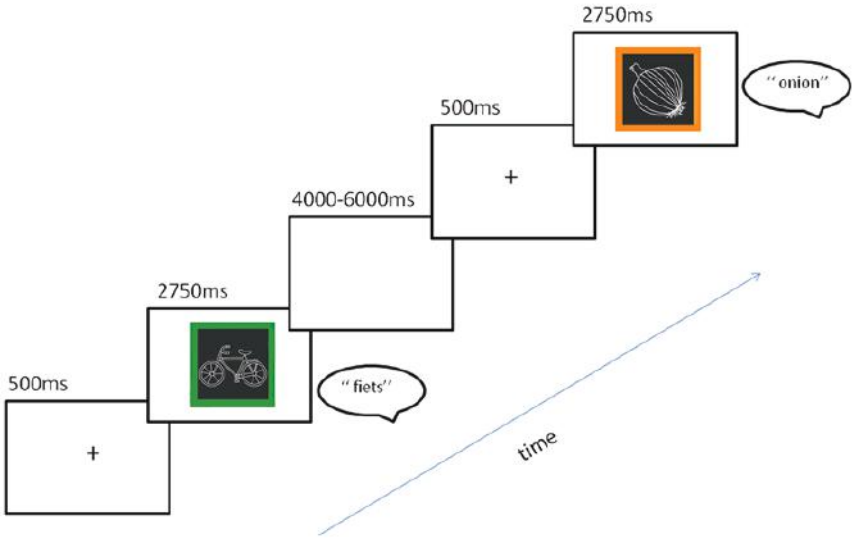


Fig. 1 Example of a trial sequence in the mixed context of the picture naming experiment. Here, the green frame indicates that the picture has to be named in L1 (Dutch), whereas the orange frame indicates that the picture has to be named in the L2 (English).

task and 45 minutes for the executive control and proficiency tasks.

2.2.3 FMRI data acquisition

Participants were scanned in a Siemens 1.5T MRI scanner. For the functional MRI data, we used a multi-echo echo-planar imaging sequence (Poser, Versluis, Hoogduin, & Norris, 2006) to reduce motion artefacts due to the language production task. Images were acquired at multiple time echoes (TEs) following a single excitation (time repetition (TR) = 2250 ms; each volume consists of 35 slices of 3.0 mm slice thickness with a slice gap of 17%; isotropic voxel size = 3.5 x 3.5 x 3.0 mm; field of view (FOV) = 224 mm). The functional images were acquired at TE 1 = 8.3 ms; TE 2 = 27.6 ms; TE 3 = 37 ms; TE 4 = 46 ms; TE 5 = 55 ms. For each subject, the first six volumes in each scan series were discarded as magnetization had not yet reached the equilibrium state. The anatomical images were acquired using a T1-weighted three-dimensional gradient-echo sequence (TR = 2300 ms; TE = 3.03 ms; FOV = 256 mm; 192 sagittal slices).

2.2.4 FMRI data analysis

We used SPM8 (Wellcome Department of Cognitive Neurology, London, UK) for image processing and statistical analysis. Image processing included realignment, slice timing correction, anatomic-functional image co-registration, segmentation, normalisation, and smoothing with a Gaussian filter of 8 mm full width at half maximum. A General Linear Model was used to estimate the effect of context for each individual subject. All contrasts were averaged, so that the two sides of the comparison were weighted equally. Naming latencies were included as a regressor of no interest. For each subject and context, significant changes in the BOLD response were assessed using t-statistics. The group averaged effects were computed with a random effects model. For group analysis, clusters with more than 10 voxels activated above a threshold of $p < 0.05$ (FWE, corrected for multiple comparisons) were considered significant. The naming latencies of both the behavioural as well as the fMRI experiment did not show any differences between L2 and L3. Therefore, we grouped the non-native languages (L2/L3) to compare them to the native language (L1) in the fMRI analysis. In the contrasts, trials were weighted so that L2 and L3 together received an equal weight compared to L1. Differences in activation were thus not due to differences in the number of trials.

To examine the brain mechanisms underlying local inhibition, we conducted two main analyses on the fMRI data: Switch versus non-switch and mixed versus blocked. Within the mixed context, we contrasted switch to non-switch trials and

specifically L2/L3 switch trials > L2/L3 non-switch trials and L1 switch trials > L1 non-switch trials. We also examined the main effect of language by contrasting L2/L3 mixed naming > L1 mixed naming. We furthermore compared the mixed to the blocked context1. To specifically address local inhibition, we also compared the mixed switch trials to the blocked context: L2/L3 mixed switch > L2/L3 blocked; L1 mixed switch > L1 blocked. Based on previous studies (Xue et al., 2008; Jahfari et al., 2011), two cortical brain regions (rIFG and pre-SMA) related to inhibition were defined as regions of interest (ROIs)2. ROIs were generated based on previously reported Montreal Neurological Institute (MNI) coordinates (Jahfari et al., 2011) in the rIFG (centre = 51, 19, 17; 16 mm radius) and the pre-SMA (centre = 9, 24, 50; 8 mm radius). We used the MarsBar ROI Toolbox (Brett, Anton, Valabregue, & Poline, 2002) to investigate the BOLD responses in these ROIs for the above-named contrasts. Contrast values (effect sizes for the ROI) were obtained from the single-subject contrast images and were exported to SPSS for group level analyses. We used a repeated measures analyses of variance (ANOVA) to identify effects of language (L1, L2/L3), trial sequence (switch, non-switch), or context (blocked, mixed) on BOLD responses in the ROIs.

3. Results

3.1 Behavioural results

The earlier behavioural experiment showed the same effects as the fMRI experiment. Here, we only report the results of the behavioural data of the fMRI experiment. We only included naming latencies of correct trials in the analysis. Trials were incorrect if there was no response, if the response was given too late or in the wrong language, or if the wrong word was selected. The experimental design included two main within-subject factors: Language (L1, L2, L3) and context (blocked, mixed). The mixed context furthermore contained the within-subject factor trial sequence (switch, non-switch trials). Error rates and naming latencies were submitted to repeated measures ANOVA for subjects (F1) as well as items (F2). We used an alpha level of .05 for all statistical tests. The error results followed the same pattern as the naming latencies and will therefore not be reported here. We first analysed the naming latencies in the blocked and mixed contexts separately, before comparing switch to non-switch trials and the mixed to the blocked context.

3.1.1 Blocked context

In the blocked context, there was a significant effect of language, $F1(2, 32) = 12.65$, $p < .001$, $\eta p2 = .44$; $F2(2, 34) = 10.40$, $p < .001$, $\eta p2 = .38$. Naming latencies thus differed among the three languages in the blocked context. A further comparison of the individual languages showed a significant difference in naming latencies between L1 and L3, $F1(1, 16) = 25.88$, $p < .001$, $\eta p2 = .62$; $F2(1, 17) = 12.89$, $p = .002$, $\eta p2 = .43$, and between L1 and L2, $F1(1, 16) = 17.81$, $p = .001$, $\eta p2 = .53$; $F2(1, 17) = 15.56$, $p = .001$, $\eta p2 = .48$. Between L2 and L3, however, there was no significant difference, $F1(1, 16) = .99$, $p = .34$; $F2(1, 17) = .050$, $p = .82$. Naming in L1 (1150 ms) was faster than naming in L2 (1267 ms) or L3 (1301 ms). Pictures were named equally fast in L2 and L3.

3.1.2 Mixed context

The mixed context showed no significant main effect of language, $F1(2, 32) = .84$, $p = .44$; $F2(2, 34) = 3.88$, $p = .065$. Whereas L1 was named faster than L2 and L3 in the blocked context, all languages were named equally fast in the mixed context (L1 = 1589 ms; L2 = 1560 ms; L3 = 1535 ms).

3.1.3 Switch versus non-switch trials

Within the mixed context, we compared switch to non-switch trials to test for effects of inhibition (see Fig. 2). Switch trials (1581 ms) were significantly slower than non-switch trials (1521 ms), $F1(1, 16) = 11.62$, $p = .003$, $\eta p2 = .70$; $F2(1, 17) = 15.70$, $p = .001$, $\eta p2 = .84$. There was no interaction of trial sequence (switch, non-switch trials) and language (L1, L2, L3), $F1(2, 32) = 2.83$, $p = 0.074$, $\eta p2 = .15$; $F2(2, 34) = 1.17$, $p = .32$, $\eta p2 = .060$. The difference between switch and non-switch trials (i.e., the switch costs) thus did not differ significantly between L1 (97 ms), L2 (31 ms) and L3 (55 ms).

3.1.4 Mixed versus blocked

To examine whether there was a difference in inhibition between languages in the blocked and the mixed context, we tested for an interaction between context (blocked, mixed) and language (L1, L2, L3). Taking the two contexts together, there was a significant main effect of language on the item, but not subject analysis, $F1(2, 32) = 1.93$, $p = .16$, $\eta p2 = .28$; $F2(2, 34) = 3.39$, $p = .046$, $\eta p2 = .30$, and a significant main effect of context, F1(1, 16)

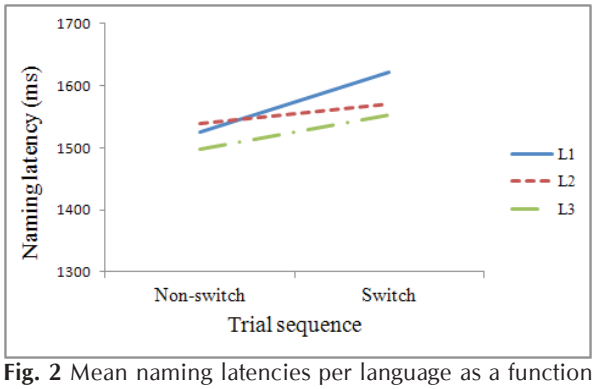


Fig. 2 Mean naming latencies per language as a function of trial sequence (non-switch, switch).

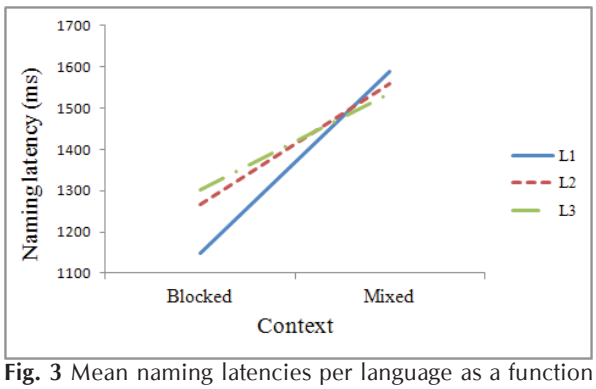


Fig. 3 Mean naming latencies per language as a function of context (blocked, mixed).

= 86.80, p < .001, ηp2 = .84; F2(1, 17) = 103.32, p < .001, ηp2 = .98. Naming was slower in the mixed context (1561 ms) than in the blocked context (1239 ms). Importantly, there was a significant interaction between context and language, F1(2, 32) = 8.70, p = .001, ηp2 = .52; F2(2, 34) = 27.96, p < .001, ηp2 = .74. The difference between naming in the mixed context minus naming in the blocked context was larger for L1 (356 ms) than for L2 (293 ms) and L3 (234 ms), suggesting that the difference in inhibition between the blocked and mixed context is larger for L1 than for L2 and L3 (see Fig. 3).

3.2 Executive control and proficiency results

Simon costs showed a trend in predicting L1 switch costs, β = -1.51, t(11) = -1.84, p = .086. This relationship was negative: The higher the Simon costs, the less time participants needed to switch to L1. Thus, the better participants performed in the Simon task, the more they inhibited L1 and consequently the larger the switch cost to L1. Simon costs were a significant predictor of L2 switch costs, β = .70, t(11) = 3.8, p = .002. This relationship was positive: The higher the Simon costs, the more time participants needed to switch to L2. Similarly, Simon costs were a significant predictor of L3 switch costs, β = .60, t(11) = 2.89, p = .012. This relationship was also positive: The higher the Simon costs, the more time participants needed to switch to L3. Performance on the stop-signal task, operation span, or proficiency scores were not significant predictors of L1, L2, or L3 switch costs.

3.3 FMRI data

We first report the ROI analysis for the rIFG and pre-SMA, followed by the whole-brain analysis.

3.3.1 ROI analysis

3.3.1.1 Switch versus non-switch trials

We expected that the difference in inhibition between switch and non-switch trials within the mixed context would be larger for L2/L3 than for L1. A repeated measures ANOVA with language and trial sequence showed no significant effect of language (p > .05 in rIFG and pre-SMA) or trial sequence (p > .05 in rIFG and pre-SMA). Importantly, however, there was a significant interaction of language by trial sequence for both the rIFG, F(16) = 8.44, p = .010, ηp2 = .35, and the pre-SMA, F(16) = 11.09, p = .004, ηp2 = .41. This interaction suggests a difference between languages in terms of inhibition in switch trials compared to non-switch trials.

We therefore analysed the differences between switch and non-switch trials for L2/L3 and L1 separately. Comparing switches to L2/L3 to L2/L3 non-switches showed significantly more activation in the rIFG, F(16) = 10.29, p = .005, ηp2 = .39, and the pre-SMA, F(16) = 5.57, p = .031, ηp2 = .26 (see Fig. 4A). Thus, there was more activation in the inhibition areas for switches to L2 and L3 compared to non-switches in L2 and L3. Comparing L1 switches to L1 non-switches, however, did not reveal significant differences in activation in either the rIFG, F(16) = 3.05, p = .10, or in the pre-SMA, F(16) = 3.46, p = .081. To summarize, switches to L2/L3 showed significantly more activation in the inhibition areas compared to L2/L3 non-switch trials. However, L1 switching did not show increased activation compared to L1 non-switch trials.

In the above analysis, we collapsed L2 and L3 trials regardless of the language of the previous trial. We thus not only compared L1-L2 and L1-L3 switches versus non-switch trials, but also included L2-L3 and L3-L2 trials. The IC model (Green, 1998), however, argues that the L1 in particular

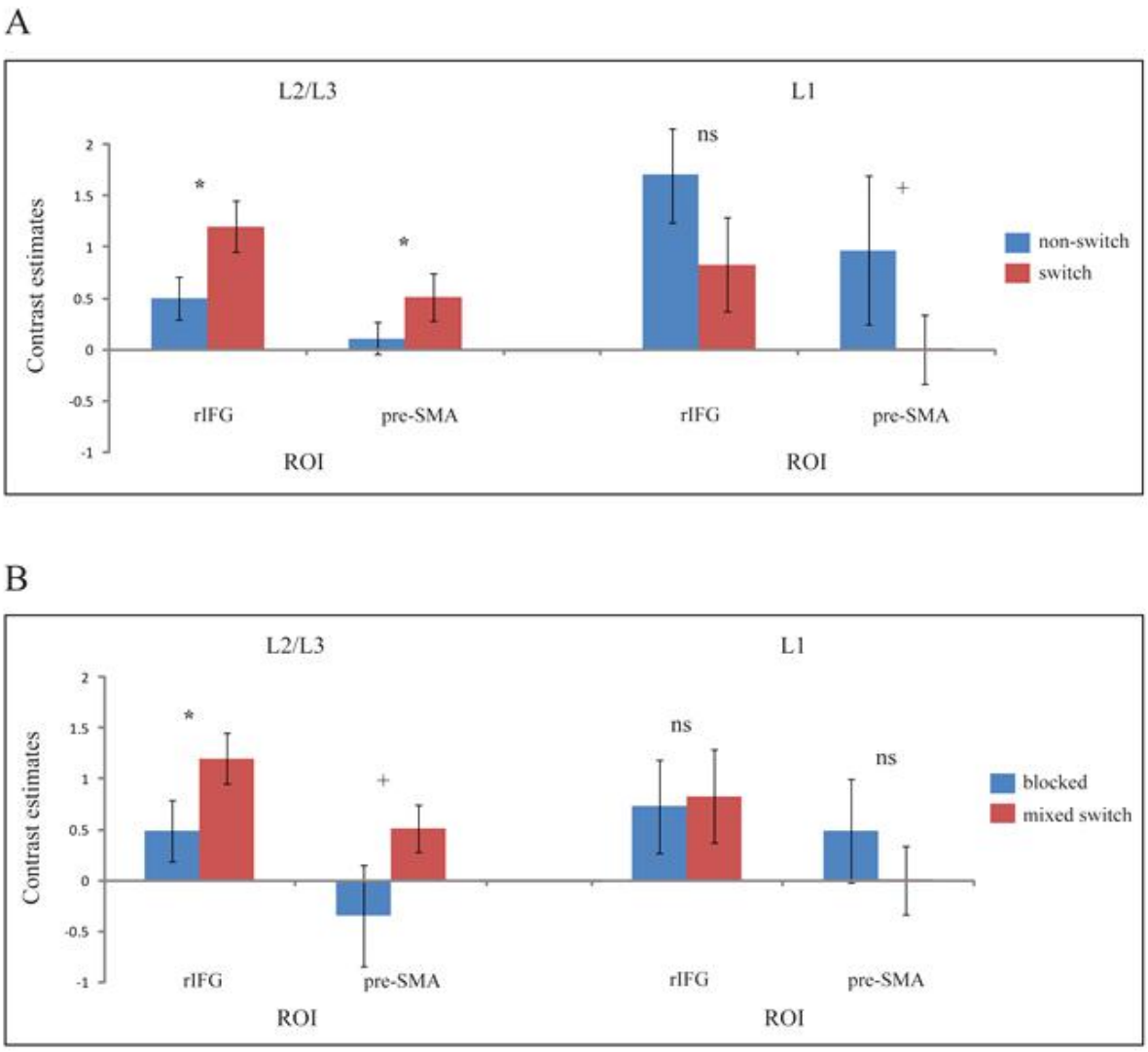


Fig. 4 Results from the ROI analysis. The graphs show the contrast estimates for the two ROIs (left: rIFG; right: pre-SMA). Error bars show the standard error of the mean. * : $p < .05$; +: $p < .1$; ns: $p > .1$. **A.** Contrast estimates for switch trials compared to non-switch trials. The left panel depicts the difference for L2/L3; the right panel for L1. **B.** Contrast estimates for mixed switch trials compared to blocked trials. The left panel depicts the difference for the L2/L3; the right panel for the L1.

needs to be inhibited. To make sure that the effect is indeed driven by inhibition-related activity for L1 in particular, we conducted the same ANOVA with language and trial sequence, but now only included L1-L2 and L1-L3 switch trials. This analysis yielded the same results, with no main effect of language or trial sequence ($p > .05$ in rIFG and pre-SMA), but again a significant interaction was found for the rIFG, $F(16) = 13.30$, $p = .002$, $\eta p2 = .45$, and the pre-SMA, $F(16) = 14.98$, $p = .001$, $\eta p2 = .48$.

We then analysed the differences between L1 and L2/L3 for the non-switch and switch trials separately. For the non-switch trials only, there was no effect of language ($p > .05$ in rIFG and pre-SMA). For the switch trials only, there was a main effect of language. Switches to L2/L3 compared to switches to L1 showed more activation in the rIFG, $F(16) = 5.68$, $p = .030$, $\eta p2 = .26$, and in the pre-SMA, $F(16) = 6.56$, $p = .021$, $\eta p2 = .29$. Compared to L1 switching, there was thus more activity in inhibition-related areas in switching to the non-native languages.

3.3.1.2 Mixed versus blocked

Following the definition of local inhibition by Guo et al. (2011), we first compared the entire mixed context (including switch and non-switch trials) to the blocked context. A repeated measures ANOVA with language (L1, L2/L3) and context (mixed, blocked) revealed no significant main effect of language ($p > .05$ in rIFG and pre-SMA) and a marginally significant effect of context in the rIFG, $F(16) = 4.30$, $p = .06$, but not in the pre-SMA, $F(16) = 3.37$, $p = .10$. There was no significant interaction between language and context ($p > .05$).

We then compared switch trials only to blocked naming to test for local inhibition. Again, there was also no significant main effect of language or context ($p > .05$ in rIFG and pre-SMA). However, there was a marginally significant interaction between language and context in the rIFG, $F(16) = 3.37$, $p = .065$, $\eta p2 = .17$, and a significant interaction in the pre-SMA, $F(16) = 7.31$, $p = .016$, $\eta p2 = .31$. This interaction suggests a difference between languages in terms of inhibition in switch trials compared to blocked trials.

We therefore analysed the differences between switch trials and blocked trials separately for L2/L3 and L1 naming. L2/L3 mixed switch trials compared to L2/L3 blocked trials showed significantly more activation in the rIFG, $F(16) = 4.10$, $p = .045$, $\eta p2 = .27$, and a marginally significant effect in the pre-SMA, $F(16) = 4.00$, $p = .050$, $\eta p2 = .26$ (Fig. 4B). Thus, there was more activation in the inhibition areas when switching to L2 and L3, compared to blocked naming in L2 and L3. Comparing L1 switch trials to L1 blocked trials did not yield significant differences in either the rIFG, $F(16) = .11$, $p = .74$, or the pre-SMA, $F(16) = .83$, $p = .38$. This suggests that, compared to blocked naming, there was more inhibition-related activity in switch trials to L2 and L3. However, for L1, the same amount of inhibition-related activity was found for mixed switching and blocked naming.

Again, we tested L1-L2 and L1-L3 switch trials only to make sure that the interaction was driven by inhibition of L1 in particular. This model yielded the same results, with no main effect of language or trial sequence ($p > .05$ in rIFG and pre-SMA), but again a significant interaction for the rIFG, $F(16) = 6.63$, $p = .020$, $\eta p2 = .29$, and the pre-SMA, $F(16) = 11.37$, $p = .004$, $\eta p2 = .42$.

3.3.2 Whole-brain analysis

3.3.2.1 Switch versus non-switch trials

Taking all three languages together, there was no brain area showing a main effect of switch trials compared to non-switch trials on FWE corrected $p < .05$. Relative to non-switches in L2/L3, however, switching to L2/L3 showed increased activation in the left precuneus (Brodmann Area/BA7), left MCC (BA24), right PCC (BA23), right cuneus (BA17), and the right ACC (BA24) (Table 2, Fig. 5A). Switching to L1 compared to non-switches in L1 did not show significant activation differences.

We also found a main effect of language in several brain areas when we compared L2/L3 naming to L1 naming (see Fig. 5B). Relative to naming in L1, naming in L2 and L3 activated the left IFG (BA45); left pre-SMA (BA6); right SMA (BA6); left pre-/postcentral gyrus (BA44/6); right Heschl's gyrus (BA41), right postcentral gyrus (BA6), and right insula (BA13); bilateral MCC (BA6); left calcarine gyrus (BA17) and cerebellum; right inferior occipital gyrus (BA19), bilateral putamen and right caudate nucleus; and right cerebellum. Relative to L2 and L3, naming in L1 did not show significant activation differences.

3.3.2.2 Mixed versus blocked

Taking the switch and non-switch trials together, the mixed context compared to the blocked context did not show significant activation differences in any brain area on FWE corrected $p < .05$. Similarly to the ROI analysis, we then compared the switch trials to blocked naming in L2/L3, which showed a

Table 2. Brain regions activated when contrasting switch trials with non-switch trials ($p < .05$, $k > 10$ voxels, FWE corrected). Multiple peaks in different brain regions within a single activation cluster are shown indented; Z refers to the highest Z score within that region. MCC = middle cingulate cortex; PCC = posterior cingulate cortex; ACC = anterior cingulate cortex.

Brain region	BA	Cluster size	MNI coordinates			Z value
			x	y	z	
L2/L3 switches > L2/L3 non-switches						
Precuneus, cingulate cortex		1251				
L precuneus	7		-8	-74	36	4.71
L MCC	24		-4	-34	38	4.41
R PCC	23		4	-28	28	4.65
R cuneus	17		22	-66	24	4.71
R ACC	24	242	4	24	28	4.72
L1 switches > L1 non-switches						
No significant activation differences						

Table 3. Brain regions activated when contrasting switch trials with blocked naming trials ($p < .05$, $k > 10$ voxels, FWE corrected). Z refers to the highest Z score within that region. MFG = middle frontal gyrus; ACC = anterior cingulate cortex

Brain region	BA	Cluster size	MNI coordinates			Z value
			x	y	z	
L2/L3 mixed switches > L2/L3 blocked naming						
R MFG	46	13	38	52	20	4.46
R ACC	24	41	4	42	12	5.70
L precuneus	7	23	-8	-66	32	4.71
L1 mixed switches > L1 blocked naming No significant activation differences						

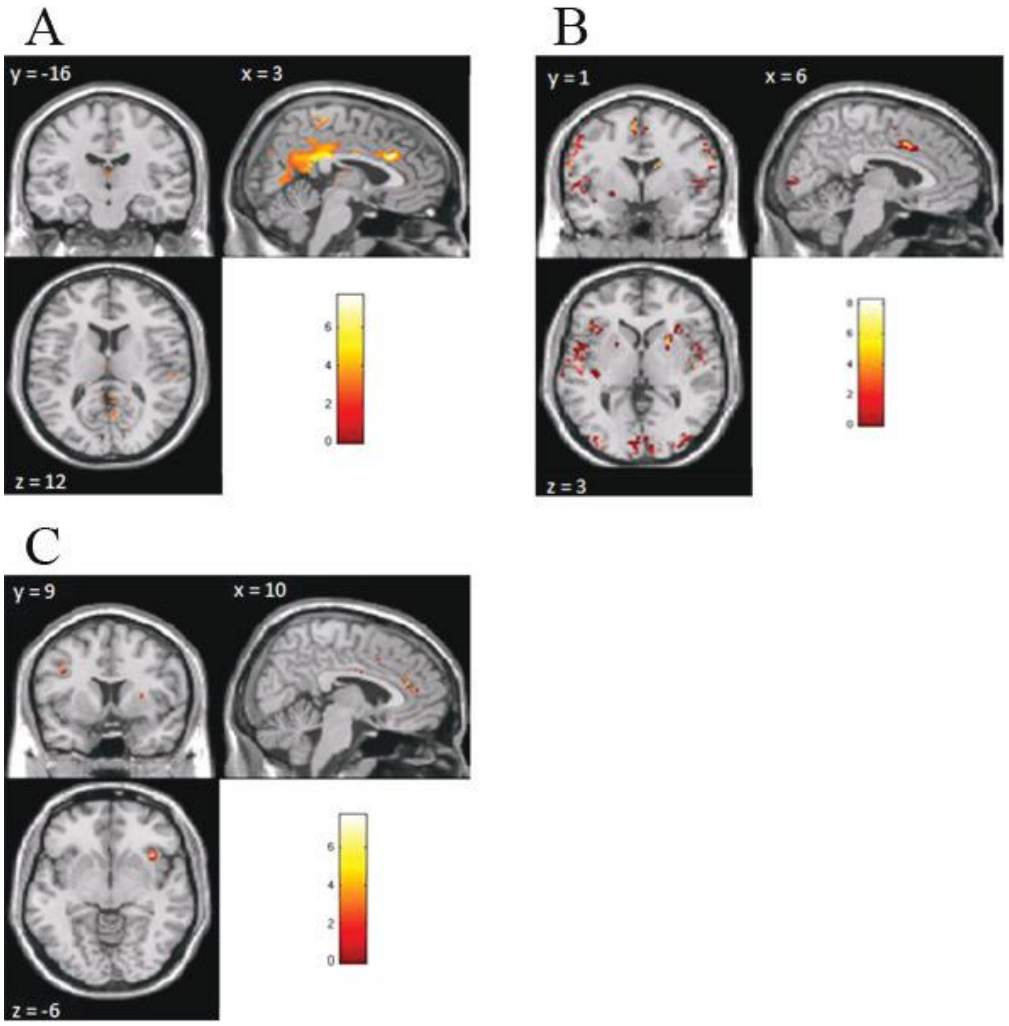


Fig. 5 Whole-brain analysis. **A.** Switches to L2/L3 compared to L2/L3 non-switch trials. **B.** Mixed naming in L2/L3 compared to mixed naming in L1. (C). Mixed switches to L2/L3 compared to blocked naming in L2/L3. Colour bars indicate Z-score.

significant effect in the right MFG (BA46), the right ACC (BA24), and the left precuneus (BA7) (Table 3, Fig. 5C). Switching to L1 in mixed naming compared to L1 blocked naming did not show significant activation differences.

3.3.3 Differences between L2 and L3

Besides the comparison between native and non-native languages, we also compared L2 and L3 in both the ROI analysis and the whole-brain analysis.

We did not observe any differences between L2 and L3, either when comparing switch to non-switch trials or when comparing mixed to blocked naming (all p values > 0.05 for both rIFG and pre-SMA). In the whole-brain analysis, we did not even observe differences between L2 and L3 when we lowered the threshold to uncorrected $p < 0.001$.

3.3.4 Simon costs

We also analysed the correlation between Simon costs and activation in the inhibition areas from the ROI analysis. There was a trend towards a negative correlation between the Simon costs and activation in the inhibition areas for switch costs (switch trials – non-switch trials) in L2 and L3 (rIFG, $r = -.28$, $p = .14$; pre-SMA, $r = -.42$, $p = .055$). The lower the Simon costs (reflecting better inhibition skills), the more activation was found in the rIFG and pre-SMA on L2/L3 switch compared to non-switch trials.

4. Discussion

The prevailing theory of language switching states that unbalanced bilinguals use inhibition to switch between their languages (Green, 1998). The present fMRI study investigated whether the brain mechanisms underlying trilingual language switching are indeed inhibitory or are rather related to other aspects of executive control. To test this, unbalanced trilinguals performed a picture naming task in the MRI scanner. Our results provide evidence that language switching recruits brain areas related to inhibition as well as areas associated with non-inhibitory aspects of executive control. We first discuss the behavioural results and the fMRI data for switch versus non-switch trials and for mixed versus blocked naming. Together, these results shed more light on the role of domain-general inhibition areas in language switching. Second, we focus on the relationship between language switching and the Simon task. Third, we address switching to L2 versus to L3. Finally, we discuss the main theoretical implications.

4.1 Domain-general inhibition areas in language switching

4.1.1 Inhibition in switching versus non-switching

Within the mixed condition, we compared switch versus non-switch trials. The behavioural data showed that non-switch trials were named faster than switch trials, but we did not observe a clear asymmetry in switch costs, contrary to other studies with unbalanced bilinguals (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999). The absence of asymmetrical switch costs is often taken as evidence for the absence of inhibition too. In our fMRI data, however, we find evidence suggesting that language switching is, at least partly, achieved by inhibitory mechanisms. We found more activation in the rIFG and pre-SMA in switches to the weaker L2 and L3 compared to non-switches, but this difference was absent for L1. The rIFG and pre-SMA are often reported to be involved in inhibition (e.g., Aron et al., 2004a, 2004b; Jahfari et al., 2011). Differences in activation in these areas are therefore suggested to reflect differences in inhibition. The activation differences for L2/L3 indicate that switches to weaker, non-native languages require more inhibition-related activity than non-switch trials. The absence of such a difference for L1 shows that L1 switch and non-switch trials are accompanied by an equal amount of activation in areas related to inhibition. These findings are compatible with Green's IC model (1998), as inhibition seems to be modulated by the strength of the language. Picture naming in the strong L1 does not require inhibition of the weaker languages, regardless of the trial type. For weaker languages, however, more inhibition of L1 is needed for switch trials than for non-switch trials.

Thus, we find an effect of switching in inhibition-related brain areas in the absence of clear asymmetrical behavioural switch costs. This suggests that (the absence of) behavioural asymmetrical switch costs cannot be taken as a reliable indicator of inhibition. The conclusions of Costa and Santesteban (2004) are therefore not compatible with our findings. Based on symmetrical switch costs, they concluded that balanced bilinguals do not use inhibition. Our study show otherwise: Even in the absence of asymmetrical switch costs, language switching can still involve inhibition. Contrary to our predictions, however, switch trials in general did not show more activation in the rIFG and pre-SMA than non-switch trials. The absence of this effect is likely to be related to L1 non-switch trials. Whereas L2/L3 switch trials showed more activation than non-switch trials, there was also slightly more activation in L1 non-switch trials than in switch trials. It is likely that these opposing effects for L2/L3 and L1 prevented the difference between switch and non-switch trials in general from reaching significance.

Our results indicate that domain-general

inhibition areas are involved in language switching, particularly in switches to non-native languages. Other brain areas related to different aspects of executive control were activated too. For switches to L2/L3 compared to L2/L3 non-switch trials, we observed a difference in activation in the ACC and in a cluster comprising the precuneus, cuneus, MCC, and PCC. Activation in the ACC likely reflects the increase in non-inhibitory aspects of executive control that are needed for switch trials and is often reported in language switching (e.g., Hosoda et al., 2012; Wang et al., 2007). Furthermore, this area is a vital part of the language switching network proposed by Abutalebi and Green (2008). The (pre) cuneus is also often found in language switching studies (Guo et al., 2011; Wang et al., 2007), but its role remains unclear. The MCC and PCC have been found in various cognitive tasks too and have been linked to response selection and error detection (cf., Torta & Cauda, 2011).

Within the mixed condition, we also compared L2/L3 naming to L1 naming. Within the switch trials, we observed more activation in the rIFG and pre-SMA for switches to L2/L3 compared to switches to L1, suggesting that switching to L2 and L3 requires more inhibition than switching to L1. We also observed other differences in brain activation between L2/L3 naming and L1 naming. Naming in L2/L3 versus L1 naming showed activation in various brain regions, including the left IFG; bilateral (pre-)SMA; left pre-/postcentral gyrus; bilateral MCC; left calcarine gyrus and right inferior occipital gyrus; bilateral putamen and right caudate nucleus; and bilateral cerebellum. Activation in the precentral gyrus is often found in switching to a weaker language (Hernandez et al., 2000; Wang et al., 2007) and might be associated with phonological retrieval or encoding (Indefrey & Levelt, 2004). Similarly, the increased activation in the cerebellum might reflect a greater need for articulatory control in L2/L3 versus L1 (Booth et al., 2007). The putamen and caudate nucleus are part of the language switching network proposed by Abutalebi and Green (2008). The putamen has been suggested to be involved in motor control of weaker languages (Abutalebi et al., 2013b). However, the role of the right caudate nucleus in language switching has been debated. Most studies (e.g., Abutalebi et al., 2008; Crinion et al., 2006) report increased activation in the left or bilateral caudate nucleus during language switching. Wang et al. (2007), on the other hand, also found right lateralized activation. The present study supports the latter finding that the right caudate nucleus might be involved in language switching too.

Summarizing, our data show that switching compared to non-switching in a mixed context recruits areas related to inhibition and areas related to other aspects of executive control, like the ACC. These areas show an increase in activation for switches to the weaker L2 and L3 compared to non-switch trials, but no such difference was found for the L1. Similarly, more activation in both inhibitory and non-inhibitory executive control areas was found for switches to L2/L3 compared to switches to L1.

4.1.2 Inhibition in mixed versus blocked context

The behavioural data showed a difference between blocked and mixed naming across the three languages. In the blocked context, naming was faster and more accurate in L1 compared to L2/L3. In the mixed context, this effect disappeared: Naming was equally fast in all three languages. The difference between mixed and blocked naming was larger for L1 than for L2/L3. This replicates other behavioural results (e.g., Christoffels et al., 2007) and is also compatible with the IC model, because it suggests that the L1 is inhibited more than the L2 and L3 in the mixed context compared to the blocked context. This is also compatible with our fMRI data that revealed more activation in the rIFG and pre-SMA for switches to L2/L3 compared to blocked trials, but no difference for the L1. This difference suggests that, relative to blocked naming, switches to weaker languages require more inhibition. These findings again support the IC model. In the blocked context, (relatively) little or no local inhibition is required. Similarly, little or no inhibition is expected for switches to L1 in the mixed context. For switches to L2 and L3, however, more local inhibition is needed, thus leading to differences between the blocked and the mixed switch trials. Again, we did not only observe differences in activation in inhibition areas, but also in other areas related to executive control. The whole-brain analysis showed increased activation in the right MFG, right ACC, and left precuneus for L2/L3 mixed compared to blocked trials. The MFG and ACC are included in the frontoparietal network that is linked to various aspects of executive control (Barbey et al., 2012; Collette et al., 2005; Niendam et al., 2012). This, again, reflects an increased need for executive control in L2/L3 trials in the mixed naming context. Contrary to our hypothesis, we did not observe a main effect of context in the rIFG and pre-SMA: In general, mixed naming did not require more inhibition than blocked naming. Still, there was

a trend in this direction. We suspect that the large amount of variation across participants prevented this effect from reaching significance. Furthermore, the blocked and mixed contexts were separated by the T1 scan, which might make a direct comparison less reliable.

In summary, our results show that switching to L2/L3 in the mixed context requires more activation of inhibition-related areas than L2/L3 blocked naming. No such difference was observed for the L1. Switching to L2/L3 compared to blocked naming also showed an increase in activation in areas related to non-inhibitory aspects of executive control, such as the ACC and MFG.

4.2 Relationship between language switching and Simon task

Our results revealed a correlation between Simon costs (reflecting inhibition skills,Van den Wildenberg et al., 2010) and both behavioural language switch costs as well as activation in the rIFG and pre-SMA. The correlation between Simon costs and L1 switch costs was negative: Better inhibition skills were associated with larger switch costs to the stronger language. In contrast, the correlation between Simon costs and L2/L3 switch costs was positive: Better inhibition skills were associated with smaller switch costs to weaker languages. Furthermore, there was a negative correlation between Simon costs and the amount of activity in the rIFG and pre-SMA in switches to L2/L3. This suggests that participants with better inhibition skills inhibit the L1 more when switching to L2 and L3. This results in smaller switch costs to L2 and L3, but in larger switch costs to L1.

4.3 Switching to L2 versus switching to L3

Our study did not find a difference between L2 and L3 naming. This is contrary to, for example, Schwieter (2013), who found an asymmetry in L2 and L3 behavioural switch costs. We also expected to find a difference between L2 and L3, with more L1 inhibition and larger switch costs for the weaker L3. The difference in proficiency between L2 and L3 in our participants might have been too small to cause an effect. Furthermore, the difference in Age of Acquisition was relatively small (2.5 years between L2 and L3) compared to other studies (e.g., Schwieter, 2013, reported a difference of 10.4 years). This small difference between L2 and L3 might have resulted in the absence of a proficiency effect in our study.

4.4 Theoretical implications

To summarize, our data show more activation in the rIFG and pre-SMA for switches to L2 and L3 compared to non-switch and blocked trials. No such differences were found for the L1. This suggests that language switching recruits domain-general inhibition areas, especially when switching to a weaker language. We furthermore observed correlations between language switching costs and Simon costs, also suggesting that language switching performance is linked to domain-general inhibition. Inhibition areas alone, however, are not sufficient. Switching to L2 and L3 also recruited brain areas related to other aspects of executive control, such as the ACC, DLPFC, and striatum.

Our results thus show inhibition-related activity during language switching, which is modulated by language proficiency (i.e., L1 versus L2/L3). These findings are compatible with Green's IC model (1998). When participants have to switch to a weaker L2 and L3, the L1 has to be inhibited. This is reflected in an increased activation in the rIFG and pre-SMA for switches to L2/L3 compared to L2/L3 non-switch trials or L2/L3 blocked naming. Furthermore, this is reflected in the increased naming latencies in L1 in the mixed context compared to the blocked context. In the blocked context, the L1 is inhibited less and therefore L1 naming is faster. The relatively longer L1 naming latencies in the mixed condition are likely to reflect the time needed to overcome inhibition. Taken together, we thus obtained evidence supporting the two main predictions of the IC model (Green, 1998). First, our fMRI data suggest that switches to weaker languages require more inhibition of the strong L1 than vice versa. Second, our behavioural data partly suggest that it takes time to overcome this inhibition: In mixed versus blocked naming, but not in switch versus non-switch trials, we found larger costs for L1 than for L2 and L3. Together, these data support the theory that unbalanced bilinguals use inhibition in language switching.

Inhibition alone, however, is not sufficient for language switching. Abutalebi and Green (2008) proposed that a domain-general network for other aspects of executive control is also recruited during bilingual language switching. Our results are compatible with this model, because they indicate that areas like the ACC, striatum, and bilateral frontal cortices are indeed involved in language switching. This is also compatible with studies showing that the frontoparietal network represents different task features, such as colour cues, rules, individual stimuli,

and responses (Woolgar et al., 2011). Our study is thus in line with the model proposed by Abutalebi and Green. We also argue, however, that both this model and Luk's meta-analysis (2012) do not give a complete picture of the brain networks involved in language switching. Besides many areas related to non-inhibitory executive control, inhibition areas are also recruited during language switching. These areas that are involved in inhibition are often not included in brain models of language switching. Our study, however, shows that these models should include, and distinguish between, areas related to inhibition in particular and to non-inhibitory aspects of executive control.

5. Conclusions

To our knowledge, our study is the first to specifically examine the involvement of domain-general inhibition areas in language switching. Our results indicate that language switching not only recruits brain regions that instantiate non-inhibitory aspects of executive control, but also areas related to inhibition in particular. This suggests that unbalanced bilinguals use inhibition in language switching, especially when they have to switch to their weaker languages. In this way, our study provides new and important neuroimaging evidence for the long-standing claim of the involvement of inhibition during language switching.

6. References

Aarts, E., Roelofs, A., & van Turennout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *The Journal of Neuroscience, 28*(18), 4671-4678.

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., & Khateb, A. (2008). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cerebral Cortex, 18*(7), 1496-1505.

Abutalebi, J., Della Rosa, P. A. D., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013a). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex, 49*(3), 905-911.

Abutalebi, J., Della Rosa, P. A. D., Castro Gonzaga, A. K., Keim, R., Costa, A., & Perani, D. (2013b). The role of the left putamen in multilingual language production. *Brain and Language, 125*(3), 307-315.

Abutalebi, D. J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes, 23*(4), 557-582.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. Neuroscientist, *13*(3), 214-228.

Aron, A. R., Monsell, S., Sahakian, B. J., & Robbins, T. W. (2004a). A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain, 127*, 1561-1573.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004b). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences, 8*(4), 170-177.

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain, 135*(4), 1154-1164.

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*, 290-303.

Brett, M., Anton, J., Valabregue, R., & Poline, J. (2002, June). *Region of interest analysis using an SPM toolbox.* Poster presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Research, 1133*, 136-144.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research, 1147*, 192-208.

Collette, F., Van der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., & Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Human Brain Mapping, 25*, 409-423.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence 30*(2), 163-183.

Cools, R. (2011). Dopaminergic control of the striatum for high-level cognition. *Current Opinion in Neurobiology, 21*, 402-407.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491-511.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., ... & Price, C. J. (2006). Language control in the bilingual brain. *Science, 312*(5779), 1537-1540.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P., & Ridderinkhof, K. R. (2008). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *Journal of Neuroscience, 28*(39), 9790-9796.

Frank, M.J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Current Opinion in Neurobiology, 21*, 381-386.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(3), 640.

Graybiel, A. M. (2000). The basal ganglia. *Current Biology, 10*(14), R509-R511.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Bilingingualism, 1*, 67-81.

Guo, T., Liu, H., Misra, M., & Kroll, J.F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *NeuroImage, 56*, 2300-2309.

Hernandez, A.E., Martinez, A., & Kohnert, K. (2000). In search of the language switch: An fMRI- study of picture naming in Spanish-English bilinguals. *Brain and Language, 73*, 421-431.

Hosoda, C., Hanakawa, T., Nariai, T., Ohno, K., & Honda, M. (2012). Neural mechanisms of language switch. *Journal of Neurolinguistics, 25*, 44-61.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1), 101-144.

Jackson, G.M., Swainson, R., Cunnington, R., & Jackson, S.R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism, 4*(2), 169-178.

Jahfari, S., Waldorp, L., van den Wildenberg, W. P., Scholte, H. S., Ridderinkhof, K. R., & Forstmann, B. U. (2011). Effective connectivity reveals important roles for both the hyperdirect (fronto-subthalamic) and the indirect (fronto-striatal-pallidal) fronto-basal ganglia pathways during response inhibition. *Journal of Neuroscience, 31*(18), 6891-6899.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). The Boston NamingTest. Philadelphia: Lea & Febiger.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*(5660), 1023-1026.

Lemhöfer, K., & Broersma, M. (2011, advance online publication). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*.

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes, 27*(10), 1479-1488.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25-40.

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience, 12*(2), 241-268.

Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic resonance in medicine, 55*(6), 1227-1235.

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism, 13*(02), 253-262.

Schwieter, J. W. (2013). Lexical inhibition in trilingual speakers. In J. Tirkkonen & E. Anttikoski (Eds.), *Proceedings of The 24th Conference of Scandinavian Linguistics. Publications of the University of Eastern Finland: Reports and Studies in Education, Humanities, and Theology* (pp. 249-260). Joensuu, Finland: University of Eastern Finland Press.

Schwieter, J. W., & Sunderman, G. (2011). Inhibitory control processes and lexical access in trilingual speech production. *Linguistic Approaches to Bilingualism, 1*(4), 391-412.

Torta, D. M., & Cauda, F. (2011). Different functions in the cingulate cortex, a meta-analytic connectivity modeling study. *NeuroImage, 56*(4), 2157-2172.

Van den Wildenberg, W. P., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: a review. *Frontiers in human neuroscience, 4*.

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods, 40*(2), 479-483.

Verhoef, K., Roelofs, A., & Chwilla, D. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition, 110*, 84-99.

Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *NeuroImage, 35*, 862-870.

Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *NeuroImage, 56*(2), 744-752.

Xue, G., Aron, A. R., & Poldrack, R. A. (2008). Common neural substrates for inhibition of spoken and manual responses. *Cerebral Cortex, 18*, 1923-1932.

Unsupervised Feature Learning Improves Prediction of Human Brain Activity in Response to Natural Images

Umut Güçlü[1]
Supervisor: Marcel A. J. van Gerven[1]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands*

Encoding and decoding in functional magnetic resonance imaging has recently emerged as an established area of research to noninvasively characterize the relationship between stimulus features and human brain activity. To overcome the challenge of formalizing what stimulus features should modulate single voxel responses, we introduce a general framework for making directly testable predictions of single voxel responses to statistically adapted representations of ecologically valid stimuli. These representations are learned from unlabeled data without any supervisory signals. To validate our framework, we develop a parsimonious computational model of (i) how early visual cortical representations are adapted to statistical regularities in natural images and (ii) how populations of these representations are pooled by single voxels. We use our model to predict single voxel responses to natural images and identify natural images from stimulus-evoked multiple voxel responses. We show that statistically adapted low-level sparse and invariant representations of natural images better span the space of early visual cortical representations and can be more effectively exploited in the supervised learning task of identification than Gabor wavelets. Our results demonstrate the potential of our framework to probe unknown visual cortical representations.

Keywords: decoding, encoding, functional magnetic resonance imaging, natural image statistics, unsupervised feature learning, visual population codes

Corresponding author: Umut Güçlü; E-mail: u.gulcu@donders.ru.nl

1. Introduction

An important goal of contemporary cognitive neuroscience is to characterize the relationship between stimulus features and human brain activity. This relationship can be studied from two distinct but complementary perspectives of encoding and decoding (Dayan & Abbott, 2005). The encoding perspective is concerned with how certain aspects of the environment are stored in the brain and uses models that predict brain activity in response to certain stimulus features. Conversely, the decoding perspective uses models that predict specific stimulus features from stimulus-evoked brain activity and is concerned with how specific aspects of the environment are retrieved from the brain.

The stimulus-response relationship has been extensively studied in computational neuroscience to understand the information contained in individual or ensemble neuronal responses, based on different coding schemes (Brown, Kass, & Mitra, 2004). The invasive nature of the measurement techniques of these studies has restricted human subjects to particular patient populations (Pasley et al., 2012; Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). However, with the advent of functional magnetic resonance imaging (fMRI), encoding and decoding has emerged as an established area of research to noninvasively characterize the relationship between stimulus features and human brain activity via localized changes in blood-oxygen-level-dependent (BOLD) hemodynamic responses to sensory or cognitive stimulation (Naselaris, Kay, Nishimoto, & Gallant, 2011).

Encoding models that predict single voxel responses to certain stimulus features typically comprise two main components. The first component is a (non)linear transformation from a stimulus space to a feature space. The second component is a (non)linear transformation from the feature space to a voxel space. Encoding models can be used to test alternative hypotheses about what a voxel represents since any encoding model embodies a specific hypothesis about what stimulus features modulate the response of the voxel (Naselaris et al., 2011). Furthermore, encoding models can be converted to decoding models that predict specific stimulus features from stimulus-evoked multiple voxel responses. In particular, decoding models can be used to determine the specific class from which the stimulus was drawn (i.e. classification) (Haxby et al., 2001; Kamitani & Tong, 2005), identify the correct stimulus from a set of novel stimuli (i.e. identification) (Kay, Naselaris, Prenger, & Gallant, 2008; Mitchell et al., 2008) or create a literal picture of the stimulus (i.e. reconstruction) (Miyawaki et al., 2008; Schoenmakers, Barth, Heskes, & van Gerven, 2013; Thirion et al., 2006).

The conventional approach to encoding and decoding is explanatory and descriptive modeling of feature spaces. As such, feature spaces are typically "hand-designed" by theorists or experimentalists (Kay et al., 2008; Kay, Winawer, Rokem, Mezer, & Wandell, 2013; Mitchell et al., 2008; Miyawaki et al., 2008; Naselaris, Prenger, Kay, Oliver, & Gallant, 2009; Nishimoto et al., 2011; Vu et al., 2011). However, this approach is a slow and laborious process that is prone to the influence of subjective expectations and restricted to a priori hypotheses. As a result, it severely restricts the scope of alternative hypotheses that can be formulated about what a voxel represents as evident by the lack of models that adequately characterize extrastriate visual cortical voxels.

A recent trend in models of visual population codes has been the adoption of natural images for the characterization of the voxels that respond to visual stimulation (Kay et al., 2008; Naselaris et al., 2009). The motivation behind this trend is that natural images admit multiple feature spaces of low-level edges, mid-level edge junctions, high-level object parts and complete objects that can modulate single voxel responses (Naselaris et al., 2011). Implicit about this motivation is the assumption that the brain is adapted to the statistical regularities in the environment (Barlow, 1961) such as those in natural images (Bell & Sejnowski, 1997; Olshausen & Field, 1996). At the same time, recent developments in theoretical neuroscience and machine learning have shown that predictive and normative models of natural image statistics learn statistically adapted representations of natural images. As a result, they predict statistically adapted visual cortical representations, based on different coding principles. Some of these predictions were shown to be similar to what is found in the primary visual cortex such as topographically organized simple and complex cell receptive fields (Hyvärinen, 2010).

Building on previous studies of visual population codes and natural image statistics, we introduce a general framework to make directly testable predictions of single voxel responses to statistically adapted representations of ecologically valid stimuli. To validate our framework, we develop a parsimonious computational model that comprises two main components (Fig. 1A). The first component is a nonlinear feature model

that transforms raw stimuli to stimulus features. The feature model learns the transformation from unlabeled data without any supervisory signals. The second component is a linear voxel model that transforms the stimulus features to voxel responses. The voxel model learns the transformation from feature-transformed stimulus-response pairs. We use an fMRI data set to show that the encoding and decoding performance of our model is significantly better than that of a Gabor wavelet pyramid (GWP) model of phase-invariant complex cells.

2. Materials and methods

2.1 Data

We used the data set (Kay, Naselaris, & Gallant, 2011) that was originally published in Kay et al. (2008) and Naselaris et al. (2009). Briefly, the data set contains 1750 and 120 stimulus-response pairs of two subjects (i.e. S1 and S2) in the estimation and validation sets, respectively. The stimulus-response pairs are grayscale natural images of size 128 × 128 pixels and stimulus-evoked peak BOLD hemodynamic responses of 5512 (S1) and 5275 (S2) voxels in the visual areas V1, V2 and V3. The details of the experimental procedures are presented in Kay et al. (2008).

2.2 Problem statement

2.2.1 Encoding

Let $x \in R^d$ and $\boldsymbol{y} \in \mathbb{R}^q$ be a stimulus-response pair where x is a vector of pixels in a grayscale natural image, and y is a vector of voxel responses. Given x, we are interested in the problem of predicting y:

$$\hat{y} = arg\ \max_{y} p(y \vee \phi(x))$$

$$\hat{\boldsymbol{y}} = \boldsymbol{B}^T \boldsymbol{\phi}(\boldsymbol{x})$$

where $\hat{y}$ is the predicted response to x, and p is the encoding distribution of y given $\boldsymbol{\phi}(\boldsymbol{x})$. $\boldsymbol{\phi}$ nonlinearly transforms x from the stimulus space to the feature space, and **B** linearly transforms $\boldsymbol{\phi}(\boldsymbol{x})$ from the feature space to the voxel space.

2.2.2 Decoding

Let $\mathbb{X}$ be a set of images that contains x. Given $\mathbb{X}$ and y, we are interested in the problem of identifying x:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x} \in \mathbb{X}} \rho_{\boldsymbol{y}, \boldsymbol{B}^T \boldsymbol{\phi}(x)}$$

where $\hat{\boldsymbol{x}}$ is the identified image from y, and ρ is the Pearson product-moment correlation coefficient between y and $\boldsymbol{B}^T\boldsymbol{\phi}(x)$.

Solving the encoding and decoding problems requires the definition and estimation of a feature model ϕ followed by a voxel model **B**.

2.3 Feature model

2.3.1 Model definition

Following Hyvärinen and Hoyer (2001), we start by defining a single-layer statistical generative model of grayscale natural image patches. Assuming that a patch is generated by a linear superposition of latent variables that are non-Gaussian (in particular, sparse) and mutually independent, we first use independent component analysis to define the model by a linear transformation of the independent components of the patch:

$$\boldsymbol{z} = \boldsymbol{A}\boldsymbol{s}$$

where $\boldsymbol{z} \in \mathbb{R}^n$ is a vector of pixels in the whitened patch, $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ is the mixing matrix, and $\boldsymbol{s} \in \mathbb{R}^m$ is a vector of the components of z such that $m \leq n$ We then define s by inverting the linear system defined by **A**:

$$\boldsymbol{s} = \boldsymbol{W}\boldsymbol{z}$$

where $\boldsymbol{W} \in \mathbb{R}^{m \times n}$ is the unmixing matrix such that $\boldsymbol{W} = \boldsymbol{A}^{-1}$. We constrain $\mathbf{W}$ to be orthonormal and $\boldsymbol{s}_i$ to have unit variance such that $\boldsymbol{s}_i$ are uncorrelated and unique, up to a multiplicative sign. Next, we define the joint log-probability of s by the sum of the marginal log-probabilities of $\boldsymbol{s}_i$ since we assumed that the $\boldsymbol{s}_i$ are independent:

$$\log p(\boldsymbol{s}) = \sum_{i=1}^{m} \log p(s_i)$$

where $p(s_i)$ are peaked at zero and have high kurtosis since we assumed that the $\boldsymbol{s}_i$ are sparse.

While one of the assumptions of the model is that the $\boldsymbol{s}_i$ are independent, their estimates are only maximally independent. As a result, residual dependencies remain between the estimates of $\boldsymbol{s}_i$. We continue by modeling the nonlinear correlations of $\boldsymbol{s}_i$ since we constrained $\boldsymbol{s}_i$ to be linearly uncorrelated. In particular, we assume that the locally pooled

energies of s_i are sparse. Without loss of generality, we first arrange the s_i on a square grid graph that has circular boundary conditions. We then define the locally pooled energies of s_i by the sum of the energies of s_i that are in the same neighborhood:

$$c = H\, s^2$$

where $c \in R^m$ is a vector of the locally pooled energies of s_i and $H \in R^{m \times m}$ is a neighborhood matrix such that $h_{i,j} = 1$ if c_i pools the energy of s_j and $h_{i,j} = 0$ otherwise. Next, we redefine $\log p(s)$ in terms of c to model both layers:

$$\log p(\boldsymbol{s}) = \sum_{i=1}^{m} G(c_i)$$

where G is a convex function. Concretely, we use $G(c_i) = -\log(1 + c_i)$.

In a neural interpretation, simple and complex cell responses can be defined as a static compressive nonlinear function of s and c, respectively. Concretely, we use $\log(1 + \boldsymbol{s})$ and $\log(1 + \boldsymbol{c})$ to define them after we estimate the model.

2.3.2 Model estimation

We use a modified gradient ascend method to estimate the model by maximizing the log-likelihood of $\mathbf{W}$ (equivalently, the sparseness of c) given a set of patches:

$$\widehat{\boldsymbol{W}} = \arg\max_{\boldsymbol{W}} L(\boldsymbol{W}|\mathbb{Z})$$

where $L(\boldsymbol{W}|\mathbb{Z}) = \sum_{\boldsymbol{z} \in \mathbb{Z}} \log p(\boldsymbol{H}(\boldsymbol{W}\boldsymbol{z})^2)$ is the log-likelihood and $\mathbf{Z}$ is the set of patches. At each iteration, we first find the gradient of $L(\boldsymbol{W}|\mathbb{Z})$:

$$\nabla_w L(\boldsymbol{W}|\mathbb{Z}) = -\boldsymbol{H}^T (1 + \boldsymbol{H}(\boldsymbol{W}\mathbb{Z})^2)^{-1} \circ (2\boldsymbol{W}\mathbb{Z})\mathbb{Z}^T$$

where $\circ$ is defined as the Hadamard (i.e. element-wise) product. We then project it onto the tangent space of the constrained space (Edelman, Arias, & Smith, 1998):

$$\bar{\nabla}_w L(\boldsymbol{W}|\mathbb{Z}) = \nabla_w L(\boldsymbol{W}|\mathbb{Z}) - \boldsymbol{W}\nabla_w L(\boldsymbol{W}|\mathbb{Z})^T \boldsymbol{W}$$

Next, we use backtracking line search to choose a step size by reducing it geometrically with a rate from (0,1) until the Armijo-Goldstein condition holds (Boyd & Vandenberghe, 2004). Finally, we update W and find the orthogonal matrix nearest it:

$$\boldsymbol{W} \leftarrow \boldsymbol{W} + \mu \bar{\nabla}_w L(\boldsymbol{W}|\mathbb{Z})$$

$$\boldsymbol{W} \leftarrow (\boldsymbol{W}\boldsymbol{W}^T)^{-\frac{1}{2}}\boldsymbol{W}$$

where μ is the step size.

2.4 Voxel model

2.4.1 Model definition

We start by defining a model for each voxel. Assuming that $p(\boldsymbol{y}|\phi(\boldsymbol{x})) \sim N(\boldsymbol{B}^T\phi(\boldsymbol{x}), \boldsymbol{\Sigma})$ where $\boldsymbol{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_q) \in \mathbb{R}^{m \times q}$ and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_q^2) \in \mathbb{R}^{q \times q}$, we use linear regression to define the models by a weighted sum of $\phi(\boldsymbol{x})$:

$$y_i = \boldsymbol{\beta}_i^T \phi(\boldsymbol{x}) + \varepsilon_i$$

where $\varepsilon_i \sim N(0, \sigma_i^2)$.

2.4.2 Model estimation

We estimate the model using ridge regression:

$$\widehat{\boldsymbol{\beta}}_i = \arg\min_{\boldsymbol{\beta}_i} \frac{1}{N} \sum_{y_j^i \in \boldsymbol{Y}} \left(y_j^i - \boldsymbol{\beta}_i^{\mathrm{T}} \phi(\boldsymbol{x}^j)\right) + \lambda_i \|\boldsymbol{\beta}_i\|_2^2$$

Where $\boldsymbol{Y} = (\boldsymbol{y}^1, \ldots, \boldsymbol{y}^N)^{\mathrm{T}} \in \mathbb{R}^{N \times q}$ are the responses and $\lambda_i \geq 0$ is a complexity parameter that controls the amount of regularization. Let $\boldsymbol{\Phi} = (\phi(\boldsymbol{x}^1), \ldots, \phi(\boldsymbol{x}^N))^{\mathrm{T}}$ denote the feature space representation of the input data an let $\boldsymbol{Y}_i$ denote all responses for a voxel i. We obtain $\widehat{\boldsymbol{\beta}}_i$ as:

$$\widehat{\boldsymbol{\beta}}_i = (\lambda_i \boldsymbol{I}_m + \boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T\boldsymbol{Y}_i$$

If $m \gg N$, we solve the problem in a new coordinate system in which only the first $\mathbf{N}$ coordinates of $\boldsymbol{\Phi}$ are nonzero (Murphy, 2012). We first factorize $\boldsymbol{\Phi}$ using singular value decomposition:

$$\boldsymbol{\Phi} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$$

where $\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{U} = \boldsymbol{I}_N, \boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_q^2) \in \mathbb{R}^{q \times q}$ and $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V} = \boldsymbol{I}_N$. The columns of $\mathbf{U}$, the diagonal entries of $\mathbf{S}$ and the columns of $\mathbf{V}$ are the left-singular vectors, the singular values and the right-singular vectors of $\boldsymbol{\Phi}$, respectively. We then obtain $\widehat{\boldsymbol{\beta}}_i$ as:

$$\widehat{\boldsymbol{\beta}}_i = \boldsymbol{V}\,\mathrm{diag}\left(\frac{\boldsymbol{s}}{\boldsymbol{s}^2 + \lambda_i}\right)\boldsymbol{U}^{\mathrm{T}}\boldsymbol{Y}_i$$

where division is defined element-wise. The rotation reduces the complexity of the problem from $O(N m^2)$ to $O(m N^2)$. To choose the optimal λ_i, we perform hyperparameter optimization using grid search guided by a generalized cross-validation approximation to leave-one-out cross-validation (Hastie, Tibshirani, & Friedman, 2009). We define a grid by sampling the effective degrees of freedom of the ridge regression fit in the range [1,…*N*] since the parameter space is bounded from above for the effective degrees of freedom. The effective degrees of freedom is defined as:

$$\mathrm{df}(\lambda_i) = \sum_{j=1}^{N} \frac{{s_j}^2}{{s_j}^2 + \lambda_i}$$

We use Newton's method to solve df for λ_i Once the grid is defined, we choose the optimal λ_i that minimizes the generalized cross-validation error:

$$\hat{\lambda}_i = \arg\min_{\lambda \in \Lambda} \left\{ \sum_{j=1}^{N} \left[\frac{y_i^j - \hat{y}_i^j(\lambda)}{1 - \frac{\mathrm{df}(\lambda)}{N}} \right]^2 \right\}$$

where $\hat{y}_i^j(\lambda)$ is the predicted response given a particular λ.

2.5 Encoding and decoding

In the case of the SC model, each (randomly sampled or non-overlapping) patch was transformed to its principal components (such that 625 components with the largest variance were retained) and whitened prior to model estimation and validation. Once the images were feature transformed, they were z-scored. The SC model of 625 simple and 625 complex cells was estimated from 50000 patches of size 32 × 32 pixels that were randomly sampled from the 1750 images of size 128 × 128 pixels in the estimation set. The details of the GWP model are presented in Kay et al. (2008). The SC1, SC2 and GWP2 models were estimated from the 1750 feature-transformed stimulus-response pairs in the estimation set.

Voxel responses to an image of size 128 × 128 pixels were predicted as follows. In the case of the SC model, each 16 non-overlapping patch of size 32 × 32 pixels of the image were first transformed to the simple (or complex) cell responses of the SC model (i.e. a total of 625 simple or 625 complex cell responses to each patch and 10000 simple and 10000 complex cell responses to the image). The 10000 simple (or 10000 complex) cell responses of the SC model were then transformed to the voxel responses of the SC1 (or SC2) model. In the case of the GWP model, the image was first transformed to the complex cell responses of the GWP model (i.e. a total of 10921 complex cell responses per image). The 10921 complex cell responses of the GWP model were then transformed to the voxel responses of the GWP2 model. Encoding performance was quantified as the coefficient of determination between the observed and predicted voxel responses to the 120 images in the validation set.

An image was identified from a set of candidate images as follows. 500 voxels were selected prior to identification without using the image. The selected voxels were those whose responses were predicted best. The image was identified as that in the set of candidate images to which the predicted voxel responses were most correlated with the observed voxel responses to the image (i.e. highest Pearson product-moment correlation coefficient between predicted and observed voxel responses). Decoding performance was quantified as the accuracy of identifying the 120 images in the validation set from the set of candidate images. The set of candidate images contained the 120 images in the validation set and the 9144 images in the Caltech 101 dataset (Fei-Fei et al., 2007).

3. Results

3.1 Feature model

To learn the feature transformation, we used a two-layer sparse coding (SC) model of 625 simple (i.e. first layer) and 625 complex (i.e. second layer) cells (Hyvärinen & Hoyer, 2001). Concretely, we first arranged the simple cells on a square grid graph that has circular boundary conditions. We then fixed the weights between the simple and complex cells such that each complex cell locally pooled 25 simple cell energies. Next, we estimated the weights between the input and the simple cells from 50000 patches of size 32 × 32 pixels by maximizing the sparseness of the locally pooled simple cell energies. Each simple cell was fully connected to the input. We randomly sampled the patches from the 1750 images of size 128 × 128 pixels in the estimation set. Finally, we defined the simple and complex cell responses of the SC model as a static compressive nonlinear function of the simple cell and locally pooled simple cell energies, respectively (i.e. a total of 625 simple and 625 complex cell responses per patch of size 32

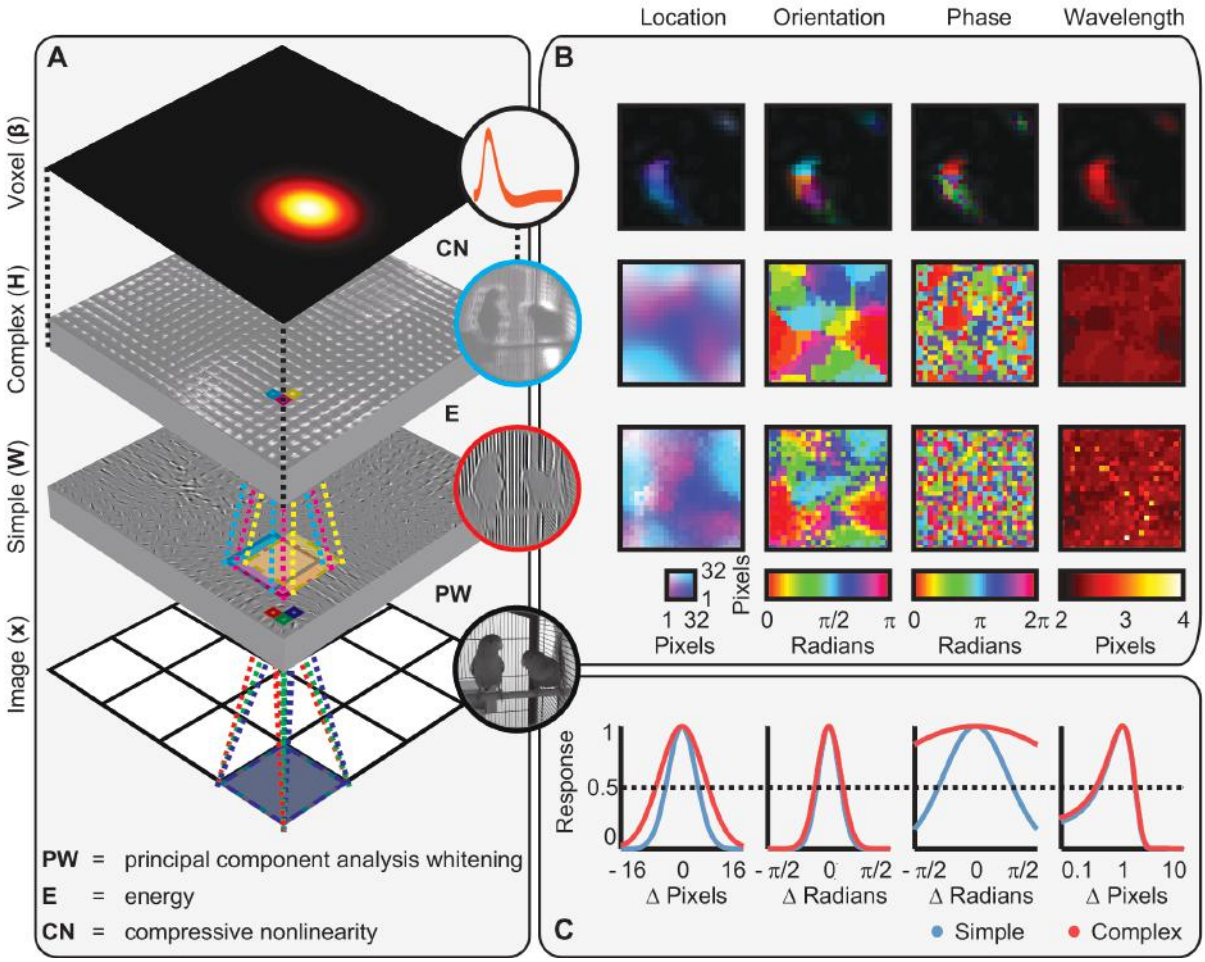


Fig. 1 Encoding model that comprises the feature and voxel models. **A.** Receptive fields. Simple cell receptive fields show the inverse of the weights between an image patch and the simple cells. Complex cell receptive fields show the locally pooled energies of the inverse of the weights between an image patch and the simple cells. Voxel receptive field shows the responses of a representative voxel to a point stimulus at different locations in the visual field. The encoding model predicts the response of the voxel to a 128 × 128 image x as follows: Each 16 non-overlapping 32 × 32 patch of the image $z^{(i)}$ is first vectorized and linearly transformed to 625 simple cell responses, i.e. $Wz^{(i)}$. Energies of each 625 overlapping 5 × 5 simple cell neighborhood are then locally pooled, i.e. $H(Wz^{(i)})^2$, and nonlinearly transformed to one complex cell response, i.e. $\log(1 + H(Wz^{(i)})^2)$. Next, 10000 complex cell responses are linearly transformed to one voxel response, i.e. $\boldsymbol{B}^T\phi(x)$ where $\phi(\mathbf{x}) = (\log(1 + \mathbf{H}(\mathbf{Wz}^{(1)})^2)^T, \ldots, \log(1 + H(Wz^{(16)})^2)^T)^T$. **B.** Maps of phase, location, orientation and wavelength. Maps for the simple and complex cells show the preferred phase, location, orientation and wavelength. Maps for the voxel show the complex cells that it pools, i.e. the maps for the complex cells that are weighted by **β**. C. Population tuning curves as a function of change in preferred phase, location, orientation and wavelength for the simple and complex cells. Responses are normalized to have a maximum of one.

× 32 pixels and 10000 simple and 10000 complex cell responses per image of size 128 × 128 pixels). The SC model learned topographically organized, spatially localized, oriented and bandpass simple and complex cell receptive fields that were similar to those found in the primary visual cortex (Fig. 1A) (De Valois, Albrecht, & Thorell, 1982; Hubel & Wiesel, 1968; Jones & Palmer, 1987; Parker & Hawken, 1988).

3.2 Topography

To analyze the topography of the simple and complex cells, we fitted Gabor wavelets to their receptive fields. As a result, we found maps of preferred phase, location, orientation and wavelength (Fig. 1B). Most adjacent simple and complex cells were selective to similar location, orientation and wavelength, whereas they were selective to different phase. These maps reproduced some of the salient features of the columnar organization of the primary visual cortex such as retinotopy and singularities. The topography suggests that the complex cells are more phase-invariant and less selective to location than the simple cells since they pool the energies of the simple cells that are selective to different phase. To analyze the invariance of the simple and complex cells, we fitted Gaussian functions to medians of their responses to Gabor wavelets. As a result, we

found population tuning curves as a function of change in preferred phase, location, orientation and wavelength (Fig. 1C). The difference between these population tuning curves confirm the invariance that is suggested by the topography. Like the simple cells, most complex cells were selective to orientation (standard deviation of 0.38 versus 0.40 radians) and wavelength (standard deviation of 0.52 versus 0.54 pixels). Unlike the simple cells, most complex cells were more phase-invariant (standard deviation of 0.82 versus 2.76 radians) and less selective to location (standard deviation of 3.70 versus 5.86 pixels). Therefore, they responded to Gabor wavelets in a specific orientation and wavelength within a relatively large receptive field, regardless of their exact phase and location.

3.3 Baseline

To establish a baseline, we used the GWP model (Daugman, 1985; Jones & Palmer, 1987; T. S. Lee, 1996) of 10921 phase-invariant complex cells (Kay et al., 2008). Variants of this model were used in a series of seminal encoding and decoding studies (Kay et al., 2008, 2013; Naselaris et al., 2009; Nishimoto et al., 2011). Concretely, the GWP model is a hand-designed population of quadrature-phase Gabor wavelets that span a range of locations, orientations and wavelengths. Each Gabor wavelet is fully connected to the input. Complex cell responses of the GWP model are defined as the square root of the pooled energies of the quadrature-phase Gabor wavelets that have the same location, orientation and wavelength (i.e. a total of 10921 responses per image of size 128 × 128 pixels).

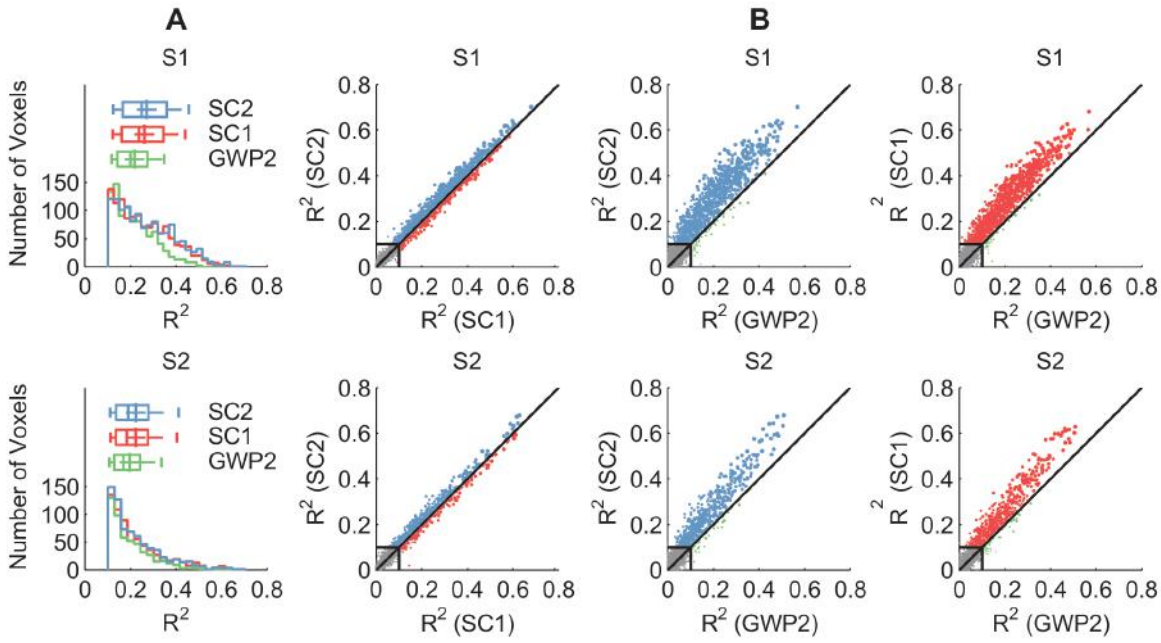
3.4 Voxel model

To learn the voxel transformation, we used regularized linear regression. We estimated the voxel models from the 1750 feature-transformed stimulus-response pairs in the estimation set by minimizing the L2-penalized least squares loss function. The combination of the voxel model with either the first layer of the SC model, the second layer of the SC model or the GWP model resulted in three encoding (i.e. SC1, SC2 or GWP2) models. The SC1 and SC2 models pooled the 10000 simple and 10000 complex cell responses of the SC model, respectively. The GWP2 model pooled the 10921 complex cell responses of the GWP model.

3.5 Encoding

We quantified the encoding performance of the SC1, SC2 and GWP models as the coefficient of determination (R2) between the observed and predicted voxel responses to the 120 images in the

Fig. 2 Encoding performance of the SC1, SC2 and GWP2 models. **A.** Prediction R^2 across the voxels that survived an R^2 threshold of 0.1. Box-and-whisker plots show the median (vertical line), mean (plus sign), 25% and 75% quantiles (box), and 9% and 91% quantiles (whiskers) of R^2. **B.** Prediction R^2 in each voxel that survived the threshold.



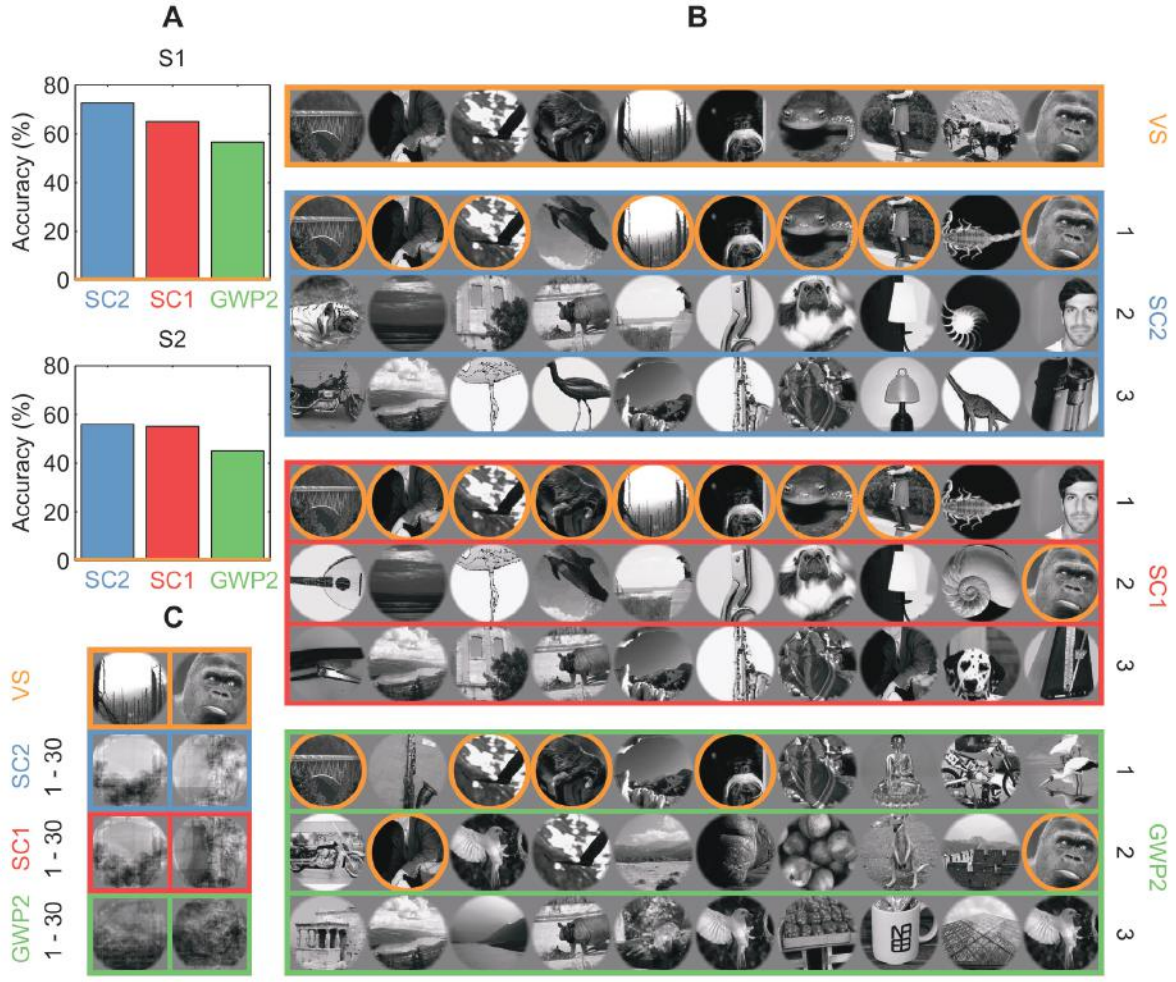


Fig. 3 Decoding performance of the SC1, SC2 and GWP2 models. **A.** Identification accuracy. Orange line shows the chance-level accuracy. **B.** Example identifications of the first 10 images in VS for the subject S1. For each of the 10 images, identifications show the three images in the set of candidate images whose observed voxel responses were first, second and third most correlated with their predicted voxel responses. **C.** Example reconstructions of the fifth and 10th images in B. For each of the two images, reconstructions show the weighted average of the 30 images in the set of candidate images except the image to be identified whose observed voxel responses were most correlated with their predicted voxel responses.

validation set. We found that the performance of the SC1 and SC2 models was significantly higher than that of the GWP2 model (binomial test, $p << 0.05$). Figure 2A compares the performance of the models across the voxels that survived an R2 threshold of 0.1. On average, the mean R2 of the SC1 and SC2 models was 0.24 across 18% of the voxels and 0.25 across 19% of the voxels, respectively. In contrast, the mean R2 of the GWP2 model was 0.21 across 14% of the voxels. Figure 2B compares the performance of the models in each voxel that survived the threshold. On average, the SC1 and SC2 models better predicted the responses of 19% of the voxels (5% and 14%, respectively), whereas those of 0.8% of the voxels that were better predicted by the GWP2 model. These results suggest that the statistically adapted low-level sparse representations of natural images better span the space of early visual cortical representations than the Gabor wavelets.

3.6 Decoding

We quantified the decoding performance of the SC1, SC2 and GWP models as the accuracy of identifying the 120 images in the validation set from a set of candidate images. The set of candidate images contained the 120 images in the validation set and the 9144 images in the Caltech 101 dataset (Fei-Fei, Fergus, & Perona, 2007). We found that the performance of the SC1 and SC2 models was significantly higher than that of the GWP2 model (binomial test, $p < 0.05$). Figure 3 compares the performance of the models. On average, the

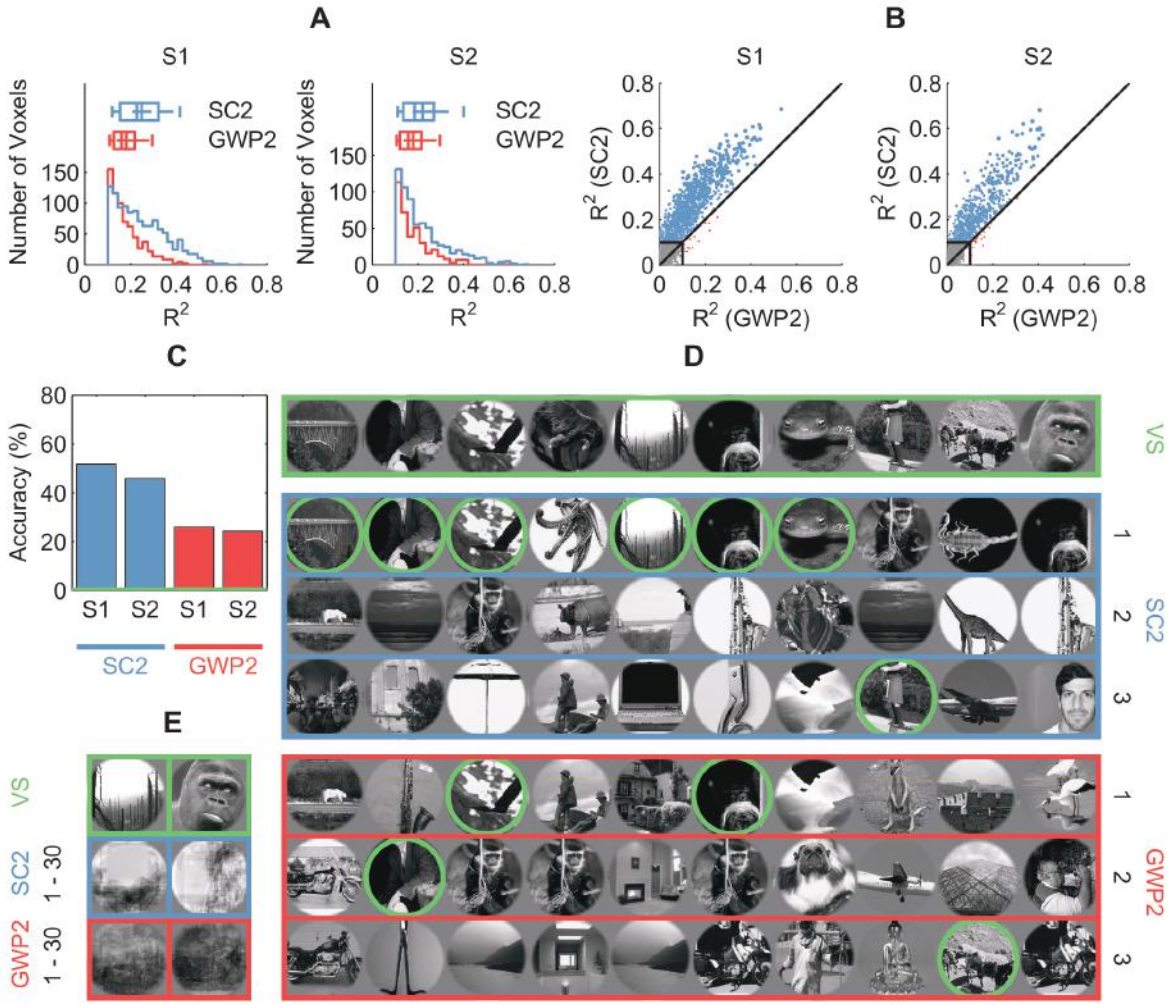


Fig. 4 Encoding and decoding performance of the SC2 and GWP2 models after translating the images in the validation set (VS) by five pixels (i.e. 0.8°) in a random dimension. Interpretation of the figure is similar to that of Figure 2 and Figure 3. **A.** Prediction R^2 of the models across the voxels that survived an R^2 threshold of 0.1. **B.** Prediction R^2 of the models in each voxel that survived the threshold. **C.** Identification accuracy. Green line shows the chance-level accuracy. **D.** Example identifications of the first 10 images in VS for the subject S1. **E.** Example reconstructions of the fifth and 10th images in D.

accuracy of the SC1 and SC2 models was 60% and 64%, respectively. In contrast, the accuracy of the GWP2 model was 51%. The chance-level accuracy was 0.01%. These results indicate that the statistically adapted low-level sparse representations of natural images can be more effectively exploited in identification than the Gabor wavelets.

3.7 Invariance

In principle, the SC2 and GWP2 models should be invariant to translations in images since the second layer of the SC model and the GWP model are phase-invariant complex cell models that have a degree of spatial invariance. Spatial invariance is of particular importance for decoding since a general visual decoder should be able to identify images, regardless of their exact location. To analyze the spatial invariance of the models, we evaluated their average encoding and decoding performance after translating the images in the validation set by five pixels (i.e. 0.8°) in a random dimension (Fig. 4). The difference in the mean R2 between the models increased to 0.06 across 7.57% of the voxels. The SC2 model better predicted the responses of 17.66% of the voxels in contrast to those of 0.63% of the voxels that were better predicted by the GWP2 model. The difference in the accuracy between the models increased to 24%. While these results demonstrate that the SC2 model is more invariant to translations in images than the GWP2 model, they raise the question of whether the performance difference between the models before the translations was a mere consequence of invariance to the translations that were inherent in the data.

Since the SC1 model performed significantly better than the GWP2 model like the SC2 model, we concluded that the performance difference between the encoding models before the translations was not a mere consequence of the invariance to the translations that were inherent in the data. Rather, the SC model is a better basis of the early visual cortical representations than the GWP model.

4. Discussion

In this study, we addressed the question of how to model feature spaces to better predict brain activity. We argued that the conventional approach constrains the formulation of hypotheses about what a voxel should represent. To overcome this constraint, we introduced a general framework for making directly testable predictions of how statistically adapted representations of ecologically valid visual stimuli are pooled by single voxels. To benchmark our framework against the conventional approach, we used a "shallow" feature model of statistically adapted low-level sparse and invariant representations of natural images. This feature model was an obvious starting point to explore the utility of our framework since we were able to compare it with a GWP model of phase-invariant complex cells – the fundamental building block of many state-of-the-art encoding and decoding models. As a result, we showed that these representations better span the space of early visual cortical representations and can be more effectively exploited in identification than the Gabor wavelets.

At the same time, what mid- or high-level representations (e.g. higher-order statistical features or semantic contents of natural images) should span the space of extrastriate visual cortical (e.g. V4 or inferior temporal gyrus) representations remain an open question. In contrast to the conventional approach, our framework is particularly suited to address this question since it lends itself to predictive and normative modeling of feature spaces. In particular, highly nonlinear multi-layer statistical generative models of natural images that learn hierarchical feature spaces (i.e. deep learning) can be used to predict unknown visual cortical representations, based on different coding principles.

One approach to estimate these models is to maximize the likelihood of all layers at the same time. However, this approach is not scalable and requires the computation of intractable partition functions that are impossible to integrate analytically and computationally expensive to integrate numerically. Nevertheless, methods such as score-matching (Hyvärinen, 2005) and noise-contrastive estimation (Gutmann & Hyvärinen, 2012) have been used to estimate unnormalized nonlinear multi-layer statistical generative models of natural images(Gutmann & Hyvärinen, 2013; Köster & Hyvärinen, 2010). An alternative approach is to use models such as deep belief networks that consist of multiple layers of restricted Boltzmann machines. Such deep models can be scaled by convolution (H. Lee, Grosse, Ranganath, & Ng, 2009) and estimated by maximizing the likelihood of one layer at a time, using the output of each layer as input for the subsequent layer (Hinton, Osindero, & Teh, 2006).

While these models are typically concerned with solving practical problems such as object recognition (Bengio, Courville, & Vincent, 2013), some of the feature spaces (e.g. mid-level edge junctions) that are learned by these models were shown to match well with the space of early visual cortical (e.g. V2) representations in quantitative comparisons (H. Lee, Ekanadham, & Ng, 2007). Models of even higher-level representations such as high-level object parts (H. Lee et al., 2009) or complete objects (Le et al., 2012) can be used to probe the unknown visual cortical representations. Furthermore, similar models were shown to learn feature spaces that are admitted by stimulus sets other than natural images, both within the visual modality such as natural movies (Le, Zou, Yeung, & Ng, 2011) and across other modalities such as auditory or somatosensory (Saxe, Bhand, Mudur, Suresh, & Ng, 2011). These models can be used to probe the unknown cortical representations in different modalities.

From a decoding perspective, statistical generative models of natural images can be used for generative decoding. For example, a deep belief network of multi-layer conditional restricted Boltzmann machines was shown to reconstruct handwritten digits from stimulus-evoked multiple voxel responses by sampling from the model after conditioning it on the voxel responses (van Gerven, de Lange, & Heskes, 2010). Statistical generative models are of particular importance for Bayesian reconstruction of natural images from stimulus-evoked voxel responses since they can be used as a prior distribution in Bayesian inference.

While providing a new approach to probe cortical representations, this study complements other developments in encoding and decoding. For example, encoding models that involve computations to account for contrast saturation or heterogeneous contrast energy were shown to improve prediction of single voxel responses to visual stimuli (Kay et

al., 2013). At the same time, these modeling work go hand in hand with the developments in fMRI such as the improvements in contrast-to-noise ratio and spatial resolution that are facilitated by the increases in magnetic field strength (Duyn, 2012). For example, the existence and spatial features of orientation-selective columns in humans were demonstrated by using high-field fMRI (Yacoub, Harel, & Uğurbil, 2008). These developments can provide insight into how cortical representations are learned, encoded and transformed.

5. Conclusion

In conclusion, we introduced a general framework that improves prediction of human brain activity in response to natural images. In particular, it enables theory-driven formulation of quantitative hypotheses about what a voxel should represent that can be confirmed or falsified using encoding models. Furthermore, it enables unsupervised learning of transformations of raw stimuli in any stimulus set to invariant representations that can be effectively exploited in a supervised learning task such as classification, identification or reconstruction using decoding models. Taken together, future prospects of our framework (e.g. deep learning) and neuroimaging techniques (e.g. ultra-high-field fMRI) herald the emergence of computational cognitive neuroscience as an established area of research to probe the unknown columnar functional organization of the human brain.

6. References

Barlow, H. W. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication* (pp. 217–234). Cambridge: MIT Press.

Bell, A. J., & Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research, 37*(23), 3327–3338.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798–828.

Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge: Cambridge University Press.

Brown, E. N., Kass, R. E., & Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience, 7*(5), 456–61.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America. A, *Optics and Image Science, 2*(7), 1160–9.

Dayan, P., & Abbott, L. F. (2005). *Theoretical Neuroscience: Computational And Mathematical Modeling of Neural Systems*. Cambridge: MIT Press.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22*(5), 545–59.

Duyn, J. H. (2012). The future of ultra-high field MRI and fMRI for study of the human brain. *NeuroImage, 62*(2), 1241–1248.

Edelman, A., Arias, T. A., & Smith, S. T. (1998). The Geometry of Algorithms with Orthogonality Constraints. *SIAM Journal on Matrix Analysis and Applications, 20*(2), 303–353.

Fei-Fei, L., Fergus, R., & Perona, P. (2007). Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. *Computer Vision and Image Understanding, 106*(1), 59–70.

Gutmann, M. U., & Hyvärinen, A. (2012). Noise-contrastive estimation of unnormalized statistical models, with applications to natural image statistics. *Journal of Machine Learning Research, 13*, 307–361.

Gutmann, M. U., & Hyvärinen, A. (2013). A three-layer model of natural image statistics. *Journal of Physiology - Paris, 107*(5), 369–98.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425–2430.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation, 18*(7), 1527–54.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195*(1), 215–43.

Hyvärinen, A. (2005). Estimation of Non-Normalized Statistical Models by Score Matching. *Journal of Machine Learning Research, 6*, 695–709.

Hyvärinen, A. (2010). Statistical Models of Natural Images and Cortical Visual Representation. *Topics in Cognitive Science, 2*(2), 251–264.

Hyvärinen, A., & Hoyer, P. O. (2001). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Research, 41*(18), 2413–2423.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology, 58*(6), 1233–58.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–85.

Kay, K. N., Naselaris, T., & Gallant, J. L. (2011). fMRI of human visual areas in response to natural images. CRCNS.org.

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature, 452*(7185), 352–355.

Kay, K. N., Winawer, J., Rokem, A., Mezer, A., & Wandell, B. A. (2013). A two-stage cascade model of BOLD responses in human visual cortex. *PLoS Computational Biology, 9*(5), e1003079.

Köster, U., & Hyvärinen, A. (2010). A two-layer model of natural stimuli estimated with score matching. *Neural Computation, 22*(9), 2308–33.

Le, Q., Ranzato, M., Monga, R., Devin, M., Chen, K., Corrado, G., Dean, J., Ng, A. (2012). Building high-level features using large scale unsupervised learning. In *International Conference on Machine Learning*.

Le, Q. V, Zou, W. Y., Yeung, S. Y., & Ng, A. Y. (2011). Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis. In *Conference on Computer Vision and Pattern Recognition*.

Lee, H., Ekanadham, C., & Ng, A. (2007). Sparse deep belief net model for visual area V2. In *Neural Information Processing Systems*.

Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In *International Conference on Machine Learning*.

Lee, T. S. (1996). Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(10), 959–971.

Mitchell, T. M., Shinkareva, S. V, Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science, 320*(5880), 1191–1195.

Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M., Morito, Y., Tanabe, H. C., Kamitani, Y. (2008). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron, 60*(5), 915–929.

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Cambridge: MIT Press.

Naselaris, T., Kay, K. N., Nishimoto, S., & Gallant, J. L. (2011). Encoding and decoding in fMRI. *NeuroImage, 56*(2), 400–410.

Naselaris, T., Prenger, R. J., Kay, K. N., Oliver, M., & Gallant, J. L. (2009). Bayesian reconstruction of natural images from human brain activity. *Neuron, 63*(6), 902–915.

Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology, 21*(19), 1641–6.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*(6583), 607–609.

Parker, A. J., & Hawken, M. J. (1988). Two-dimensional spatial structure of receptive fields in monkey striate cortex. *Journal of the Optical Society of America. A, Optics and Image Science, 5*(4), 598–605.

Pasley, B. N., David, S. V, Mesgarani, N., Flinker, A., Shamma, S. A., Crone, N. E., Chang, E. F. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology, 10*(1), e1001251.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102–1107.

Saxe, A. M., Bhand, M., Mudur, R., Suresh, B., & Ng, A. Y. (2011). Unsupervised learning models of primary cortical receptive fields and receptive field plasticity. In *Neural Information Processing Systems*.

Schoenmakers, S., Barth, M., Heskes, T., & van Gerven, M. (2013). Linear reconstruction of perceived images from human brain activity. *NeuroImage. 83*, 951–961.

Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J.-B., Lebihan, D., & Dehaene, S. (2006). Inverse retinotopy: inferring the visual content of images from brain activation patterns. *NeuroImage, 33*(4), 1104–1116.

Van Gerven, M. A. J., de Lange, F. P., & Heskes, T. (2010). Neural decoding with hierarchical generative models. *Neural Computation, 22*(12), 3127–42.

Vu, V. Q., Ravikumar, P., Naselaris, T., Kay, K. N., Gallant, J. L., & Yu, B. (2011). Encoding and decoding V1 fMRI responses to natural images with sparse nonparametric models. *The Annals of Applied Statistics, 5*(2B), 1159–1182.

Yacoub, E., Harel, N., & Uğurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America, 105*(30), 10607–12.

Social and Monetary Reward Processing in Autism Spectrum Disorders With and Without Attention-Deficit Hyperactivity Disorder

Jana Kruppa[1]
Supervisors: Anouk Scheres[2], Liesbeth Hoekstra[1,3,4], Jan Buitelaar[1,3,4] , Hilde Geurts[5], Nanda Lambregts- Rommelse[3,4]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behavior, The Netherlands
[2]Radboud University Nijmegen, Department of Developmental Psychology, Behavioral Science Institute (BSI), The Netherlands
[3]Karakter, Child and Adolescent Psychiatry University Centre, Nijmegen, The Netherlands
[4]Radboud University Medical Centre Nijmegen, Department of Psychiatry, Donders Institute for Brain, Cognition and Behavior, The Netherlands
[5]University of Amsterdam, Department of Psychology, Dutch Autism and ADHD research center (d'Arc), The Netherlands

According to the social motivation theory, social dysfunction in Autism Spectrum Disorders (ASDs) can be attributed to a deficit in reward processing and motivation specific to social stimuli. However, reward processing abnormalities may also extent beyond the social domain, may constitute a familial trait and may be explained by the presence of Attention-Deficit Hyperactivity Disorder (ADHD) comorbidity. These aspects were examined in the present study using functional magnetic resonance imaging assessing blood-oxygen level-dependent (BOLD) activation during anticipation of social and monetary reward in children (8-11 years) with ASD (+/-ADHD) and their unaffected siblings, relative to matched controls. All children showed a higher neural activation in regions belonging to the reward system (ventral striatum) during anticipation of high social and monetary rewards. No differences between groups were found. Behavioral results suggest that children with ADHD comorbidity are less motivated to perform well if no monetary rewards are expected but are extra motivated when receiving high social rewards. Taken together, these findings reveal that reward processing dysfunctions are present in other domains than only social, do not constitute a familial trait and are altered by the presence of ADHD comorbidity.

Keywords: ASD, ADHD, social reward, monetary reward, anticipation, endophenotype, fMRI, siblings

Corresponding author: Jana Kruppa; E-mail: janaakruppa@gmail.com

1. Introduction

Autism Spectrum Disorders (ASDs) are a group of neurodevelopmental disorders, characterized by impairments in reciprocal social interaction, communication and repetitive stereotypic behavior (American Psychiatric Association, 1994). Of these three areas of dysfunction, deficits in social interaction are considered as the most central to the disorder (Scott-Van Zeeland, Dapretto, Ghahremani, Poldrack, & Bookheimer, 2010). According to the social motivation theory of autism (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012), social dysfunction in ASDs can be attributed to a deficit in reward processing and motivation specific to social stimuli. An early deficit in social attention leads to less social learning experiences and therefore to an imbalance in attending to social and non-social stimuli. Eventually, this chain of events may result in the development of disrupted social skills (Chevallier et al., 2012). In non-clinical contexts, reward processing activates brain structures constituting the brain reward system, including the nucleus accumbens (NAcc), caudate, putamen, amygdala and ventromedial prefrontal cortex (vmPFC) (Knutson, Adams, Fong, & Hommer, 2001). Also, the anterior cingulate cortex (ACC), orbifrontal cortex (OFC) and ventral striatum (VS) were nominated as being key structures of the reward network (Scott-Van Zeeland et al., 2010). In ASDs, however, neurobiological responses to reward have been shown to be deviant (Dichter et al., 2010). Reward in response to social stimuli has been investigated in several studies supporting the social motivation theory about a deficit in reward processing of social stimuli in ASDs. In their functional magnetic resonance imaging (fMRI) study, Scott-Van Zeeland et al. (2010) have found a disproportionate diminished neural response in the ventral striatum to social reward in children with ASDs. Also, Delmonte et al. (2012) found support for the theory, by demonstrating reduced activity in the left dorsal striatum in male adolescents with ASDs as compared to a matched control group.

However, other literature suggests that reward deficiency might be broader than only the social domain. Dichter and colleagues (2010) found a reward-circuitry (specifically, nucleus accumbens) hypoactivation in response to monetary incentives in adult male participants with ASDs. A hypoactivation of amygdala and ACC was found even in response to both monetary and social reward in children with ASDs in the study of Kohls, Schulte-Ruether et al. (2012). A possible explanation for the findings may be the high comorbidity between ASD and Attention-Deficit Hyperactivity Disorder (ADHD). ADHD is characterized by an age-inappropriate hyperactivity, inattentiveness, and impulsivity (APA, 2000) and its comorbidity rates with ASDs range from approximately one-third (Simonoff et al., 2008) to three-quarters (Lee & Ousley, 2006), most likely caused by shared etiological factors underlying both disorders (Rommelse, Franke, Geurts, Hartman, & Buitelaar, 2010). Studies investigating reward processing in ADHD generally show reduced neural activation in the reward circuits during monetary reward anticipation (Scheres, Milham, Knutson, & Castellanos, 2007; Stroehle et al., 2008; for a review see: Cubillo, Halari, Smith, Taylor, & Rubia, 2012; see for a recent meta-analysis Plichta & Scheres, 2013). In sum, the key question whether reward dysfunction in ASDs is specific to the social domain, as stated by social motivation theory, or reflects a more general reward deficit possibly explained by the high comorbidity with ADHD, is still unsolved.

Another uninvestigated issue in clarifying the role of a social reward deficit in ASD is the presence of similar deficits in siblings of ASD probands. Siblings share on average 50% of their genetic make-up with the proband and are therefore vulnerable to displaying similar cognitive characteristics as their affected sibling. Particularly unaffected siblings are of interest in this context, because it has been proposed that this group is of use to investigate the concept 'endophenotype'. Endophenotypes are defined as heritable features with an intermediate position between genotype and phenotype (Gottesman & Gould, 2003). Endophenotypes are associated with a condition, are present in affected individuals regardless of whether the condition is manifested and co-segregate with the condition in families (Gottesman & Gould, 2003). Importantly, endophenotypes are found in unaffected relatives at a higher rate than in the general population (Gottesman & Gould, 2003; Saresella et al., 2009) and are therefore proposed to be more sensitive in picking up genetic risks than phenotypic measures (Gottesman & Gould, 2003). The concept of endophenotypes, which has originated in the field of biology, has also become popular in neuropsychiatric research in the recent years. Surprisingly, siblings of children with ASD have not often been subject of investigation in neuroimaging research, although recent reports suggest that this may be a fruitful area of research. For example, Spencer et al. (2011) showed that the neural response to facial expression of emotion in unaffected siblings did not differ

significantly from the response in the group with ASD. Hence, in order to shed more light on the role of a social reward dysfunction in ASDs it is expedient to also investigate unaffected siblings, given the high heritability for ASD.

In the present study, we want to investigate whether 1) children with ASD show a specific or general reward processing dysfunction, as reflected by impairments in a social and/or monetary reward paradigm, 2) this depends on ADHD comorbidity and 3) whether this pattern is familial. In order to answer these questions, we invited children with ASDs (+/- ADHD), their unaffected siblings and typically developing children between 8-11 years to complete a social and monetary incentive delay (MID/SID) task (original task: see Knutson, Westdorp, Kaiser, & Hommer, 2000) in the MR scanner. For this task, participants are instructed to press a button when a target appears on screen, which is preceded by a cue indicating the amount of reward that can be won when executing the button press on time. The MID/SID task has proven to be sensitive in assessing brain responses to monetary and social incentives during reward anticipation and outcome in both clinical and non-clinical populations via functional magnetic resonance imaging (Delmonte et al., 2012; Rademacher et al., 2010). As the social motivation theory proposes a motivational ('wanting') deficit in ASDs as opposed to the hedonic sensation of 'liking' (for a review of the 'wanting'/'liking' framework, see: Kohls, Chevallier, Troiani, & Schultz, 2012), we will focus on the anticipation phase of reward processing in this work.

2. Methods

2.1 Subjects and screening

Forty children between 8-11 years old with ASDs (IQ>70), 18 of their unaffected siblings and 19 controls were selected (based on their age) from an ASD family genetics project [Biological Origins of Autism (BOA)]. The BOA project aims at examining the genetic, biochemical and cognitive origins of ASD as well as studying the overlap between ASD and ADHD on these levels. The BOA project included ASD families with at least one child between 2 and 20 years old with an ASD diagnosis, at least one biological sibling (regardless of ASD-status) and at least one biological parent. In addition, control families participated with no family history of ASD or ADHD. Healthy controls were matched as closely as possible with the patient and unaffected sibling group on age, IQ, and gender. Exclusion criteria for all children included a diagnosis of organic brain disorder (epilepsy, traumatic brain injury), sensory impairments, a diagnosis of a defined non-genetic or genetic cause of ASD (Rett's syndrome, fragile-X syndrome), known genetic disorders (Down syndrome), and MRI contraindications.

Diagnoses of ASD and ADHD were based on a history of clinical diagnosis confirmed by proband assessment by questionnaires/interview: for ASD screening, the Social Communication Questionnaire (SCQ, Rutter, M., Bailey, A. & Lord, C., 2003) and the algorithm of the Autism Diagnostic Interview-Revised (ADI-R, Le Couteur, Lord, & Rutter, 2003) were used. The SCQ was completed by parents. For all children scoring above the SCQ cutoff (>10) score, the ADI-R was administered by a certified clinician to reconfirm the diagnosis of ASD. For ADHD screening, the Conners long version Rating Scales-Revised (CRS-R Parent and Teacher Rating Scale: Conners, 1997) and the parental account for childhood symptoms (PACS, Taylor, Sandberg, Thorley, & Giles, 1991) were administered. The CRS-R was filled in by parents and teachers, and for all children scoring above cutoff on any of the ADHD subscales (T>63), the PACS was administered by a certified clinician to make a formal ADHD diagnosis. For a more detailed description of the assessment procedure, see van Steijn et al. (2012).

For the final analyses, data of a number of participants were excluded. From behavioral analyses, data of 11 participants were excluded: from the affected group, 5 children ended the session prematurely and 3 children were excluded due to 0% performance on at least 2 out of 3 reward levels in at least 1 out of 3 runs on at least 1 of the 2 tasks. For the latter reason, also 1 unaffected sibling and 1 control participant were excluded. Another participant from the control condition ended the session prematurely. In addition to the same sample that was excluded from behavioral analysis, 17 more (in total 28) participants were excluded due to motion above 3.5 mm on at least one of the 2 tasks: 10 children from the affected group, 3 unaffected siblings and 2 children from the control group.

All children were asked to withdraw from their medication at least 24 hours before the test session, if possible. Of the children included for behavioral analyses, 5 children from the affected group were taking psychotropic mediations (one child took Prozac®, one Risperdal®, one Risperdal® and melatonin, one Risperdal® and methylphenidate/

Table 1. Characteristics of the final participant sample included for fMRI and behavioral analyses. (TIQ, Total Intelligence Quotient; Autism Quotient, measure of the extent of autistic traits; ASD, autism spectrum disorder; ADHD, attention-deficit/hyperactivity disorder; ADI, autism diagnostic interview – revised; PACS, parental account for childhood symptoms)

	Imaging sample			**Behavioral sample**		
	Probands N = 19	**Siblings N = 15**	**Controls N = 15**	**Probands N = 36**	**Siblings N = 18**	**Controls N 18**
	M(SD)	**M(SD)**	**M(SD)**	**M(SD)**	**M(SD)**	**M(SD)**
% Male	89,5	80,0	80,0	88,9	83,3	83,3
Age in years	10,4 (1.2)	10.3 (1.1)	10.0 (1.1)	10.5 (1.2)	10.3 (1.2)	10.0 (2.4)
% Medications	5.3	-	-	13.8	-	-
TIQ	97.9 (28.6)	109.4 (20.9)	110.2 (11.1)	106.5 (25.2)	109.4 (21.3)	113.1 (13.7)
Social Communication Questionnaire	17.9(5.9)	2.9 (2.3)	3.8 (2.6)	18.7 (6.7)	3.3 (2.6)	3.7 (2.4)
Autism Quotient	92.1 (17.6)	35.3 (13.1)	21.0 (22.0)	90.6 (19.7)	36.3 (14.6)	20.6 (22.5)
Conners inattentive scores	63.0 (9.4)	50.2 (11.2)	45.2 (5.6)	63.6 (8.2)	51.6 (10.8)	45.2 (5.2)
% Above cut-off T>63	47.4	6.7	-	50.0	5.6	-
Conners hyperactive scores	68.4 (10.6)	51.5 (9.2)	46.9 (5.0)	68.3 (9.2)	51.2 (8.7)	46.7 (4.6)
% Above cut off T>63	73.7	6.7	-	69.4	5.6	-
% ASD (ADI-R)	100	-	-	100	-	-
% ADHD (PACS)	26.3	6.7	-	33.3	5.6	-

Concerta®, and the other child methylphenidate/ Concerta®) on the test day. From the participants included for imaging analyses, only one child of the affected group was taking medication (Risperdal®) on the test day. Controls and unaffected siblings were not taking any psychotropic medications at the time of scanning. All participants had normal or corrected-to-normal visual acuity. Participant characteristics of the sample included for behavioral and fMRI analyses are displayed in Table 1.

2.2 Procedure

Families potentially satisfying inclusion criteria (i.e. confirmed diagnosis of probands and proper age), based on known data from the BOA project, were phoned and asked whether they would be interested in participating in a MRI-study of the BOA-project. Only families that had given consent for being contacted for possible follow up studies of the BOA project were contacted. On the test day, children first completed a mock scan session prior to imaging to accustom them to the scanning environment and to practice the fMRI (MID/SID) task. Participating families were paid 10€ for each child that underwent the procedure, in addition to 15€ travel expenses. These expenses were provided independent of withdrawal during the experiment. Children who completed the MID task earned some pocket money (15-20€) depending on their task performance. All procedures were conducted with the understanding and written consent of participants and their legal guardians as well as formal approval of the 'Commissie Mensgebonden Onderzoek' (CMO), Region Arnhem-Nijmegen.

2.3 fMRI task

The fMRI task comprised of two different tasks: the 'monetary incentive delay' (MID) task as introduced by Knutson et al. (2000) and the 'social incentive delay' (SID) task, an adaptation of the former (Spreckelmeyer et al., 2009). In both tasks the children practiced 27 trials during the mock scan session, which were followed by 3 experimental runs of 27 trials each (in total 81 trials) for the MID and the SID task in the MR scanner. Each trial consisted of: a cue (250 ms), followed by a fixation cross (randomly varying duration of 2000-2500 ms), a target (duration varying between 166-333 ms), a blank screen (4093 ms – delay duration – target duration) and a feedback screen (1650 ms) (see Fig. 1 for an illustration of the gain context in the SID and MID task). The cue consisted of an empty circle (no reward), a circle with one horizontal line (low reward) or a circle with three horizontal lines (high

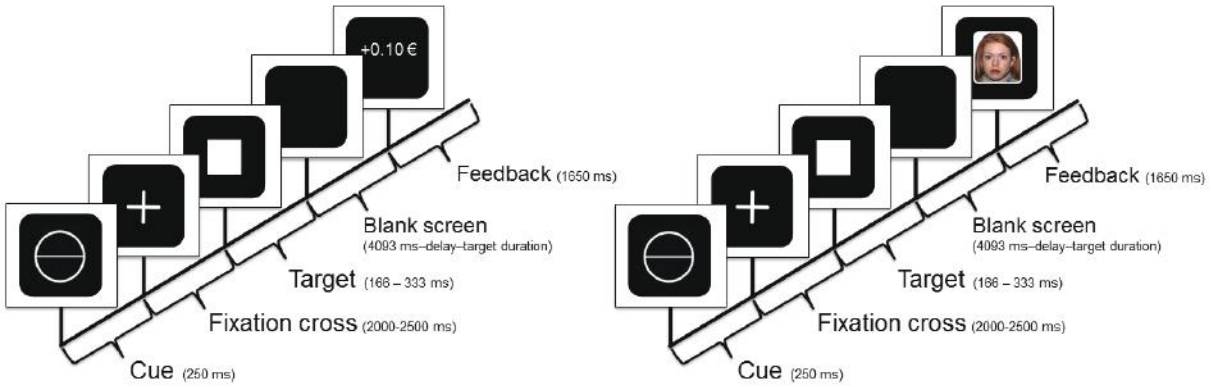


Fig. 1 Example of trial sequences of the MID (left) and SID (right) for the low reward condition. Participants were instructed to press a button as soon as the target, a filled white square, would appear on the screen. If a response was made as long as the target was on screen, a reward was won (here, during low reward trials, 0.10 € on the MID task and a happy face with a moderate smile on the SID task). The circle with one horizontal line (cue) at the beginning of each trial indicated the amount of reward that could be won if the button is pressed on time.

reward). Purpose of the cue was to generate reward anticipation and to signal the potential outcome, varying on three levels. Each run contained 3 reward conditions with 9 trials each, resulting in a total of 27 trials per condition. The three levels of monetary reward were 0.00€ (n = 9 per run, cue = empty circle), 0.10€ (n = 9 per run, cue = circle with one horizontal line) and 1.00€ (n = 9 per run, cue = circle with three horizontal lines). The three levels of social reward were a neutral face (n = 9 per run, cue = empty circle), a happy face with a moderate smile with a closed mouth (n = 9, cue = circle with one horizontal line) and a happy face with a more intense smile with an open mouth (n = 9, cue = circle with three horizontal lines). The order of appearance of the 3 conditions was counterbalanced and did not differ across blocks. Participants were instructed to press a button as fast as possible when the target, a solid white square, would appear on the screen. Importantly, a button press was required following all cue types, even though nothing could be won after the cue signaling no reward. If the button was pressed on time, that is, before the target had disappeared from the screen, participants could either win money (MID task) or a happy face (SID task). On the MID task, success was acknowledged by presenting a feedback screen with the respective amount of money (see Fig. 1). In the case of no outcome, due to a failure of pressing the button in time or a cue signaling no reward, a screen presenting a zero amount of money was shown. On the SID task, success was acknowledged by presenting a screen with a smiling face (depending on the cue either a moderate or intense smile). For no outcome, a neutral face was presented. At the end of each run, a screen with the amount of money (MID) or the amount of compliments (SID) the participant had won in total appeared. Task difficulty was manipulated by adjusting the target duration to individual reaction times, in order to achieve a hit rate of ~66% for all participants. Reaction times (RT) achieved during the practice session were used to determine target durations for the task performed in the scanner. Prior to the start of each run, target durations for all trials of that run were adjusted based on the mean RT of correct hits of the previous run. To make the task realistic and to motivate the subjects, the money gained during the MID task (max. 20€) was actually paid, on top of the normal participation gratification of 10€. Face stimuli of the SID task were color photographs taken from a standard database of facial expressions (NimStim set of Facial Expressions; available at: http://www.macbrain.org (Tottenham et al., 2009)). Trial types were pseudo-randomly ordered. Participants completed 3 times the MID and 3 times the SID task. The order of completing first the MID or the SID task was counterbalanced. Task duration in total was approximately 30 minutes.

2.4 fMRI setting

For both tasks, stimuli presentations and recording of reaction times were performed using the software E-prime (Psychology Software Tools, Inc., Pittsburgh). Children indicated their response by pressing a button of a response box with the index finger of their right hand.

2.5 Image acquisition

Scanning was performed on a 3-Tesla head-dedicated MRI system (MagnetomTrioTim; Siemens). Participants lay in a supine position, while

head movement was restricted using foam cushions and a tape attached to the head of participants. The functional scans for the MID/SID task were acquired using a multi-echo gradient pulse sequence (TR = 2490 ms; TE = 9.8 ms, 22.2 ms, 35 ms, 47 ms, 60 ms; flip angle = 80°). Each volume consisted of 31 transversal slices with a thickness of 3 mm. The voxel resolution was 3.5 x 3.5 x 3.0 mm. Structural images were acquired using a T1-weighted 3D MPRAGE sequence (TR = 2300 ms; TE = 3.03 ms; 192 sagittal slices; slice thickness = 1 mm; voxel size = 1.0 x 1.0 x 1.0 mm).

2.6 Behavioral data analysis

Analyses were performed with Statistical Package for the Social Science version 19 (SPSS 19; IBM Corporation, Armonk, NY, USA). Repeated measure ANOVAs were used separately for speed (mean reaction time based on correct hits) and accuracy (percentage correct hits), with group (3 levels: proband, unaffected sibling, control) as between-subjects factor and reward magnitude (3 levels: no, low, high) as within-subjects factor. Hit rates were analyzed in order to see whether the intended manipulation to homogeneize hit rates between participants was successful. Reward magnitude was averaged across participant's performance on 3 different runs. Analyses were run for the MID and SID task apart. Post-hoc comparisons between probands with and without ADHD were performed. Following Cohen's guidelines (Cohen, 1988), effect sizes were defined in terms of the percentage of explained variance: 1, 9 and 25% were used to define small, medium, and large effects. This translates into $\eta p2$-values of 0.01, 0.06 and 0.14.

2.7 Image analysis

Functional data were preprocessed and analyzed using Statistical Parametric Mapping (SPM) version 8 (Wellcome Department of Cognitive Neurology, London, UK), implemented in MATLAB 7.9 (Mathworks Inc., Sherborn, MA, USA). Time series for fMRI were realigned spatially to the first volume, and the signal in each slice was realigned temporally to the first slice using a sinc interpolation. Mean intra-participant head motion was below 3.5 mm translation and 3°rotation. Resliced volumes were normalized to a standard EPI template based on the Montreal Neurological Institute (MNI) reference brain in Talairach space. Normalized images were spatially smoothed with an 8-mm FWHM isotropoic Gaussian Kernel and high-pass filtered to 1/128 Hz.

For the fMRI analyses, first a general linear model for each participant was specified in which the preprocessed fMRI data were coupled to the vectors of anticipation onset of each condition (no reward, low reward, high reward) in all sessions of the SID and MID task. Also, six realignment parameters were included for each run of the two tasks to account for translation and rotation variability. Subsequently, six t-contrasts were computed: 1) MID no reward – implicit baseline, 2) MID low reward – implicit baseline, 3) MID high reward- implicit baseline, 4) SID no reward – implicit baseline, 5) SID low reward – implicit baseline, 6) SID high reward – implicit baseline.

For the second level analysis, first a region of interest (ROI) analysis was performed using MarsBar 0.42 (MARSeille Boîte À Région d'Intérêt; Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. , 2002). Striatal subregions were subdivided by defining 3 seed regions: ventral caudate (inferior)/nucleus accumbens (±9, 9, -8), ventral caudate (superior) (±10, 15, 0) and dorsal caudate (±13, 15, 9). The radius of the spheres that were used to build the ROIs was 8 mm. Regions were defined anatomically in the left and right hemisphere separately with coordinates adopted from Di Martino et al. (2008). These areas were defined as ROIs because they are known for their involvement in reward processing, especially the more ventral regions. The ventral striatum has been reported repeatedly as being activated during anticipation of monetary and social reward (Dichter et al., 2010; Rademacher et al., 2010; Spreckelmeyer et al., 2009). Beta-values were extracted from the fMRI data for each reward level. For the ROIs, the extracted beta-values of each participant were exported to SPSS, and subsequently analyzed for each task separately in a repeated measures ANOVA with group (affected, sibling, control) as between-subjects factor and reward level (no reward, low reward, high reward) as within-subjects factor. Only statistically significant effects ($p < .05$) and marginal effects of interest ($.05 \leq p < .10$) are reported.

In order to see which possible other regions than ventral striatum are activated during anticipation of rewards in each task, whole brain analyses were performed using a full factorial design with group as between-subject factor and reward level as within subject factor. Main effects of group and reward as well as interactions between group and reward were tested for the MID and SID separately. The main effect of reward was further explored by looking at different contrasts: 'high reward minus no reward', 'high reward minus low reward', 'high & low

reward (combined) minus no reward'. Significant voxels and clusters are reported as significant if p < .05 corrected with the family-wise error (FWER) approach. MRIcron was used to label the significant clusters and voxels.

3. Results

3.1 Behavioral results

3.1.1 SID task

For percentage correct hits, no significant main effect of group was observed ($F(2) = 0.80$, $p = .46$, partial $\eta 2 = 0.02$). However, a main effect of magnitude of reward was found ($F(2, 63) = 4.36$, $p = .02$, $\eta p2 = 0.12$). In the post-hoc tests, this effect was evident when comparing no reward vs. high reward trials ($F(1) = 8.72$, $p = .004$, $\eta p2 = 0.12$). Children obtained a significantly higher percentage correct hits on high reward trials (M = 37.30) than on no reward trials (M = 31.92) (Fig. 2A). No significant group by reward interaction was found ($F(4, 126) = 0.53$, $p = .71$, $\eta p2 = 0.02$) on percentage correct hits. The post-hoc test for examining potential effects within the affected group revealed no significant main effect of group ($F(1) = 0.49$, $p = .49$, $\eta p2 = 0.02$) on percentage correct hits. However, a significant group by reward interaction ($F(2, 29) = 3.84$, $p = .03$, $\eta p2 = 0.20$) was found. Differences between groups were present in the high reward condition ($F(1) = 11.46$, $p = .002$, $\eta p2 = 0.28$), in which children with ASD+ADHD (M = 44.79) performed better as compared to children with ASD only (M = 34.39), whereas there were no differences between groups in the low ($F(1) = 2.79$, $p = .11$, $\eta p 2= 0.08$)

Table 2. Main performance variables of the SID task. (%hits, percentage correct hits; MRT, mean reaction time)

Measures	Affected (n=32)	ASD only (n=21)	ASD + ADHD (n=11)	Siblings (n=18)	Control (n=19)
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
% hits					
No reward	30.74 (12.29)	32.73 (10.90)	26.91 (14.36)	35.29 (11.09)	29.74 (12.40)
Low reward	37.33 (15.95)	35.80 (15.05)	40.24 (17.99)	35.73 (13.98)	32.06 (10.67)
High reward	37.97 (16.15)	34.39 (16.02)	44.79 (14.73)	39.00 (9.45)	34.95 (13.00)
MRT					
No reward	165.01 (82.04)	178.47 (84.03)	139.59 (75.12)	193.53 (65.28)	191.73 (72.41)
Low reward	174.84 (73.16)	175.83 (77.21)	172.94 (68.28)	172.13 (57.36)	179.75 (67.05)
High reward	169.01 (70.56)	162.99 (74.19)	180.73 (64.80)	180.13 (64.59)	184.90 (70.49)

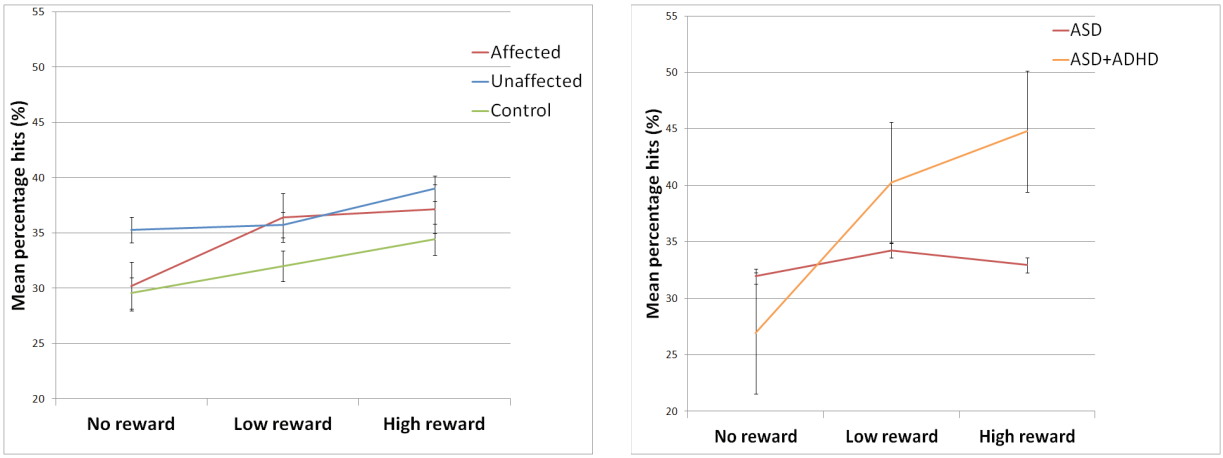


Fig. 2A. Mean percentage hits (SID): main effect of reward. Children across all groups obtained a higher percentage correct hits on high reward trials than on no reward trials. **B.** Mean percentage hits (SID): post-hoc group comparison. Differences between groups were present in the high reward condition, in which children with ASD+ADHD performed better as compared to children with ASD only, whereas there were no differences between groups in the low and high reward conditions.
Note: error bars refer to standard errors.

and high reward (F(1) = 1.10, p = .30, ηp2 = 0.04) conditions (Fig. 2B). For mean reaction time, no significant differences found between groups (F(2) = 0.41, p = .67, ηp2 = 0.01) and reward levels (F(2, 63) = 0.58, p = 5.62, ηp2 = 0.02). Main performance variables of the SID task are summarized in Table 2.

3.1.2 MID task

For percentage correct hits, no main effect of group was obtained (F(2) = 0.74, p = .48, ηp2 = 0.02). However, a main effect of reward was found (F(2, 66) = 3.76, p = .03, ηp2 = 0.10), with a significantly higher percentage correct hits on high reward trials (M = 53.63) than on low reward (M = 51.12) trials (F(1) = 5.54, p = .02, ηp2 = 0.08) and no reward trials (M = 50.09) (F(1) = 6.23, p = .02, ηp2 = 0.09), respectively (Fig. 3A). The percentage correct hits did not differ on no reward versus low reward trials (F(1) = 0.64, p = .43, ηp2 = 0.01). A marginally significant interaction between reward level and group was found (F(4, 132) = 2.02, p = .10, ηp2 = 0.06): the affected group had a lower mean percentage correct hits (M = 46.46) than control children (M = 54.03) on no reward trials only (F(1) = 8.08, p = .01, ηp2 = 0.14) (Fig. 3A). The post-hoc test for examining potential effects within the affected group revealed no significant main effect of group (F(1) = 0.94, p = .34, ηp2 = 0.03) on percentage correct hits. A marginally significant interaction was found between group and reward (F(2, 32) = 2.74, p = .08, ηp2 = 0.15). Differences between groups were present in the no reward condition (F(1) = 13.61, p = .001, ηp2 = 0.29), in which children with ASD+ADHD (M = 41.03) performed worse as compared to children with ASD only (M = 49.66), whereas there were no differences between groups in the low and high reward conditions (Fig. 3B). No significant differences in mean reaction time were found between groups (F(1) = 0.29, p = .51, ηp2 = 0.01) and reward levels (F(2, 32) = 0.05, p = .95, ηp2 = .003). Main performance variables of the MID task are summarized in Table 3.

3.2 fMRI results

3.2.1 SID task

ROI analysis

The ROI analyses were focused on 3 regions of the ventral striatum: the ventral caudate (inferior)/ nucleus accumbens, ventral caudate (superior), and dorsal caudate. For all of these regions, both left and right, no significant difference between groups in activation during the anticipation of social incentives was found (all p-values > .55). However, a significant main effect of reward magnitude was found in the right ventral caudate (inferior) /nucleus accumbens (F(2, 45) = 3.21, p = .05, ηp2 = 0.13) (Fig. 4). A higher neural activation was found during the anticipation of high rewards (M = 0.34) as compared to low rewards (M = 0.07) (F(1) = 6.50, p = .01, ηp2 = 0.12) in all groups. No interactions between group and reward level were found (all p-values > .51).

Whole brain analysis

No main effects of group and reward level were found. However, a main effect of reward was found for the contrast 'no reward<high reward'. All

Table 3. Main performance variables of the MID. (%hits, percentage correct hits; MRT, mean reaction time)

Measures	Affected (n=35)	ASD only (n=22)	ASD + ADHD (n=13)	Siblings (n=18)	Control (n=19)
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
% hits					
No reward	46.46 (12.43)	49.66 (9.33)	41.03 (15.31)	49.79 (11.93)	54.03 (6.42)
Low reward	51.75 (8.16)	51.85 (7.14)	51.59 (9.97)	50.62 (5.82)	50.98 (5.93)
High reward	53.12 (9.46)	52.86 (8.79)	53.56 (10.86)	53.50 (9.61)	54.26 (9.33)
MRT					
No reward	222.29 (56.08)	217.84 (54.36)	229.83 (60.34)	225.80 (37.68)	250.38 (55.07)
Low reward	221.51 (56.76)	218.94 (59.34)	225.85 (54.14)	227.05 (30.79)	252.35 (56.38)
High reward	223.23 (52.23)	219.65 (57.04)	229.29 (44.40)	217.88 (42.39)	245.68 (57.79)

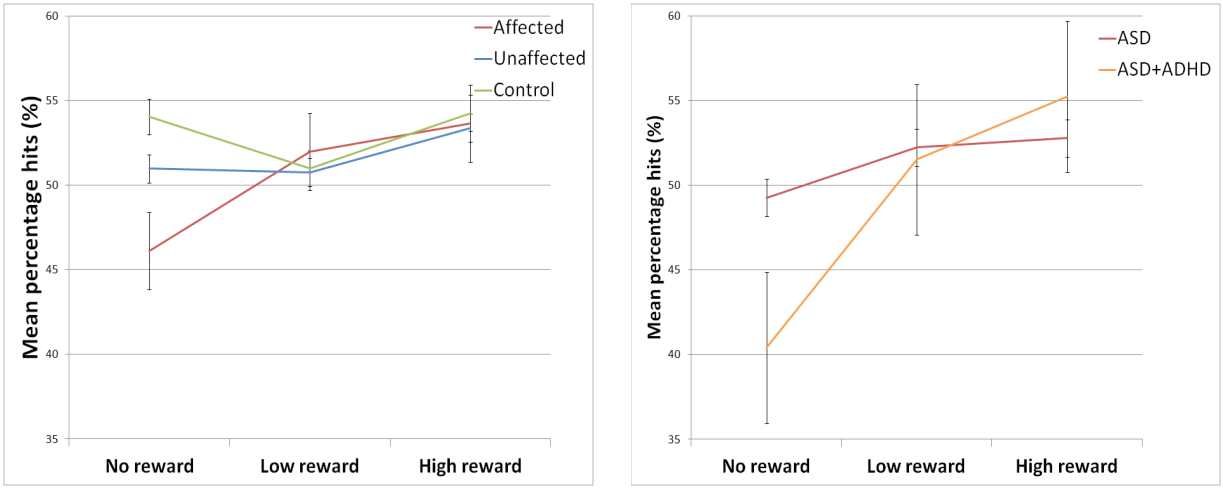


Fig. 3 Mean percentage hits (MID) **A.** Main effect of reward & marginal significant group x reward interaction. All children obtained a higher percentage correct hits on high reward trials than on low reward trials and no reward trials, respectively. Children in the affected group obtained a lower percentage correct hits than children in the control group on no reward trials. **B.** Post-hoc group comparison. A marginal significant interaction between reward and diagnosis is shown. Differences between groups were present in the no reward condition, in which children with ASD + ADHD performed worse as compared to children with ASD only, whereas there were no differences between groups in the low and high reward conditions. Note: error bars refer to standard errors.

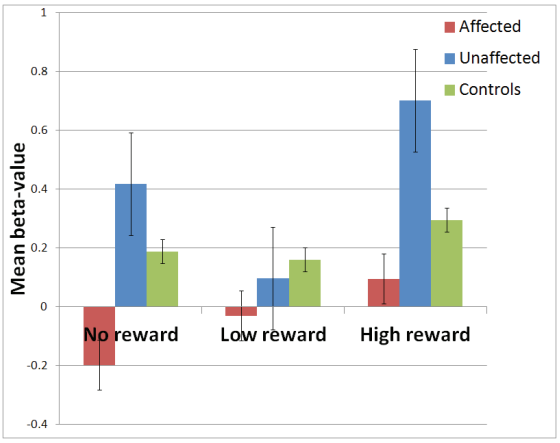


Fig. 4 Bar graph of the right ventral caudate (inferior) showing the beta values for anticipation of no reward, low reward, and high reward on the SID task

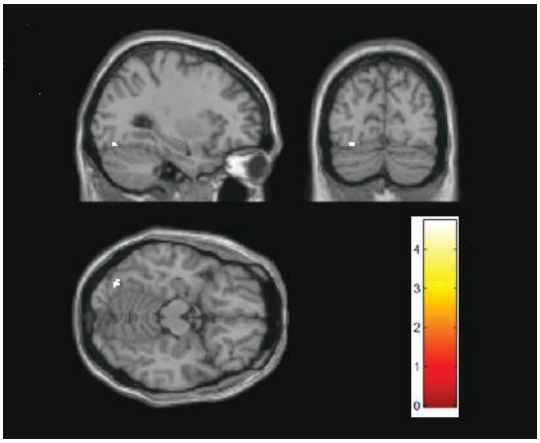


Fig. 5 Activation map for the left fusiform gyrus in the contrast 'high reward minus no reward' in the sagittal, coronal and horizontal plane.

groups showed a higher neural activation in the left fusiform gyrus during anticipation of high social incentives as compared to no rewards (Fig. 5). No interaction effects between groups and reward levels were found. See Table 4 for a summary of results.

3.2.2 MID

ROI analyses

For the three ROIs, ventral caudate (inferior) / nucleus accumbens, ventral caudate (superior), and dorsal caudate, no main effect of group was found (all p-values > .30). However, a main effect of reward level was found. In the left and right ventral caudate (inferior) /nucleus accumbens, neural activity was higher during anticipation of high reward (left: M = 0.58; right: M = 0.54) than low reward (left: M = 0.18; right: M = 0.10) and no reward (left: M = 0.10; right: M = -0.06), respectively (all p-values < .05). In the left ventral caudate (superior), the main effect of reward was only marginally significant ($F(2, 45) = 2.88$, $p = .07$, $\eta p2 = 0.11$). Higher activity was present during the anticipation of high reward (M = 0.60) than no reward (M = 0.20) ($F(1) = 5.83$, $p = .02$, $\eta p2 = 0.11$), and activation during anticipation of high rewards was marginally significantly higher than low reward (M = 0.32) ($F(1) = 3.33$, $p = .08$, $\eta p2 = 0.07$). In the right ventral caudate (superior), activation was higher during anticipation of high rewards (M = 0.47) than no rewards (M = -0.10) ($F(1) = 10.20$, $p = .003$, $\eta p2 = 0.18$) and low rewards (M = 0.09) ($F(1) = 5.38$, $p = .03$, $\eta p2 = 0.11$). Activation during anticipation of low reward

Table 4. Cluster showing significant reward level (no<high) differences to social incentives.

Area	L/R	Cluster size (mm^3)	Z	MNI coordinates		
				x	y	z
Fusiform gyrus	L	17	4.61	-32	-76	-16

Table 5. Cluster showing significant reward level differences to monetary incentives.

Area	L/R	Cluster size (mm^3)	Z	MNI coordinates		
				x	y	z
No reward < high reward Nucleus accumbens	R	22	4.83	10	8	-2
No reward < low reward Postcentral gyrus	R	3	4.54	56	-4	20

(M = 0.09) was marginally significantly higher than no reward (M = -0.10) ($F(1) = 3.48$, $p = .07$, $\eta p2 = 0.07$). In the right dorsal caudate, activation during anticipation of high reward (M = 0.51) was higher than no reward (M = 0.13). The difference between anticipation of high (M = 0.51) versus low (M = 0.33) ($F(1) = 3.87$, $p = .06$, $\eta p2 = 0.08$) and low reward (M = 0.33) versus no reward (M = 0.13) ($F(1) = 3.13$, $p = .08$, $\eta p2 = 0.06$) was only marginally significant. No main effect of reward was present in the left dorsal caudate ($F(2, 45) = 1.62$, $p = .21$, $\eta p2 = 0.07$). No interactions between group and reward level were found (all p-values > .20).

Whole brain analysis

No main effect of group was found. In all groups, no differences in anticipation across reward levels could be obtained. However, a main effect of reward was found when comparing anticipation of no reward versus high reward and no reward versus low reward: a higher activity was present in the right NAcc during anticipation of high rewards as compared to no rewards (Fig. 6A). Part of the postcentral gyrus was activated during anticipation of low rewards versus no rewards (Fig. 6B). No significant interaction between group and reward level was found. Results of the activated brain area during anticipation of 'high reward >no reward' and 'low reward>no reward' are summarized in Table 5.

4. Discussion

In the present study we administered a social and monetary incentive delay (SID/MID) task to children with ASD with and without ADHD comorbidity, their unaffected siblings and healthy controls to explore 1) whether reward processing abnormalities in ASDs extend beyond the social domain, 2) can be explained by the presence of ADHD comorbidity and 3) constitute a familial trait. On the behavioral level, we found that differences in reward processing in ASDs were present during non-rewarded trials concerning monetary but not social incentives. More specifically, children from the affected group were less sensitive than children from the control group to non-rewarded trials of the MID. The presence of ADHD comorbidity was associated with a better performance in a high reward context in the social task and a relatively poor performance in a no reward context in the monetary task, when compared to children with ASD diagnosis only. No familiar patterns were found; all children

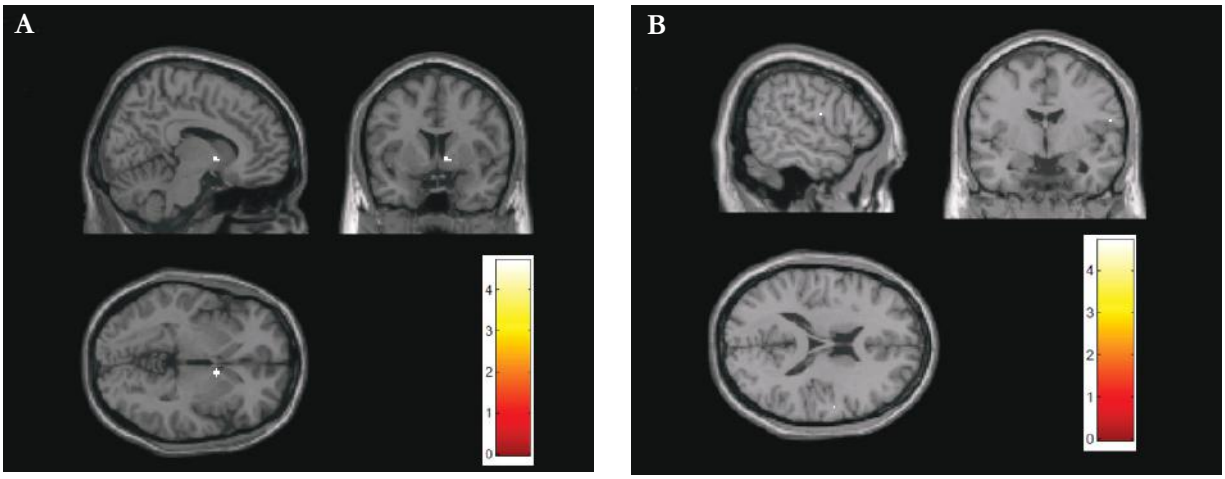


Fig. 6 Activation maps in the sagittal, coronal and horizontal plane. **A.** Activation of the right NAcc in the contrast 'high reward minus no reward'. **B.** and for the right postcentral gyrus in the contrast 'high reward minus low reward'.

performed better on rewarded than non-rewarded trials in both tasks. On the neural level, no reward processing abnormalities in ASDs and no patterns of familial traits could be obtained. However, on both tasks, all children were sensitive to the amount of reward, as reflected by a higher neural activation during anticipation of higher rewards (SID: fusiform gyrus, right NAcc; MID: postcentral gyrus, ventral striatum) than no rewards. Due to a limited number of children with a combined diagnosis in the affected group, possible effects of ADHD comorbidity were not examined on the neural level.

4.1 Behavioral responses to monetary and social reward

Our behavioral finding concerning better performance with increasing reward magnitude across groups is in line with the study of Delmonte et al. (2012). However, as described above (see Methods), mean accuracy was intended to be held constant across conditions and participants by adjusting the duration of the target from trial to trial to ensure that BOLD-activity would be specific to reward anticipation and not contaminated by responses to the target. As the task was supposed to serve as a behavioral baseline, the difference in accuracy is a rather surprising finding and suggests that the task manipulation did not alter task performance as intended. However, the differences in accuracy on the different reward levels confirm that the reward manipulation itself worked. Also in the study of Delmonte et al. (2012) differences in accuracy on different reward levels were found despite efforts to keep performance constant, demonstrating that the effect of reward was strong.

Our finding of less sensitivity in the affected group as compared to the control group on no reward trials on the MID is likely to be driven by children with ADHD comorbidity, as this group performed worse in the no reward condition when compared to children with ASD only. This finding is in line with the finding of Demurie, E., Roeyers, H., Wiersema, J. R. and Sonuga-Barke, E. (2013): in a monetary incentive delay go/no-go task, children with ASD+ADHD made more omission errors compared with children with ASD and controls. As in the present study, this group difference disappeared in the rewarded trials. Demurie et al. (2013) concluded that attention problems, which omission errors during response execution are linked to (Bezdijian, Baker, Lozano, & Raine, 2009), in children with combined diagnosis are worse compared with children with ASD when they are not rewarded. These attention problems appear to normalize in rewarded contexts. Also Slusarek, Velling, Bunk, and Eggers (2001) showed that children with ADHD performed well on an inhibition task when properly motivated. Compared to controls, children with ADHD performed worse under a low reward condition but performed as well as controls when given high levels of reward. Similarly, on the SID we found differences in performance between the two affected groups. Here, children with ASD+ADHD performed better than children with ASD in the high reward condition, suggesting an advantage of having an ADHD diagnosis in addition to ASD in social contexts. The postulated hyposensitivity to social rewards in ASD (Chevallier et al., 2012) and hypersensitivity to social rewards (Kohls, Herpertz-Dahlmann, & Konrad, 2009) as well as a stronger tendency to seek reward (Scheres, Oosterlaan, & Sergeant, 2001) in children with ADHD may explain the advantageous combination of both diagnoses in this specific context. More research on the effects of social reward in ADHD is needed, as most current studies investigating reward processing in ADHD administer monetary incentives.

Taken together, our behavioral results suggest that 1) differences in reward processing in ASDs are present in domains other than only social, as shown by a less sensitivity to non-rewarded trials on the MID as compared to controls. Children with ASD seem to be less motivated than typically developing children when no monetary reward is expected. 2) This pattern is aggravated by ADHD comorbidity, as shown by the direct comparison with children with ASD. In a high rewarded social context, having ADHD as a comorbidity seems to be advantageous for children with ASD. 3) No familial patterns in responses to social and monetary rewards were obtained.

4.2 Neural responses to social rewards

For the SID, a higher neural activation in the right ventral caudate (inferior)/nucleus accumbens was found during anticipation of high rewards as compared to low rewards in all groups. Also, in the left fusiform gyrus activity was higher during anticipation of high rewards as compared to no rewards. In contrast to the results of Scott-Van Zeeland and colleagues (2010), we did not find malfunctions in the reward circuitry for social rewards in ASD. Scott-Van Zeeland et al. (2010) used a learning paradigm and combined facial rewards

with verbal praise as social reinforcement. The addition of verbal praise might have been perceived as a stronger reinforcement and the learning context a higher motivation to perform well, making it more likely that differences in activation in the NAcc between children with ASD and controls were obtained. In addition, Scott-Van Zeeland et al. (2010) focused on reward outcome rather than anticipation, two processes which were associated with dissociable neural networks (Rademacher et al., 2010). Delmonte et al. (2012) looked at reward anticipation and outcome in ASD and controls, and did find reduced activity in dorsal striatum for social rewards in ASD as compared to controls only during consumption of reward and not during anticipation. Delmonte's finding is in line with results of Dichter et al. (2012), who also did not obtain reduced ventral striatal activity during anticipation, but during consumption of social rewards in children ASDs. Hence, findings suggest that reward processing abnormalities in ASDs are not present during anticipation of social rewards but rather associated with consumption of reward. However, another explanation for the fact that no differences between the ASD and control group could be obtained, might be the overall low number of correct hits on the SID in the present study. The NAcc is innervated by mesostriatal dopamine (Plichta & Scheres, 2013), which signals highest between cue and reward when reward outcomes occur under conditions of highest uncertainty (i.e. probability = 0.5) (Fiorillo, Tobler, & Schultz, 2003). Therefore, a probability of wins of approximately 0.3 – 0.4 on the SID task, as reflected by the respective percentage correct hits, might have led to lower VS-responses in the present study. Many studies (e.g. Dichter et al., 2012) using manipulations of target durations in order to keep accuracy constant like in the present study do not report actual accuracy rates and hence it is unknown whether task manipulations had the intended effect. Future studies on VS-responsiveness during reward anticipation using the SID task should report accuracy in addition to mean reaction time to clarify this important aspect.

The fusiform gyrus (FFA), the second region responding to higher social rewards in the present study, is a region critical to social perception (Neuhaus, Beauchaine, & Bernier, 2010). Among healthy controls, the FFA is associated with face detection and is shown to be less activated in people with ASD during emotional expression of faces (Pelphrey, Morris, McCarthy, & LaBar, 2007). In several studies, the FFA was reported to be active to monetary (Scott-Van Zeeland et al., 2010; Kohls, Schulte-Ruether et al., 2012) as well as social (Kohls, Schulte-Ruether et al., 2012) rewards across groups. However, previous studies finding activation in FFA do not pay much attention to this region, although recent reports suggest that visual brain regions are part of a widespread network signaling forthcoming reward. In their magnetoencephalography (MEG) study, Apitz and Bunzeck (2012) demonstrated in a face/scene discrimination task with monetary rewards versus no rewards that neural effects of the difference between faces and scenes were observed in event-related magnetic fields, amplified in the context of reward with a source lateralized to the lateral occipital cortex and fusiform gyrus. Hence, the authors concluded that reward motivation has a modulatory effect on neural computations of complex visual information. Furthermore, in a recent meta-analysis about reward-related processing in major depressive disorder, the role of fusiform gyrus in response to positive stimuli in MDD is highlighted (Zhang, Chang, Guo, Zhang, & Wang, 2013). Future studies on reward processing in ASDs could be strengthened by paying careful attention to other components of reward processing in addition to the common regions associated with the reward circuit. In conclusion, present imaging results of the SID task indicate that 1) no reward processing abnormalities in ASDs are present during the anticipation of social rewards in the context of a low probability in achieving rewards, 2) the pattern of social reward processing is not familial, as no differences between groups were found and 3) other brain regions like FFA frequently activated along with the reward circuitry merit further attention.

4.3 Neural responses to monetary rewards

For the MID, in all seed regions but the left dorsal caudate, as well as postcentral gyrus, a higher neural activation was found during anticipation of higher rewards. The finding of stronger effects in the ventral as compared to dorsal part of the striatum together with even an absence of reward effects in the left dorsal caudate might be explained by a ventral-to-dorsal gradient within the striatum found by Tanaka et al. (2004). Tanaka et al. (2004) found that immediate rewards were mainly processed within ventral regions whereas future rewards were processed in the dorsal regions of the striatum during a Markov decision task administered to healthy subjects. Immediate rewards translate to the monetary rewards gained in the present study, which were acknowledged immediately after correct

responses and disbursed by cash after completion of the task. Hence, present results are in line with Tanaka and colleague's suggestion of differential involvement of the cortico-basal ganglia loops in reward prediction at different time scales.

However, unlike other studies (e.g. Dichter et al., 2012), we did not find deviant reward circuitry (specifically, NAcc) function in the affected group. One possible explanation concerns the amount of reward provided in the high reward condition. Original studies of Knutson and colleagues delivered $5 as highest reward compared to 1€ in the present study. Also in the study of Scheres et al. (2007), who administered the task to participants with ADHD and healthy controls, notably the $5 condition caused a distinct increase in VS activation. Accordingly, the proportionally lower reward in the high reward condition of the present study might not have been rewarding enough in order to account for possible group differences. A second aspect, like for the SID task, is the percentage correct hits achieved on the task. Dichter et al. (2012) reported an adaptation of the task such that accuracy was maintained to a hit rate of approximately 66%. Similar to the SID task of the present study, attempts to keep mean accuracy constant at 66% failed for the MID task. Lowest hit rates were about 45% of the trials, which is lower than chance level and lower than in the study of Dichter et al. (2012), assumed that their manipulation worked. In the study of Kohls, Schulte-Ruether et al. (2012), participants were rewarded for successful task performance with a probability of even 80%. Possibly, the large difference in the percentage of correct hits between the present study and the study of Dichter et al. (2012) and Kohls, Schulte-Ruether et al. (2012) may account for the finding of no differences between groups. Thirdly, the heterogeneity of our affected group in combination with a small sample size should be taken into account. The present study included children with ASD as well as children with ASD and ADHD comorbidity, with and without psychotropic medication, resulting in a total of only 19 participants in the affected group. The study sample of Kohls, Schulte-Ruether et al. (2012), in contrast, consisted of comorbid-free and mediaction-naive participants with ASD. The disadvantage of including comorbidity- and medication-free participants only, however, is that the population of ASD patients is not accurately represented as a variety of comorbid emotional or behavioral disorders in ASDs are reported (e.g. Simonoff et al., 2008) and many of children with ASD are being treated with psychotropic medication (Frazier et al., 2011). Future studies should include a number of participants high enough in order to compare possible differential effects due to status of medication.

In sum, the imaging results of the MID task indicate that 1) no reward processing abnormalities in ASDs are present during anticipation of monetary rewards in a context of below chance level probability to achieve a (in comparison with the original task settings low) reward, 2) the pattern of reward processing of monetary incentives is not familial, as no group differences were present and 3) more studies with larger and more homogeneous groups are needed.

4.4 Study limitations

The present study has some limitations that should be considered. All of our participants with ASD were high-functioning, so the inference drawn to the autism population in general is limited, given the high heterogeneity within the autistic spectrum. Hence, it is possible that reward processing dysfunctions manifest differently in individuals across the spectrum. To be able to generalize results to the broader spectrum of ASD, replications with larger samples including also participants with lower functioning autism are needed. As in the current study only a very young age group was tested, results are also limited to children and hence can't be generalized to adolescents or adults with ASD. Furthermore, our proband sample consisted of medicated and unmedicated children, making the groups more inhomogeneous. Psychotropic medication, as used by some of our participants, have repeatedly been shown to induce an up- or down-regulation of the mesocorticolimbic reward system activity and thereby possibly biasing results (see for example Rubia et al., 2009). However, as stated above, including participants who use medication makes the study sample more representative of the ASD population. Our affected group included in addition to children with ASD only also children with ADHD comorbidity. In order to shed more light on the effects of ADHD comorbidity, future studies should include an additional group of participants with ADHD diagnosis only to compare groups.

In addition to the sample itself, also the administered task showed some limitations. The social stimuli used in the SID task (i.e. happy faces with increasing intensity of happiness) are relatively artificial compared to real-life social encounters. To circumvent possible difficulties for children

with ASD in face processing (Grelotti, Gauthier, & Schultz, 2002), Demurie, E., Roeyers, H., Baeyens, D., and Sonuga-Barke, E. (2012) operationalised social reward as pictograms of an interaction between two persons, with one person approving the performance of the other person with thumbs up and a compliment in a text balloon. The disadvantage of this task, however, is that in addition to points were earned in the social condition. Furthermore, it is difficult to totally separate the effects of social and monetary reward, as money might be associated with social interactions as well (for example, when paying at the cash desk). Another important difference between the monetary and social task in the present study is the fact that the money which children gained during the task was actually paid, but they didn't receive any 'actual' social reward at the end of the social task. Follow-up studies should keep on improving tasks in order to create more ecologically valid designs.

5. Conclusion

To conclude, our findings demonstrate that reward processing abnormalities in children with ASD were not present during the anticipation phase of reward and instead might be rather associated with the consumption of reward. No familial patterns of reward processing were obtained. The data further indicate that reward processing in ASD is altered by the presence of ADHD comorbidity: when no monetary reward is expected, children with ASD+ADHD seem to be less motivated to perform well but are extra motivated when receiving high social rewards. Further research on reward processing in ASDs with and without ADHD as comorbidity is needed.

6. References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th edn (DSM-IV)*. Washington, DC: APA.

Apitz, T., & Bunzeck, N. (2012). Reward modulates the neural dynamics of early visual category processing. *Neuroimage, 63*(3), 1614-1622.

Bezdijian, S., Baker, L. A., Lozano, D. I., & Raine, A. (2009). Assessing inattention and impulsivity in children during the go/no-go task. *British Journal of Developmental Psychology, 27*(2), 365-383.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using an SPM toolbox [abstract]. Presented at the Eights International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. *Neuroimage, 16*.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Science, 16*(4), 231-239.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Conners, C. K. (1997). *Conners' rating scales revised user's manual*. North Tonawanda, New York: Multi-Health Systems.

Cubillo, A., Halari, R., Smith, A., Taylor, E., & Rubia, K. (2012). A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex, 48*(2), 194-215.

Grelotti, D. J., Gauthier, I., & Schultz, R. T. (2002). Social interest and the development of cortical face specialization: what autism teaches us about face processing. *Developmental Psychobiology, 40*(3), 213-225.

Delmonte, S., Balsters, J. H., McGrath, J., Fitzgerald, J., Brennan, S., Fagan, A. J., & Gallagher, L. (2012). Social and monetary reward processing in autism spectrum disorders. *Molecular Autism, 3*(7), 1-13.

Demurie, E., Roeyers, H., Baeyens, D., & Sonuga-Barke, E. (2012). Temporal discounting of monetary rewards in children and adolescents with ADHD and autism spectrum disorders. *Developmental Science, 15*(6), 791-800.

Demurie, E., Roeyers, H., Wiersema, J. R., Sonuga-Barke, E. (2013). No Evidence for Inhibitory Deficits or Altered Reward Processing in ADHD: Data From a New Integrated Monetary Incentive Delay Go/No-Go Task. J*ournal of Attention Disorders*, 1–15.

Dichter, G. S., Felder, J. N., Green, S. R., Rittenberg, A. M., Sasson, N. J., & Bodfish, J. W. (2010). Reward circuitry function in autism spectrum disorders. *Social cognitive and affective neuroscience, 7*(2), 160-172.

Dichter, G. S., Richey, A., Rittenberg, A. M., Sabatino, A., & Bodfish, J. W. (2012). Reward circuitry function in autism during face anticipation and outcomes. *Journal of autism and developmental disorders, 42*(2), 147-160.

Di Martino, A., Scheres, A., Marqulies, D. S., Kelly, A. M., Uddin, L. Q., …, & Milham, M. P. (2008). Functional connectivity of human striatum: a resting state fMRI study. *Cerebral Cortex, 18*(12), 2735-2747.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Behavioral Brain Research, 299*, 165-170.

Frazier, T. W., Shattuck, P. T., Narendorf, S. C., Cooper, B. P., Wagner, M., Spitznagel, E. L. (2011). Prevalence and correlates of psychotropic medication use in adolescents with an autism spectrum disorder with and without caregiver-reported attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology, 21*(6), 571-579.

Gottesman, I. I. & Gould, T. D. (2003). The

endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry, 160*(4), 636-645.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience, 21*(16),1–5.

Knutson, B., Westdorp, A., Kaiser, E., & Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage, 12*, 20-27.

Kohls, G., Chevallier, C., Troiani, V., & Schultz, T. (2012). Social ‚wanting' dysfunction in autism: neurobiological underpinnings and treatment implications. *Journal of Neurodevelopmental Disorders, 4*(10), 1-20.

Kohls, G., Herpertz-Dahlmann, B., & Konrad, K. (2009). Hyperresponsiveness to social rewards in children and adolescents with attention-deficit/hyperactivity disorder. *Behavioral and Brain Functions, 5*, 5-20.

Kohls, G., Schulte-Ruether, M., Nehrkorn, B., Mueller, K., Fink, G., Kamp-Becker, I., Herpertz-Dahlmann, B., Schultz, R.T., & Konrad, K. (2012). Reward system dysfunction in autism spectrum disorders. *Social Cognitive and Affective Neuroscience, 8*(5), 565-572.

Le Couteur, A., Lord, C. & Rutter, M. (2003). *The Autism Diagnostic Interview-Revised (ADI-R)*. Los Angeles, CA: Western Psychological Services.

Lee, D. O. & Ousley, O. Y. (2006). Attention-deficit hyperactivity disorder symptoms in a clinic sample of children and adolescents with pervasive developmental disorders. *Journal of Child & Adolescent Psychopharmacology, 16*(6), 737-746.

Neuhaus, E., Beauchaine, T.P., Bernier, R. (2010). Neurobiological correlates of social functioning in autism. *Clinical psychology review, 30*, 733-748.

Pelphrey, K. A., Morris, J. P., McCarthy, G., LaBar, K. S. (2007). Perception of dynamic changes in facial affect and identity in autism. *Social Cognitive Affective Neuroscience, 2*(2), 140-149.

Plichta, M. M. & Scheres, A. (2013). Ventral striatal responsiveness during reward anticipation in ADHD and its relation to trait impulsivity in the healthy population: A meta-analytic review of the fMRI literature. *Neuroscience & Biobehavioral Reviews*, (in press).

Rademacher, L. Krach, S., Kohls, G., Irmak, A., Gruender, G., & Spreckelmeyer, K. N. (2010). Dissociation of neural networks for anticipation and consumption of monetary and social rewards. *NeuroImage, 49*(4), 3276-3285.

Rommelse, N. N. J., Franke, B., Geurts, H. M., Hartman, C. A., & Buitelaar, J.K. (2010). Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *European child & adolescent psychiatry, 19*(3), 281-295.

Rubia, K., Halari, R., Cubillo, A., Mohammad, A. M., Brammer, M., & Taylor, E. (2009). Methylphenidate normalises activation and functional connectivity deficits in attention and motivation networks in medication-naive children with ADHD during a rewarded continuous performance task. *Neuropharmacology, 57*(7), 640-652.

Rutter, M., Bailey, A. & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.

Saresella, M., Marventano, I., Guerini, F. R., Mancuso, R., Ceresa, L., Zanzottera, M., …, & Clerici,M. (2009). An autistic endophenotype results in complex immune dysfunction in healthy siblings of autistic children. *Biological Psychiatry, 66*(10), 978-984.

Scheres, A., Milham, M. P., Knutson, B., & Castellanos, F. C. (2007). Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit/hyperactivvity disorder. *Biological Psychiatry, 61*(5), 720-724.

Scheres, A., Oosterlaan, J., & Sergeant, J.A. (2001). Response inhibition in children with DSM-IV subtypes of AD/HD and related disruptive disorders: the role of reward. *Child Neuropsychology, 7*(3), 172-189.

Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(8), 921-929.

Scott-Van Zeeland, A. A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., Bookheimer, S. Y. (2010). Reward processing in autism. *Autism Research, 3*(2), 53-67.

Slusarek, M., Velling, S., Bunk, D., & Eggers, C. (2001). Motivational effects on inhibitory control in children with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(3), 355-363.

Spencer, M. D., Holt, R. J., Chura, L. R., Suckling, J., Bullmore, E. T., & Baron-Cohen, S. (2011). A novel functional brain imaging endophenotype of autism: the neural response to facial expression of emotion. *Translational Psychiatry, 1*, 2-7.

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad. K., Kircher, T., & Gruender, G. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social cognitive and affective neuroscience, 4*(2), 158-165.

Steijn, van, D., Richards, J. S., Oerlemans, A. M., Ruiter, de, S. W., Aken, van, M. A. G., Franke, B., Buitelaar, J. K., & Rommelse, N. N. J. (2012). The co-occurrence of autism spectrum disorders and attention-deficit/hyperactivity disorder symptoms in parents of children with ASD or ASD with ADHD. *Journal of Child Psychology and Psychiatry, 53*(9), 954-963.

Stroehle, A., Stoy, M., Wrase, J., Schwarzer, S., Schlagenhauf, F., Huss, M., Hein, J., Nedderhut, A., Neumann, B., Gregor, A., Knutson, B., Lehmkuhl, U., Bauer, M., & Heinz, A. (2008). Reward anticipation and outcomes in adult males with attention-deficit/hyperactivity disorder. *NeuroImage, 39*(3), 966-972.

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature neuroscience, 7*(8), 887-893.

Taylor, E., Sandberg, S., Thorley, G., & Giles, S. (1991). *The Epidemiology of Childhood Hyperactivity*. Maudsley Monograph No. 33. Oxford: Oxford University Press.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., …, & Nelson, C. (2009). The NimStim set of facial expressions: judgements from untrained research participants. *Psychiatry Research, 168*(3), 242-249.

Zhang, W. N., Chang, S. H., Guo, L. Y., Zhang, K. L., Wang, J. (2013). The neural correlates of reward-related processing in major depressive disorder: a meta analysis of functional magnetic resonance imaging studies. *Journal of affective disorders, 151*(2), 531-539.

Representations During Visual Mental Imagery: Effects of Low- and Higher-Level Features

Brónagh McCoy[1]
Supervisors: Anke Marit Albers[1], Ivan Toni[1], Floris de Lange[1]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour The Netherlands

In contrast to perception, visual imagery of a stimulus is induced within the brain itself. Much research based on conventional univariate analysis of the visual cortex has proven unsuccessful in defining consistent activation during imagery and the extent to which it is perception-like in nature. Multivariate pattern analysis (MVPA) techniques have recently provided a means to assess the similarities and differences between mental imagery, working memory and perception, by highlighting representational involvement based on the patterns of responsive voxels within a region of interest. In the current functional magnetic resonance imaging (fMRI) study, we are interested in the imagery-related functions of areas making up the visual hierarchy. We look at how imagery representations in the visual cortex differ, by directly comparing in one experiment imagery of both a low- and a higher-level feature of letter stimuli. With this approach we examine whether imagery representations are dependent on the low- or higher-level demands of a task, in the same way that perceptual representations are. Using the same stimuli, we also examine whether there is a difference in information content between familiar and unfamiliar letters. We find that familiarity with a stimulus plays a role in the decoding of representations in the visual cortex. Our imagery results agree with the reverse hierarchy theory of perceptual learning, suggesting that representations of low-level features are no longer required when a higher-level representation has already been established. Our results also indicate that it is difficult to maintain global representations of unfamiliar stimuli, thus leading to better representations of the finer detail.

Keywords: functional magnetic resonance imaging, multivariate pattern analysis, visual cortex, imagery, working memory, familiarity

Corresponding author: Brónagh McCoy; E-mail: mccoyb4@gmail.com

1. Introduction

Mental visual imagery is a process of seeing in the "mind's eye". Visual imagery is induced within the brain itself, whereas perception is motivated directly by external stimuli through the sensory system. Although these are different processes, a leading theory, the "depictive" account (Kosslyn, Ganis, & Thompson, 2003) assumes that imagery involves the same mechanisms as those used in the early phases of perception and hence predicts activation in the same early visual processing regions, i.e. the striate (primary visual area, V1) and extra-striate (higher visual areas, V2-V5) cortex. Much research has focused on the extent to which these perceptual areas are activated during imagery (Podgorny & Shepard, 1978; Kaas, Weigelt, Roebroeck, Kohler, & Muckli, 2010; Klein, Paradis, Poline, Kosslyn, & Le Bihan, 2000; Knauff, Kassubek, Mulack, & Greenlee , 2000) but has shown mixed results, such as either an increase or a decrease in V1 activity. An important criterion placed on the depictive account is that if imagery does indeed activate the primary visual cortex, then the topographical organization of V1 should reflect the spatial properties of the image (Pylyshyn, 2002). Slotnick, Thompson and Kosslyn (2005) presented a study in support of the depictive account; visual mental imagery was shown to evoke topographically organized activity in striate and extrastriate cortex, leading to retinotopic maps of imagery similar to those of perception. Another study examining reaction times to real and imagined stimuli at different retinal eccentricities also found that visual perception and imagery share a similar visuotopic organization (Marzi, Mancini, Metitieri, & Savazzi, 2006).These analyses support the depictive account and activation of the visual cortex during visual imagery. The experiment we propose targets the specific conditions under which there may be representational involvement of early visual areas in imagery.

Representational involvement or informational content is based on the patterns of responses within a region of interest (ROI), and may be found using multivariate pattern analysis techniques (MVPA). Instead of focusing on activity in only one voxel at a time, as is the case with conventional univariate analysis, MVPA approaches provide information based on the combined activity of distributed voxels (Haynes & Rees, 2006). Visual selectivity for low-level details of a stimulus, such as the orientation of a grating, has been found to be represented by activity patterns in the visual cortex, even when overall levels of activity are low (Haynes & Rees, 2005). Using MVPA decoding techniques in the context of imagery, it is possible to train a classifier on perceptual data and to test the classifier on associated imagery data, thereby allowing inferences to be made about the representations involved when one performs mental imagery. Kamitani and Tong (2005) first made use of the decoding method for perceived stimuli, successfully predicting eight different orientations in the visual cortex. This work was extended further by Harrison and Tong (2009) to include working memory (WM), where classification performance significantly above chance level was found across all early visual areas V1-V4 . WM is utilized when a recently presented image is maintained in mind and thus has an inherent link with mental imagery. A recent study has shown that early visual representations involved in both mental imagery and visual WM of the orientation of gratings are similar to those involved in perception (Albers, Kok, Toni, Dijkerman, & de Lange, 2013). Since the earlier mentioned mixed results in V1 were based on univariate activity, it is thus recently becoming clear that MVPA techniques may be a better method of assessing the similarities and differences between mental imagery, WM and perception.

In the current study we are interested in the imagery-related functions of areas making up the visual hierarchy. For perception, V1 is mainly concerned with orientation encoding and local contrasts, while higher visual areas process increasingly complex stimulus features, such as faces or places (O'Craven & Kanwisher, 2000). This is analogous to the concept of local and global feature processing; meaningful global forms may be achieved by combining local image features into a more unified percept. This involves a complex interplay of feed-forward, feedback, and horizontal propagation of information between and within lower- and higher-level areas (Spillmann, 1999). Previous decoding studies on imagery and WM have focused either on grating orientation and the ability to predict a stimulus based on patterns within early visual areas (Harrison & Tong, 2009), or on more global features such as food, tools, faces or places and the decoding of these using patterns in the ventral-temporal cortex (Reddy, Tsuchiya, & Serre, 2010). We would like to look at how imagery representations in the brain differ, by directly comparing in one experiment imagery of both a low-level and a higher-level feature. With this approach we examine whether imagery representations are dependent on the low- or high-level demands of a task, in the same way that perceptual representations are.

A visual area found to be involved in higher-level perception is the Visual Word Form Area (VWFA). The VWFA forms part of the ventral visual stream and has been reproducibly located across individuals, scripts and cultures (Bolger, Perfetti, & Schneider, 2005; Cohen et al., 2000). It has been shown to respond to words more than to other control shapes, such as checkerboards (Cohen et al., 2002) or objects (Szwed et al., 2011) and displays adaptations to orthographic constraints (Binder, Medler, Westbury, Liebenthala, & Buchanan, 2006). These characteristics imply an underlying conceptual nature to the processing carried out by the VWFA. Importantly, Vinckier and collegues also found a gradient of selectivity throughout the area, indicating it has a high degree of functional and spatial organization (Vinckier et al., 2007). Another study targeted activity associated with single letters, locating a region just lateral to the VWFA (Flowers et al., 2004). This suggests perhaps an extension of the VWFA selectivity gradient, as words are fragmented into smaller and smaller components. It has been also been shown for perception that the region responsive to individual letters has greater selectivity for Roman letters than Chinese characters in English speakers (James, James, Jobard, Wong, & Gauthier, 2005), indicating an effect of familiarity on letter perception. It is important to note that univariate activity in these regions is not indicative of representational content. That which is of most interest to us, decoding in the VWFA, has been demonstrated by Braet and colleagues (Braet, Wagemans, & Op de Beeck, 2012), where they retrieved selectivity patterns for visual words over objects. The aforementioned findings provide us with good reason to locate the VWFA (extending to the individual-letter region) as our main Region Of Interest (ROI) for conceptual processing. Thus, we use letter stimuli in our experiment and manipulate these on both a local and global level.

The aim of this study is to decipher the way in which the visual cortex is employed when keeping stimuli in mind. Specifically, we want to know whether the involvement of early visual areas is always required during maintenance of an image, whereby V1 acts as a type of 'sketchpad' (Likova, 2012). According to this view, representations of images are always formed in V1, regardless of what is being imagined. This was a motivation for the earlier mentioned univariate analyses of V1 activity during imagery. This view is logical for perception, since stimuli in the outside world are mapped retinotopically or 'sketched' onto the visual cortex. However, imagery does not necessarily depend on this same mechanism. An alternative to the sketchpad view is that V1 is only required during imagery when zooming in on fine details. According to the reverse hierarchy theory of perceptual learning, learning begins at higher levels of the visual system, and gradually progresses backwards to lower levels only when there is insufficient information (Ahissar & Hochstein, 2004). In this way, V1 may not be needed for imagining stimuli that already have a more conceptual representation in the brain. Therefore, in order to test these alternative theories, we explicitly adjust the focus or goals of the person who is keeping a stimulus in mind. We manipulate the level of detail of imagery by having participants imagine the same stimuli across two different tasks. Previous studies on perception have shown better decoding of orientation than contrast of a grating when orientation is the task-relevant feature (Jehee, Brady, & Tong, 2011; Kok, Jehee, & de Lange, 2012). Thus, one of our tasks targets orientation discrimination of small gratings making up the letter stimuli. During this task we expect decoding accuracy of the orientation to be significantly above chance in V1. The second task involves a feature that should be processed in more conceptual visual areas, specifically the shape of the letter stimulus in its entirety. For this shape task, we would expect to find that, since fine details are not needed to perform the task, decoding accuracy of orientation in the primary visual cortex is decreased during the orientation task, or even diminished to chance level. In this condition, we would thus expect significant classification performance of the letter shape in both V1 and the VWFA (or single-letter) ROI.

A second aspect of this research targets whether there is a difference between known stimuli (real letters) and unfamiliar stimuli (pseudo letters). Since we have seen that familiarity plays a role in letter processing during perception, we would like to examine whether familiarity with a stimulus affects imagery representations across our two tasks. With this approach, we attempt to dissociate the more global content from the lower-level characteristics; for unknown stimuli we expect a diminished conceptual representation, which might therefore demand more informational resources in V1 during the shape task, i.e. better orientation discrimination. This prediction is therefore in accordance with the reverse hierarchy theory. Finally, we also anticipate better decoding of real letters than pseudo letters in the VWFA/single-letter ROI, since the region is associated with such selectivity. In summary, our experimental design targets predictions regarding the patterns of voxels in visual regions during imagery,

with the aim to dissociate these patterns based on ROI, task demands, and familiarity with a stimulus.

2. Materials and methods

2.1 Participants

Thirty-seven right-handed university students participated in the behavioural training session. Participants were healthy, with normal or corrected-to-normal vision and no reading problems. All participants gave written informed consent and received payment for their participation. Twenty participants progressed to the functional magnetic resonance imaging (fMRI) scanning sessions based on behavioural training results. Of those who did not progress, two participants underwent an initial sub-optimal stair-casing procedure (subsequently updated) (García-Pérez, 1998), one took medication with undetermined side-effects on memory and attention, and the remaining fourteen did not meet the required behavioural performance threshold. Two earliest fMRI participants were excluded from all analyses due to technical failure with the fMRI beamer when used in conjunction with a PC running with Linux OS. All subsequent participants underwent the scanning sessions with the stimulus PC running on Windows. Two further participants could not complete the fMRI sessions due to personal circumstances. Therefore, complete analyses were performed individually on data of the remaining sixteen participants (five male, age 18-31 years, M= 23.75 years).

2.2 Procedure

The full experiment consisted of three separate sessions, held on different days. The first session was a 45-minute behavioural training session, where participants performed six blocks of sixteen trials each, to get accustomed to the tasks. The second session was an fMRI task session, where participants performed eight blocks of the task (totalling 128 trials) while in the scanner. This lasted approximately 90 minutes. If time permitted, a T1 weighted anatomical scan was also collected here, otherwise it was carried out in the next session. The second fMRI session consisted of a localizer task in which the task stimuli were presented in a blocked design (15-20 minutes), and population-based receptive field (pRF) mapping measurements were taken (20 minutes). The second fMRI sessions were all held within 14 days of the first. Eye-tracking data were collected for all scanning sessions.

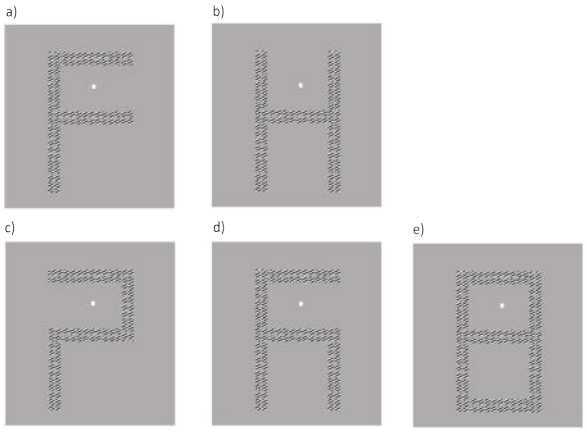


Fig. 1 Stimuli presented during experiment. Real letters **(a)**,**(b)**, and pseudo letters **(c)**,**(d)** were used during the task session. The final pseudo letter **(e)** is only used during the localizer session; it is a template which contains structure of all other letters. The fixation dot is enlarged here for illustrative purposes

2.3 Stimuli

Stimuli consisted of two real letters 'F' and 'H', and two pseudo letters (Fig. 1), all subtending 16.7 degrees of visual angle in the vertical direction and 11.4 degrees horizontally. The letter textures were made up of small grayscale luminance-defined sinusoidal gratings, 0.75 cm in diameter, generated using MATLAB (Mathworks, Natick, MA) with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The gratings were presented as complete circles, rather than in standard annulus form, since they were already in the periphery and therefore were not restricted by the central fixation dot. The gratings had a spatial frequency of 0.18 cycles per degree (cpd), a Michelson contrast of 100% ((Lmax - Lmin) / (Lmax + Lmin), where Lmax is maximum luminance and Lmin is minimum luminance of a grating), and an orientation of either 63 or 142 degrees from the vertical axis. Each grating making up a stimulus had a random phase (0 to 2π) with respect to the other gratings of that stimulus, and had a random phase each time a stimulus was presented. Stimuli were presented on a rear-projection screen using an EIKI projector ((EIKI, Rancho Santa Margarita, CA, 1024 x 768 resolution, 60 Hz refresh rate).

2.4 Experimental paradigm

We conducted a visual mental imagery experiment, in which participants were briefly presented with two

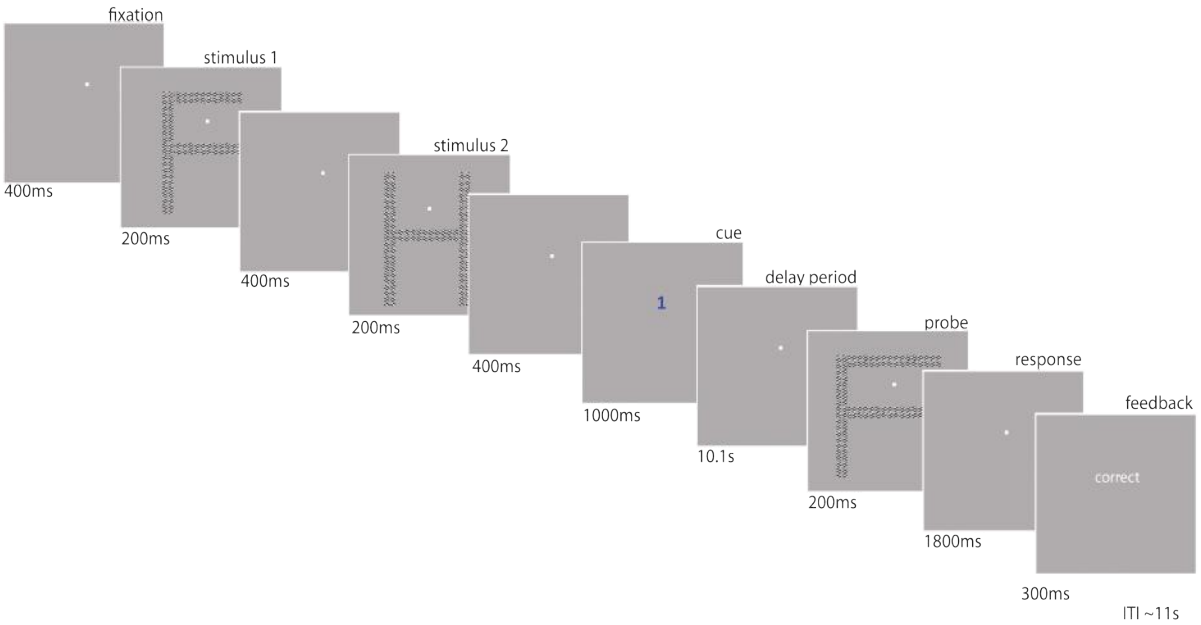


Fig. 2 Experimental design. This is a shape-task-real-letter trial; two letter stimuli are presented at the start of the trial, and then participant is cued to keep the first stimulus in mind (blue cue indicated shape task). Participant maintains image of this for ~10s duration of the delay period, and then judges whether the probe is the same or different from the imagined stimulus. Feedback follows their response. An inter-trial interval of ~11s separates each trial.

stimuli and were cued ("1" or "2") to keep one of the stimuli in mind for a ten second imagery period (Fig. 2). The cue was also coloured red or blue to indicate an orientation- or shape-task trial respectively. At the end of the imagery period, a probe was presented, which was either the same as or different from the stimulus participants kept in mind. Participants were instructed to respond to the probe with either an index- or middle- finger button-press of the right hand, indicating that the probe was the same as or different from the maintained stimulus, respectively. The probe could be different from the maintained stimulus in one of two ways; either the orientation of the lines of the gratings making up the probe could be different (rotated clockwise or anti-clockwise with respect to the gratings of the stimulus), or the shape of the probe could be different (stretched or compressed in the horizontal or vertical dimension). These changes occurred 50% of the time in each of the tasks. All probe stimuli could be changed on either or both of these dimensions for any given trial, but performance was based on making correct responses in the cued dimension only.

2.5 Staircase procedure

In order to match difficulty across the two tasks, a two-down/one-up fixed-step-size (FSS) adaptive staircase procedure was used (García-Pérez, 1998). There were two staircases for each of the orientation and shape tasks: one for real letters and one for pseudo letters, rendering four staircases in total. These staircases were established to converge at a performance level of 75-80% correct answers. In practice, if participants began to perform above this point, the difference between the imagined and test stimulus in subsequent trials became smaller and thus harder to detect, and vice versa for underperformance. For the training session, the staircase was seeded with an orientation difference of 20 degrees and a total shape difference of 3 gratings in either the horizontal or vertical dimension (divided equally left/right or top/bottom respectively, i.e. 1.5 gratings on either side). With this staircase procedure every trial was balanced for difficulty, regardless of task or stimulus to be imagined.

Training performance was thus based on three related factors: number of correct answers, orientation staircase values (for real and pseudo letters), and shape staircase values (for real and pseudo letters). These staircase values were fed from each block to the next so that convergence could be reached not only within, but also across blocks. Participants were included in the fMRI experiment if, during training, they had a minimum of 11/16 correct trials, a maximum of 30 degrees detectable difference between the orientation of gratings of the starting stimulus and the probe, and a maximum detectable difference of 3 gratings in each of the horizontal or vertical dimensions (1.5 gratings added to or removed from the left and right of the stimulus, or added to or removed from the top and

bottom). These thresholds were taken as reasonable measures of visual sensitivity and understanding of the tasks involved.

2.6 Training

Training consisted of a behavioural session held in a darkened, sound-attenuated room, using a PC running on Linux. Participants underwent six blocks of 16 trials each during the training session. In the first two blocks participants were trained on the trial sequence and tasks involved. Participants were presented with orientation-cued trials only in the first block, and shape-cued trials only in the second block. This helped to build up separate representations for each of the task cues. The third block (and all further training blocks) contained randomly intermixed orientation- and shape-cued tasks. In the first three blocks, stimuli were presented for one second. In the fourth block, the presentation duration of the stimuli was reduced to 0.2 seconds, as this was the duration of stimulus presentation during the fMRI scanning session. The first four blocks each began with a maintenance period of 5 seconds, increasing to 10 seconds as the block progressed. The final two blocks had a fixed maintenance period of 10 seconds, the same as during the fMRI session.

2.7 Localizer

During the second fMRI session, participants underwent a localizer scanning run. They carried out a 2-back fixation task which required an index finger response each time the fixation dot turned the same color as it was two colors ago. The same letter stimuli as were used during the task session were presented in the periphery. The fixation task ensured attention was kept at the centre, so that localizer brain activity was not representative of fluctuations in attention towards the stimuli. Exactly the same stimuli from the task were used; the shape extended 16.7 x 11.4 degrees of visual angle, with a grating orientation of either 63 deg or 142 deg for each of the four real/pseudo letters, in order to maintain a position-invariant reference frame for participants while fixating. This is important for decoding analyses, since individual voxels are sensitive to specific spatial positions of the visual field, with activity in each of these voxels contributing to the overall pattern analysis. The stimuli were presented in a blocked design, each letter for 12 seconds at a frequency of 4 Hz, i.e. the letters flashed on and off every 250ms. In addition to the four task stimuli, an extra stimulus (resembling a digital form of the number '8', Fig. 1E.) was included. This was, in fact, the template from which all the other letters were constructed, i.e. the top and bottom horizontal lines were removed to produce the letter 'H' for example. Each block included 12 seconds presentation of each letter and the template stimulus. The fixation dot alone (still changing colour) was presented for the final 12 seconds of the block. Thus, each block lasted 72 seconds. Twelve blocks were collected, with a random order of the letter stimuli in each. The two possible grating orientations were applied to each letter equally, for a total of six presentations of each orientation per letter across the entire session. The fixation task was clear to all but one participant, who stopped making button presses as the run progressed. At the end of the run the participant reported that he maintained fixation at all times, and the eye-tracker data from the session confirmed this.

2.8 pRF mapping

We used population-based receptive field (pRF) mapping to estimate the visual field map (Dumoulin & Wandell, 2008). Five short blocks (3.5 minutes per block) of translating bars, consisting of black and white checkerboards were presented during the pRF session, beginning as a small wedge at one edge of the screen and growing in size as the bars moved across the fixation dot. Within each block, the bars moved across each of the vertical, horizontal and diagonal planes. Participants had to maintain fixation and make a response each time the fixation dot changed from red to green or vice versa. The moving bars covered 11.5 degrees of the visual field, which is equivalent to the horizontal size of the letter stimuli used during the task fMRI session, but did not cover the extremities of the letter stimuli in the vertical direction. This lack of coverage was only for the bottom corner extremities of the letters, as the fixation dot during the task was positioned higher than that during the pRF mapping.

2.9 fMRI data acquisition

Participants were scanned using a 3 Tesla Trio Magnetic Resonance Imaging (MRI) system (Siemens, Erlangen, Germany). Once the participant's position in the scanner was established using a localizer, the brain was automatically outlined by AA Scout. AA Scout takes two low-resolution whole-head scans, and compares this to a reference atlas on the scanner. This is used to position all future scans in

the session. This outline was manually adjusted for one participant and then used for all subsequent participants, to collect data in a consistent manner across participants. Sometimes the cerebellum or part of the ventral temporal lobe were not enclosed within the AA Scout outline, however, since the regions of interest were visual processing regions, this was not a concern. Functional images were acquired using a T2* weighted 3D gradient-echo Echo Planar Imaging (EPI) sequence, with the following parameters: repetition time (TR) = 1500ms, echo time (TE) = [25 ms], flip angle (FA) = 15 deg, field of view (FOV) = 224 x 224 x 52, voxel size = 2x2x2 mm. To control for T1 equilibrium effects the first six volumes (1.5TR * 6 volumes = 9 seconds) were discarded from the task and localizer sessions, and the first eight volumes (1.5TR * 8 volumes = 12 seconds) were discarded from the pRF mapping session. Structural images were acquired at 1mm isotropic resolution using a T1-weighted Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequence (192 slices, TR = 2300ms, TE = 3.03ms, FA = 8 deg, voxel size = 1 x 1 x 1 mm).

2.10 Behavioural data analysis

We collected reaction times (RTs) and accuracies of responses for all participants during the task session. Each of these was separately analyzed under the following three conditions: task (orientation or shape), letter (real or pseudo), and cued stimulus (first or second). Analysis involved carrying out two-tailed paired samples t-tests for these conditions. RTs were also compared to the accuracies by calculating the Pearson correlation coefficient for the different conditions, to assess if there was a speed-accuracy trade-off.

2.11 fMRI data preprocessing

The functional sessions (task, localiser, and pRF mapping) and structural scan were preprocessed on an individual, per subject basis using SPM12b (http://www.fil.ion.ucl.ac.uk/spm, Wellcome Trust Centre for Neuroimaging, London UK). After conversion to nifty file format, functional images were realigned to a mean functional image and coregistered to the structural scan.

2.12 fMRI data analysis

2.12.1 Localizer

The blocked design of the localizer provided five conditions; four different letter stimuli and one "template" for all letters. The fixation dot changed colour throughout the entire session, thus, the fixation-only presentations at the end of each block were an implicit baseline. The localizer was modelled by convolving the blocked durations (12 seconds per condition per block) with a canonical haemodynamic response function. This model was specified in a general linear model (GLM), which included the five conditions and six motion parameters. We collapsed all five conditions into one regressor to create a statistical t-map based on the perception of all letter stimuli versus fixation.

2.12.2 ROI selection

As explained in the introduction, we were interested in the visual processing of local or global details of stimuli, therefore we had two ROIs; one related to the processing of fine details such as orientation (early visual regions), and one related to the perception of whole letters on a more

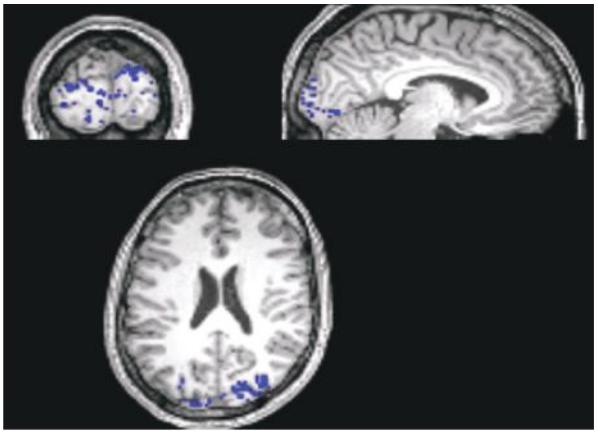


Fig. 3A Early Visual ROI for participant SB17. This was made using an inclusive retinotopy mask ($p < 0.001$, 42 voxel clusters) over localizer activity ($p < 0.001$).

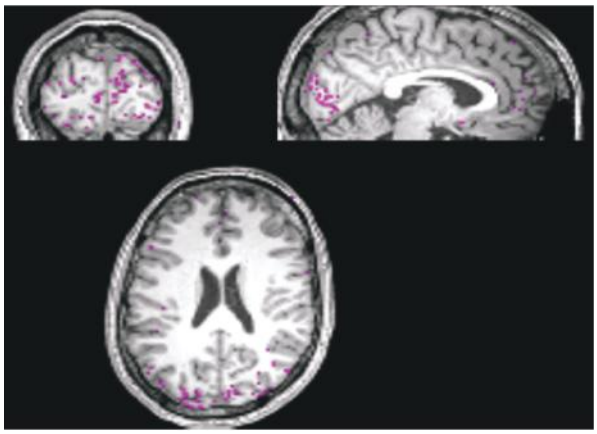


Fig. 3B Letter-relevant ROI for participant SB17. This was made using an exclusive retinotopy mask ($p < 0.001$, 42 voxel clusters) over localizer activity ($p < 0.001$).

conceptual level. In order to define our early visual ROI, we firstly used the data collected from the pRF session to find voxels whose receptive fields were most responsive to the spatial locations covered by the moving bar stimuli. The onsets and durations of the moving bars from each of the five pRF blocks were specified in a GLM, again using fixation as the implicit baseline. A t-contrast was made for moving bars>fixation at an uncorrected $p < 0.001$ for clusters with a spatial extent > 42 voxels. Next, we took the t-map created from the localizer session (where participants viewed blocked presentations of the task stimuli) and used the t-contrast from the pRF mapping as an inclusive mask over the localizer t-contrast. This combined map was analysed at an uncorrected $p < 0.001$ for all voxels (since clusters of 42 voxels had already been specified for the pRF map). This map thus located voxels responsive to both the localizer letter stimuli and the pRF checkerboards.

As previously mentioned, the VWFA has been shown to process words and letters. In order to locate this, the pRF t-map was used here as an exclusive mask over the localizer t-map. Again, this combined map was analysed at an uncorrected $p < 0.001$ for all voxels. The resulting activation thus highlighted significant activation caused by the letter stimuli, excluding the perception of simply any stimulus at those spatial locations covered by the letter stimuli. We used this statistical map as our conceptual (letter-relevant) ROI. Fig. 3 shows an example of the ROIs obtained for one participant (SB17).

2.12.3 Task multivariate analysis

MVPA provides a method of extracting stimulus-specific information in the pattern of activity of distributed voxels within a region of interest. We trained several linear support vector machines (SVMs) based on the patterns of blood oxygen level dependent (BOLD) activity over voxels in each of our ROIs. There were two different dimensions in which we could train our SVMs; to discriminate between the two possible orientations or between the four letters. This was carried out for either the orientation- or the shape-task (task-level), and in either the early visual or conceptual ROI (ROI-level). Thus, our classification performance measures describe the amount of orientation or letter information available in the BOLD response pattern over voxels for a particular task. For all decoding analyses, the 150 most active voxels were chosen for each ROI per participant. All trials regardless of a correct or incorrect response were taken for decoding purposes. The minimal differences in the orientation and shape of the probe as given by the staircase procedure meant that voxels most responsive to one particular orientation (e.g 63 deg) would still be activated by a small inaccuracy in the mental image of this orientation, based on the properties of those voxels' tuning curves, and would be sufficiently different from those activated by another distinct orientation (142 deg). First, we both trained and tested the classifier on the delay period. For the training and testing sets we averaged activity over time window 5 to 9.5 seconds after the onset of the imagery period. This time window was chosen to maximize the temporal distance from stimulus related activity evoked by the presentation of the two stimuli and cue at the start of each trial. After averaging, data were normalized by z-scoring. For each of the orientation and shape tasks the classifier was trained and tested on 64 trials per task. This was broken down into 32 v 32 trials for orientation decoding (63 deg v 142 deg), and 16 v 16 trials for letter decoding (real letters 'F' v 'H', or pseudo letters 'pseud1' v 'pseud2'). For orientation decoding, the classifier was trained using a 32-fold cross-validation procedure; first it was trained on stimuli in all but two trials for the condition we were interested in and then tested on stimuli for the remaining two trials. For letter decoding, a 16-fold cross-validation procedure was used, in the same manner. Since we were also interested in any differences in decoding accuracies of orientation when either a real or pseudo letter was presented, the orientations were then also divided up per real or pseudo letter. The trials were not equally balanced across these comparisons, so in order to keep classification unbiased, i.e. instead of training the classifier on one class more than the other (KrishnaVeni & Sobha Rani, 2011), we took the condition with the lowest number of unbiased trials per participant and used this as the size of the class training sets for the rest of the biased classification conditions. This led to a reduction in the power of classification; however, the size of the training set was never below 20 trials per classification analysis. Once decoding accuracies were calculated, we also assessed whether there were any correlations between these accuracies and average staircase values, to see if greater sensitivity in the perception of a stimulus was correlated with a better representation of the stimulus during the delay period. Since we wished to carry out statistics on the resultant classification results based on the assumption of parametric data, we also analysed the skewness and kurtosis of these classification distributions. These measures were converted to

z-scores and compared with the expected z-scores from a normal distribution. An absolute value of less than 1.96 for skewness and kurtosis z-scores was taken as insignificant, allowing subsequent gaussianity-based statistics to be carried out.

Finally, we trained the classifier on perception data and tested on the delay period. We took the univariate activity associated with the two orientations or the four letters during the localizer scan, and used this for training. Thus, there was no need for a cross-validation procedure; all of the delay period data could be tested at once. Here we looked at individual scans across the whole trial, and focused on one time point at the start of the imagery period (6.5 seconds after the onset of imagery) and another time point after the probe was presented (14 seconds after the onset of imagery period). The 6.5s time-point was chosen to avoid picking up activity related to the stimuli presented at the start of the trial (allowing for the BOLD delay). The 14s time point was chosen also allowing for the BOLD delay in picking up activity directly related to the presentation of the probe at the end of the trial.

Our first step for each training-set in multivariate analyses was to check whether decoding was indeed possible in our two ROIs. This was carried out by comparing all decoding classification accuracies with chance level. Following that we then choose specific conditions to compare, to assess if our hypotheses and predictions were validated by the data.

2.13 Combined behavioural and fMRI analyses

Behavioural accuracies were compared with classification accuracies for all task conditions to check if better decoding was achieved in participants who made correct responses to the probe. Assuming correct answers are an indication of precise imagery representations, we would expect a positive correlation between response and decoding accuracies.

The staircases were used to match task difficulty and ensure that participant's responses converged at 75% correct answers. For all four staircases (orientation thresholds for real or pseudo letters, shape thresholds for real or pseudo letters), larger values signified less sensitivity towards the stimuli, i.e. a greater difference in the stimulus and probe was necessary for the participant to detect it, and vice versa. Thus, we calculated average staircase values per participant and compared real against pseudo letters for each of the orientation and shape tasks, using a two-tailed paired samples t-test.

3. Results

3.1 Behavioural results

The sixteen participants responded to all probe stimuli, so all 128 trials per participant could be used for behavioural analyses. Reaction times (RTs) of responses did not differ across task (orientation or shape), letter (real or pseudo), or cued stimulus (first or second stimulus) conditions (all $p > .05$). Accuracies also did not show any difference for cued stimulus, however, there was a distinction between the orientation and shape task (76 % and 64 % correct respectively; $t(26) = 4.37$, $p = .0001$). Splitting these into real or pseudo letters per task show that task influenced accuracies of both real letters (orientation: 78 %, shape: 63 %; $t(26) = 4.09$, $p = .0003$) and pseudo letters (orientation: 75 % , shape: 64 %; $t(26) = 3.35$, $p = .002$). Comparing RTs and accuracies revealed no significant correlations, thus, there was no speed-accuracy trade-off. Both correct and incorrect trials were taken for decoding purposes, since incorrect trials did not necessarily mean participants did not have a sufficiently robust mental representation for decoding purposes, only that it was not accurate enough for our specific task.

3.2 fMRI results

A full list of relevant figures and conditions is provided in Table 1 and 2.

3.2.1 Training on imagery

For imagery training, we trained and tested the classifiers on averaged imagery activity over time points 5.0 to 9.5 seconds after the onset of the imagery period. Skewness and kurtosis z-scores were calculated from the classification distribution per condition and were found to be insignificant when compared to expected z-scores from a normal distribution.

Decoding orientation

The stimulus gratings from all trials were divided equally into 63 degrees and 142 degrees orientations.

Table 1. Indicating relevant figures for the various decoding conditions using the imagery classifier.
Significance levels: * p<0.05, ** p< 0.01, *** p<0.001. Curly bracket denotes a significant difference of p<0.05 between conditions.

ROI		Early visual				Letter relevant			
Classification feature		Orientation		Shape		Orientation		Shape	
Task	**Letter stimuli**								
Orientation	Real	Fig. 4	Fig. 6*	Fig. 5	Fig. 7	Fig. 4	Fig. 6*	Fig. 5*	Fig. 7
	Pseudo		Fig. 6		Fig. 7		Fig. 6		Fig. 7
Shape	Real	Fig. 4	Fig. 6	Fig. 5*	Fig. 7	Fig. 4*	Fig. 6	Fig. 5	Fig. 7
	Pseudo		Fig. 6**		Fig. 7		Fig. 6*		Fig. 7

Table 2. Indicating relevant figures for the various decoding conditions using the perception classifier. Significance levels: * p<0.05, ** p< 0.01, *** p<0.001. Curly bracket denotes a significant difference of p<0.05 between conditions.

ROI		Early visual				Letter relevant			
Classification feature		Orientation		Shape		Orientation		Shape	
Time point in task		Image	Probe	Image	Probe	Image	Probe	Image	Probe
Task	**Letter stimuli**								
Orientation	Real	Fig. 8	Fig. 8	Fig. 10	Fig. 10 ***	Fig. 9	Fig. 9	Fig. 11	Fig. 11***
	Pseudo	Fig. 8	Fig. 8	Fig. 10	Fig. 10 **	Fig. 9	Fig. 9	Fig. 11	Fig. 11***
Shape	Real	Fig. 8	Fig. 8	Fig. 10	Fig. 10 ***	Fig. 9	Fig. 9*	Fig. 11	Fig. 11**
	Pseudo	Fig. 8*	Fig. 8	Fig. 10	Fig. 10 **	Fig. 9*	Fig. 9	Fig. 11	Fig. 11***

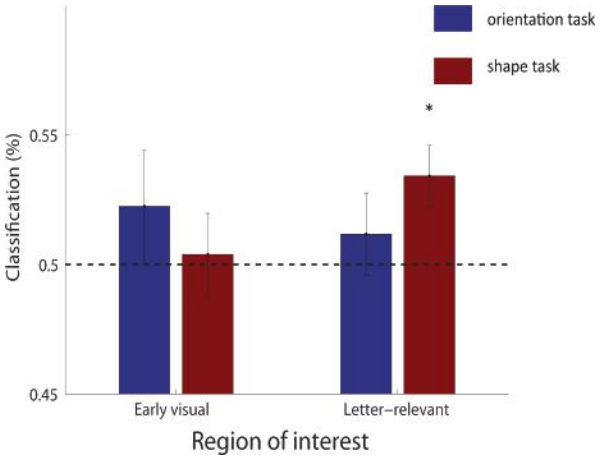


Fig. 4 Decoding orientation during the orientation and shape task in both ROIs. Real and pseudo letters are not distinguished in this analysis. Error bars represent SEM. Significance levels: * p < 0.05, ** p < 0.01, *** p < 0.001.

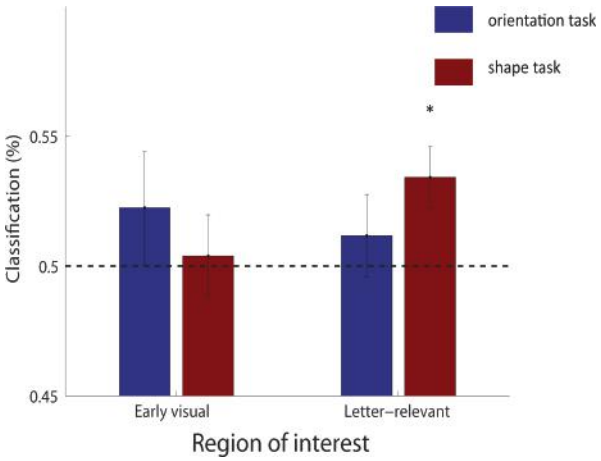


Fig. 5 Decoding letter shape during the orientation and shape task in both ROIs. Real and pseudo letters are not distinguished in this analysis. Error bars represent SEM. Significance levels: * p < 0.05, ** p < 0.01, *** p < 0.001.

The classifier was trained and tested on each of these, for the two tasks (orientation or shape task), and in the two ROIs (early visual ROI or letter-relevant ROI). We could decode significantly above chance level (>50% accuracy) for one condition only, during the shape task in the letter-relevant ROI (decoding accuracy 53 %, $t(15) = 2.88$, $p = .0057$). See Figure 4 for classification accuracies of all orientation decoding conditions.

Decoding letters

The stimuli from all trials were equally divided across the four letters. To decode all letters, irrespective of familiarity, we used six classifiers per orientation or shape task. For both tasks the mean classification accuracy was calculated from the accuracies of each letter compared to each other letter, i.e. the mean accuracy of 'F' v 'H', 'F' v 'pseud1', 'F' v 'pseud2',

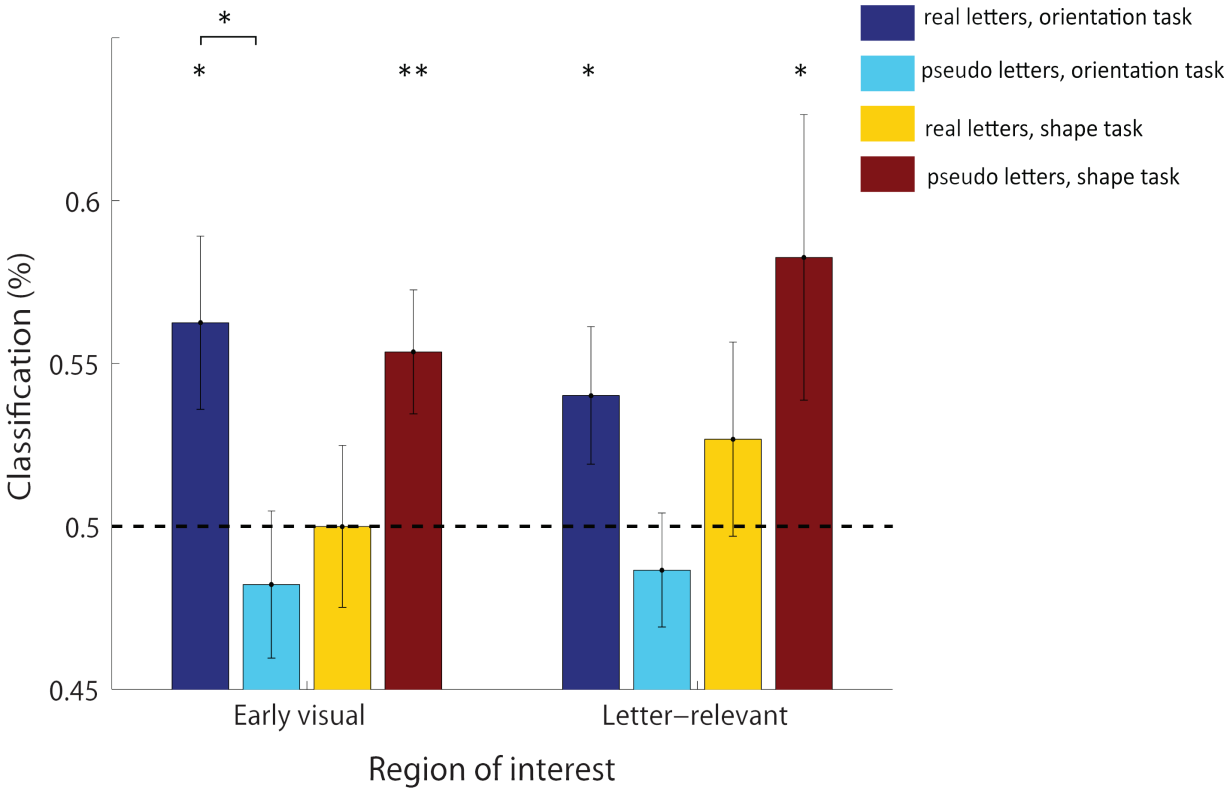


Fig. 6 Decoding orientation of real and pseudo letters during the orientation and shape task in both ROIs. Error bars represent SEM. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

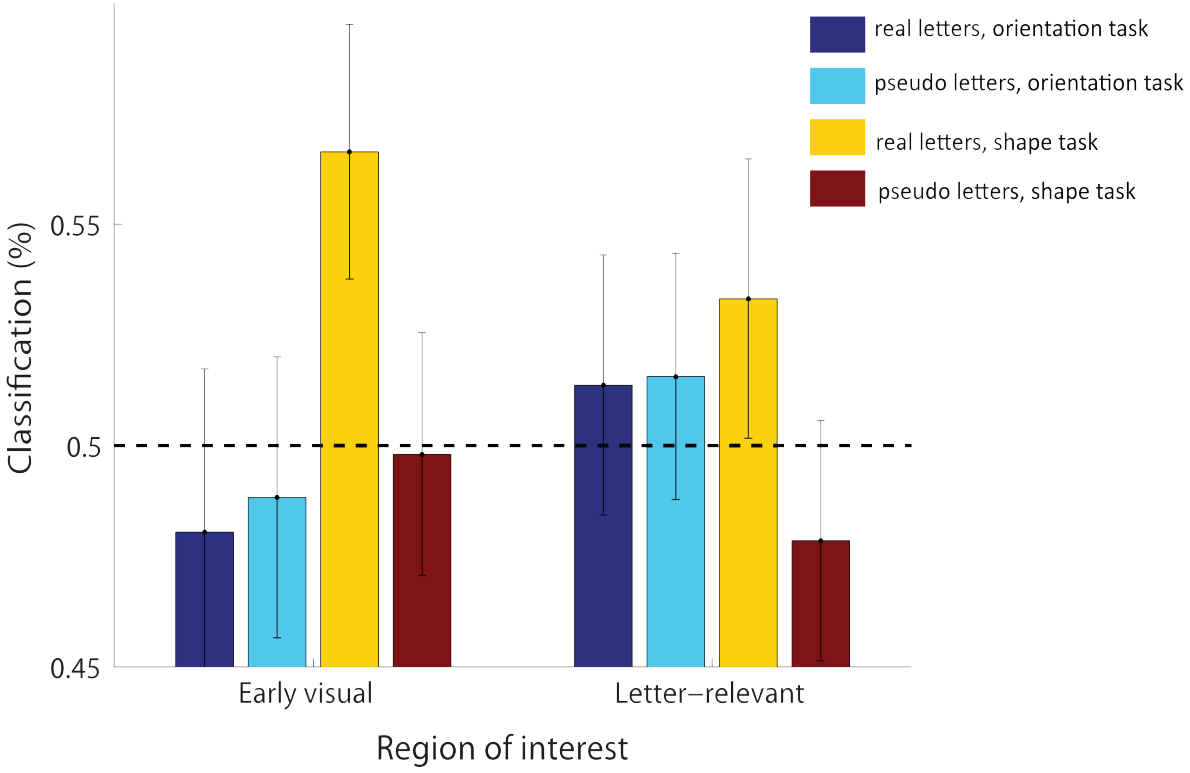


Fig. 7 Decoding letter shape of real and pseudo letters during the orientation and shape task in both ROIs. Error bars represent SEM. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

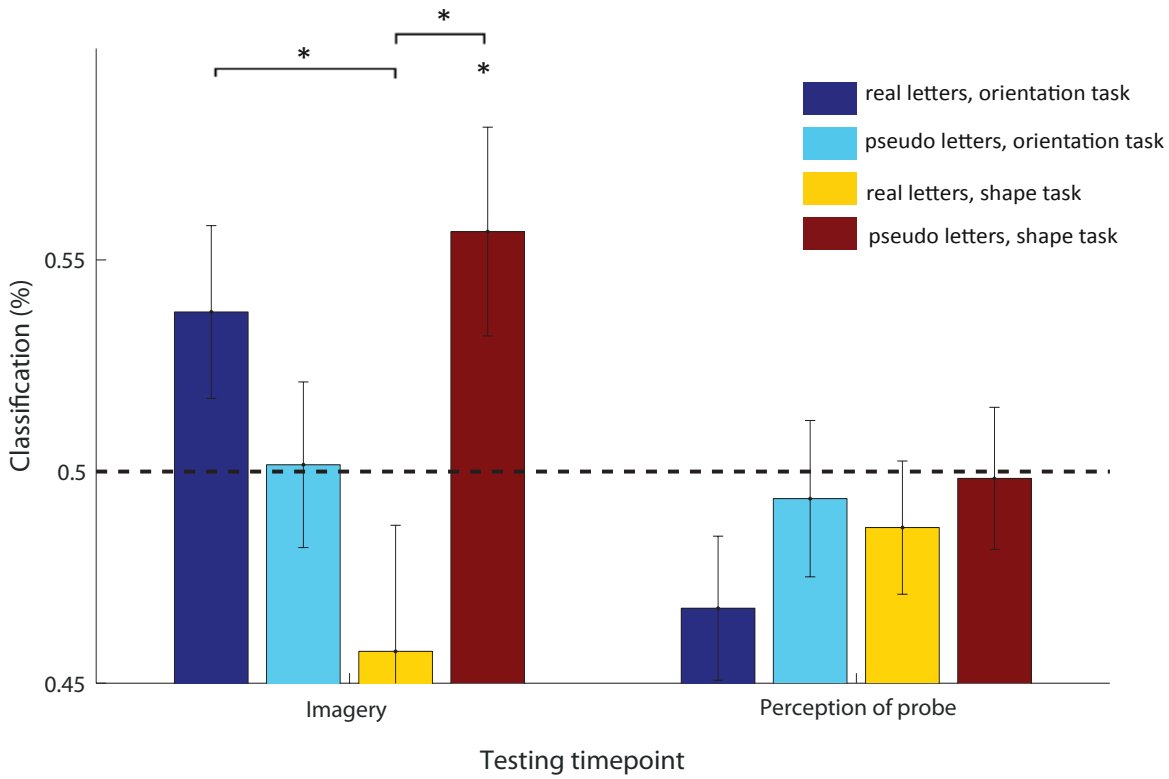


Fig. 8 Decoding orientation of real and pseudo letters in the early visual ROI, when training on stimulus-driven activity during the localizer. Imagery decoding is taken 6.5 seconds after the onset of the imagery period, decoding during perception of the probe is taken 4 seconds after probe is presented (allowing for BOLD delay). Error bars represent SEM. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

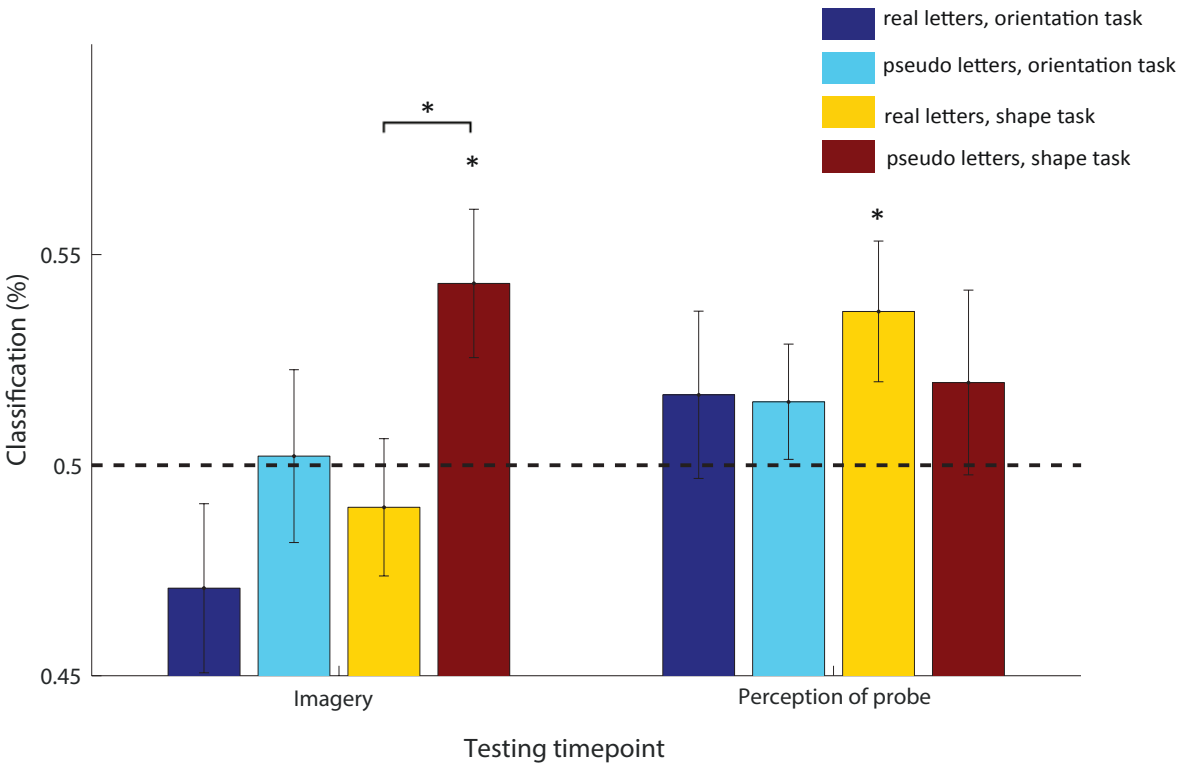


Fig. 9 Decoding orientation of real and pseudo letters in letter relevant ROI, when training on stimulus-driven activity during the localizer. Imagery decoding is taken 6.5 seconds after the onset of the imagery period, decoding during perception of the probe is taken 4 seconds after probe is presented (allowing for BOLD delay). Error bars represent SEM. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

'H' v 'pseud1', 'H' v 'pseud2', 'pseud1' v 'pseud2'. Here we found information during the shape task in the early visual ROI (decoding accuracy 53 %, t(15) = 2.26 , p = .02) and during the orientation task in the letter-relevant ROI (decoding accuracy 53 %, t(15) = 2.36, p = .016) . Figure 5 contains all results for letter decoding.

Decoding orientation based on familiarity

To tackle our second research question, we again tried to decode orientation of the stimulus gratings, however, this time the orientations were categorized according to which letter they belonged to, and classification was performed between the two orientations for all cued real letters 'F' and 'H' and between the two orientations for all cued pseudo letters 'pseud1' and 'pseud2'. Real and pseudo letters were also counterbalanced across the two tasks, and so the impact of familiarity could be fairly compared with results from the first research question (decoding in general, regardless of familiarity). Decoding accuracies were obtained using four classifiers per ROI; classifying orientation during the orientation task for real letters, classifying orientation during the orientation task for pseudo letters, classifying orientation during the shape task for real letters, and classifying orientation during the shape task for pseudo letters (Fig. 6). We found that the early visual ROI contained information about the orientation of real letters during the orientation task (decoding accuracy 56 %, t(15) = 2.35, p = .016), and this was significantly higher than decoding of pseudo letters during the orientation task in this region (t(30) = 2.30, p = .029). In the early visual ROI we also found representations for pseudo letters during the shape task (decoding accuracy 55 %, t(15) = 2.82, p = .0065). In addition, we found similar classification in the letter-relevant ROI; with information about real letters during the orientation task (decoding accuracy 54 %, t(15) = 1.90, p = .038) and information about pseudo letters during the shape task (decoding accuracy 58 %, t(15) = 1.88, p = .04).

Decoding letter shape based on familiarity

The same division between real and pseudo letters was applied to classifiers which were then trained to decode letter shape rather than orientation, again using four classifiers per ROI. Letter information was found only for real letters during the shape task in the early visual ROI (decoding accuracy 57 %, t(15) = 2.31, p = .018), see Figure 7. Pseudo letter information in this ROI was found to be non-significant, along with all conditions for the letter-relevant ROI.

3.2.2 Training on perception

We also used stimulus-driven activity from the localizer to train the classifier, and test across all scans of the trial for each condition. We focused on time points 6.5 s (imagery representations) and 14 s (probe-related representations).

Decoding orientation for real and pseudo letters Imagery

Our early visual ROI showed imagery period information (Fig. 8). This again included information on pseudo letters during the shape task (decoding accuracy 56 %, t(15) = 2.30, p = .036). This was significantly more information than for real letters during the same condition (pseudo letters: 56 %, real letters: 46 %; t(30) = 2.57, p = .016). There was also significantly better decoding of real letters during the orientation task than during the shape task (orientation task: 54 %, shape task: 46 %; t(30) = 2.22, p = .034). The letter-relevant ROI (Fig. 9) contained imagery-related information also for pseudo letters during the shape task (decoding accuracy 54 %, t(15) = 2.45, p = .014), and better decoding of pseudo over real letters during the shape task (t(30) = 2.22, p = .034). When we compared imagery period results from training on perception with training on imagery, we found more information in real letters from training on imagery. This was for both orientation decoding during the orientation task in the letter-relevant ROI (imagery training decoding accuracy: 61 %, perception training decoding accuracy: 47 %; t(30) = 2.38, p = .024), and for letter decoding during the shape task in the early visual ROI (imagery training decoding accuracy: 56 %, perception training decoding accuracy: 49 % ; t(30) = 2.18, p = .038).

Probe

All decoding analyses in the early visual ROI related to perception of the probe were insignificant (Fig. 8). Decoding of real letters during the shape task in the letter-relevant ROI was significant however, during perception of the probe (Fig. 9; decoding accuracy 54 %, t(15) = 2.19, p = .023).

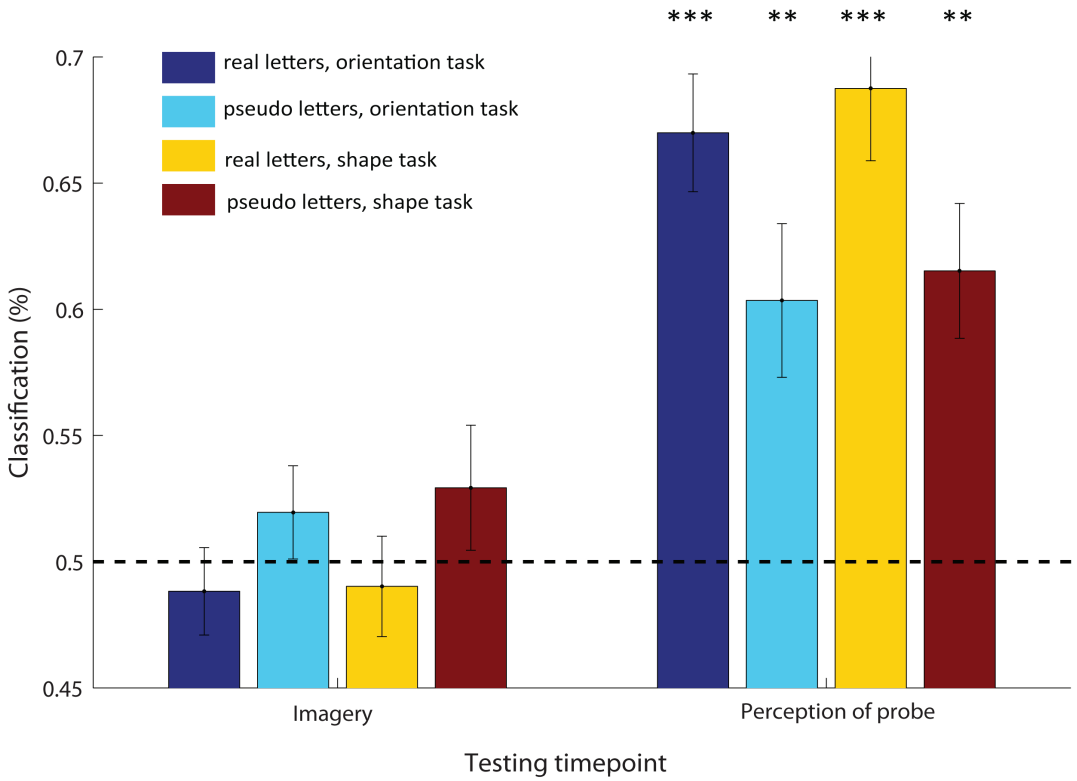


Fig. 10 Decoding letter shape of real and pseudo letters in the early visual ROI, when training on stimulus-driven activity during the localizer. Imagery decoding is taken 6.5 seconds after the onset of the imagery period, decoding during perception of the probe is taken 4 seconds after probe is presented (allowing for BOLD delay). Error bars represent SEM. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

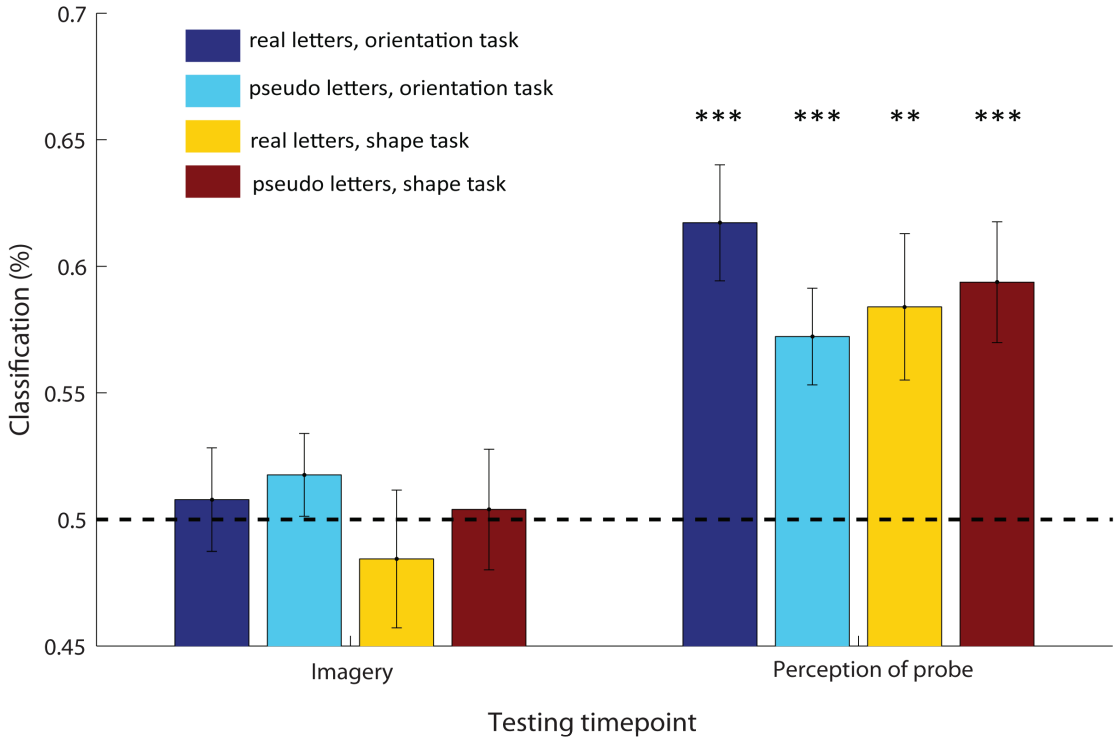


Fig. 11 Decoding letter shape of real and pseudo letters in letter-relevant ROI, when training on stimulus-driven activity during the localizer. Imagery decoding is taken 6.5 seconds after the onset of the imagery period, decoding during perception of the probe is taken 4 seconds after probe is presented (allowing for BOLD delay). Error bars represent SEM. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Decoding letter shape for real and pseudo letters imagery

Decoding of letter shape showed no significant classification during the imagery period, for either ROI (Figs. 10, 11).

Probe

In contrast to imagery, we found highly significant shape representations during perception of the probe in both the early visual ROI (Fig. 10) and letter-relevant ROI (Fig. 11) for real letters during the orientation task (early visual: decoding accuracy 67 %, $t(15) = 7.28$, $p < .0001$; letter-relevant: decoding accuracy 62 % , $t(15) = 5.12$, $p < .0001$), pseudo letters during the orientation task (early visual: decoding accuracy 60 %, $t(15) = 3.40$, $p = .002$; letter-relevant: decoding accuracy 57 %, $t(15) = 3.79$, $p = .0009$), real letters during the shape task (early visual: decoding accuracy 69 %, $t(15) = 6.54$, $p < .0001$; letter-relevant: decoding accuracy 58 %, $t(15) = 2.90$, $p = .005$), and pseudo letters during the shape task (early visual: decoding accuracy 62 %, $t(15) = 4.31$, $p = .0003$; letter-relevant: decoding accuracy 59 %, $t(15) = 3.93$, $p = .0007$). Lastly, a 2 (feature: orientation or shape decoding) x 2 (ROI) x 2 (task) x 2 (real or pseudo letters) repeated measures ANOVA revealed a highly significant main effect of feature ($F(1,15) = 53.255$, $p < .0001$) and a significant feature*ROI interaction ($F(1,15) = 12.25$, $p = .003$).

3.3 Combined behavioural and fMRI results

Analyses of decoding accuracies from imagery training and behavioural response accuracies showed no significant correlations. A positive correlation was found, however, between orientation decoding accuracies of real letters during the orientation task in the early visual ROI and the mean orientation staircase value ($r = .63$, $p = .01$), signifying that we could decode more real letter orientation information in the early visual ROI for participants who had a greater difference between the stimulus and probe grating orientation.

4. Discussion

In this study, we used MVPA to investigate how imagined stimuli are represented in the visual cortex. We researched how these representations differ, depending on whether participants focused on a low-level aspect of the stimulus, i.e. the orientation of gratings making up the stimulus, or on a higher-level feature, i.e. the shape of the stimulus. We further assessed whether familiarity plays a role in these representations, using letters known to participants and used on a daily basis, and using pseudo letters, which participants had no prior experience with. We hypothesized that these low- and high-level features might be better represented in different areas of the visual cortex, and so used stimulus-driven activity to define two ROIs; one associated with the perception of gratings at the earliest visual level (Kamitani & Tong, 2005), and one associated with the perception of letters at a more conceptual level (Braet et al., 2012; Cohen et al., 2002; Flowers et al., 2004). We also performed decoding analyses in two ways; by training the classifier on orientation/shape during imagery from other trials of the same task condition, and by training the classifier on the stimulus-driven activity induced by the presentation of orientation/shape during the localizer session.

4.1 Representations based on task demands

Decoding orientation during the orientation task was not possible in the early visual ROI (Fig. 4). This is not in agreement with our prediction, which was based on previous studies showing successful decoding of grating orientation during working memory and imagery (Albers et al., 2013; Harrison & Tong, 2009). Instead, we found orientation decoding in an unexpected condition and area; during the shape task in our letter-relevant ROI (Fig. 4). This is a surprising result since it has been demonstrated for perception that task-relevant features can be decoded better than task-irrelevant features (Jehee et al., 2011; Kok et al., 2012). For the shape task, we indeed found significant classification of letter shape in the early visual ROI, as predicted (Fig.5). This is in line with the 'visuo-spatial sketchpad' theory (Likova, 2012); we have shown that when people imagine the global aspect of a stimulus when required to do so, then this feature is represented in the early visual cortex. We also found decoding of letter shape in our letter-relevant ROI, but during the orientation rather than shape task. This may be explained by the effect of "global precedence" (Hughes, Fendrich, & Reuter-Lorenz, 1990), whereby people still process global structure even when it is required to focus on local details. Lastly, we could not decode letters in the letter-relevant ROI during the shape task. Since this region was defined specifically for the decoding

of letter shape, this outcome is again unexpected, and not in agreement with our predictions.

Our second research question sheds some light on the aforementioned surprising results. We subsequently found that familiarity with a stimulus does indeed play a role in the decoding of representations in the visual cortex. Specifically, for orientation decoding, our gratings were placed within familiar or unfamiliar global structures, which may explain why results of our first research question do not show any orientation decoding in the early visual ROI. In contrast, previous orientation decoding studies on imagery used large annulus gratings as stimuli (Albers et al., 2013; Harrison & Tong, 2009).

4.2 Representations based on familiarity

In both the early visual and letter-relevant ROIs we were able to classify the orientation of real letter gratings during the orientation task, and the orientation of pseudo letter gratings during the shape task (Fig. 6). Also in the early visual ROI, we found significantly better decoding of the orientation of real over pseudo letters during the orientation task. Thus, for familiar stimuli we can reliably decode the fine detail of the stimulus when this is the task-relevant feature. As predicted, we could not decode grating orientation in familiar stimuli during the shape task. This is in accordance with the reverse hierarchy theory of perceptual learning (Ahissar & Hochstein, 2004), which suggests that when a higher-level representation has already been established, i.e. global shape, representations of low-level features are no longer required. In contrast, we cannot decode the fine detail of unfamiliar stimuli when this is the task-relevant feature. This orientation information can only be classified in unfamiliar stimuli when people are required to focus on the global aspect of the stimulus. Thus, it may be that we maintain information about both the finer and more global details of unfamiliar stimuli when required to focus on the shape of the stimulus. However, when we tried to decode the shape of pseudo letters during the shape task, this was not possible, and classification of the shape of pseudo letters was significantly lower than classification of their grating orientation during the shape task, in both ROIs (Figs. 6, 7). Taken together, these findings resonate with our prediction that it is difficult to maintain a more conceptual representation of unfamiliar stimuli (James et al., 2005). However, rather than seeing this dissociation illustrated between ROIs as we predicted, i.e. less representation of unfamiliar stimuli in the letter-relevant ROI, leading to better representations in the early visual ROI, we instead see this being captured in each of our ROIs. The fact that we can decode the shape of real letters during the shape task in the early visual ROI (Fig. 7) clarifies that it is indeed possible to decode representations of maintained familiar stimuli in the early visual cortex, and again conforms to the 'sketchpad' view.

4.3 Training on perceptual representations

In addition to training on imagery, we also trained classifiers solely on the perception of orientation/shape during the localizer session. For this we looked at the temporal evolution across scans and analyzed the information content at the start of the imagery period and during perception of the probe. Orientation decoding during imagery in the early visual ROI showed a similar profile to orientation decoding from imagery training (Fig. 8 and Fig. 6 respectively). However, decoding orientation of real letters was no longer possible in the letter-relevant ROI when training on perception. One might assume a trade-off with real letter identification during this task, whereby letter shape is more easily represented, however, there were also no imagery period representations found for letter shape during the shape task in this ROI. Assessing the shape decoding results during perception of the probe, we found high classification accuracies for real and pseudo letters during both tasks in both ROIs, and these were significantly higher than corresponding orientation decoding accuracies in these regions (Figs. 10, 11 and Figs. 8, 9 respectively). This suggests that the perception of letter shape strongly dominated over the perception of grating orientation during the localizer session. This finding reflects the widely accepted notion of seeing the 'forest before the trees' (Navon, 1977). Assuming this, it is likely that this is the reason why imagery period orientation information of real letters is lost from the letter-relevant ROI. However, pseudo letter orientation information is not lost during imagery; again we could decode pseudo letter grating orientation during the shape task, the same as when we trained on imagery. Why is real letter orientation information lost during the orientation task in this ROI, when pseudo letter information is not lost during the shape task, as compared to when training on imagery? This again may be explained by the prediction that it is more difficult to imagine the shape of unfamiliar stimuli. The reliable representations during the perception of pseudo letter shape do not change the fact that

people cannot maintain a good image of unfamiliar stimuli shape and thus, they ultimately focus on the finer detail. In contrast, strong perceptual representations of real letter shape may hamper or interfere with one's ability to imagine the finer detail when required.

The aforementioned conclusion would suggest that we can expect better decoding of real letter shape during the imagery period in the letter-relevant ROI; this was found not to be the case. However, we have two reasons to believe that representations of letter stimuli during perception are not the same as working memory representations. Firstly, when training on imagery, we can decode real letter shape for the shape task during the imagery period in the early visual ROI (Fig. 7). In contrast, when training on perception, we cannot decode this shape information (Fig. 10). If this were due to a problem with analyzing the localizer perception activity, we would expect no decoding to be possible at all, however, this is clearly not the case, since shape information during subsequent perception of the probe can easily be classified. Secondly, when training on perception, shape classification results from both ROIs distinctly demonstrate no shared imagery period representations. All shape classification conditions showed high decoding accuracies during the perception of the probe. It is likely that univariate analysis of the task session would reveal differential BOLD responses between the imagery period and perception of the probe (Albers et al., 2013), suggesting that lower univariate activity during the imagery period might account for the decreased decoding performance. However, previous research has shown that decoding is possible even when overall activity levels are low (Christophel, Hebart, & Haynes, 2012; Riggall & Postle, 2012; Serences, Ester, Vogel, & Awh, 2009). Since orientation information can still be deciphered during imagery from perceptual training data, we therefore suggest that working memory representations of the shape of letter stimuli, specifically for real letters, are different to perceptual representations in the visual cortex.

4.4 Future directions

One of the main drawbacks of this study was that our letter-relevant (higher-level) ROI was not defined appropriately. Stimuli used during the localizer session were the same as the task stimuli, displaying each of the four letters balanced across the two possible grating orientations. Since we were interested in both the finer and more global features, an issue that arose was that these features were integrated in the localizer, making it difficult to separate out the activity associated purely with the perception of grating orientation and purely with the perception of letters. As mentioned earlier, we found that perceptual representations of shape strongly dominated over the perceptual representations of grating orientation. Ideally, we would use two localizers; one displaying a large annulus grating, and the other presenting letters with no "finer detail", i.e. coloured uniformly in grey. This would result in better-defined ROIs for early visual and letter-relevant cortex. When we tried to create a real letter over pseudo letter contrast using our current localizer, we could not find a sufficient number of significantly active voxles in most participants, i.e. the 150 most active voxels required for decoding analyses. Thus, instead we used an exclusive mask of pRF checkerboard activity over the perception of all stimuli to create our letter-relevant ROI. These pRF checkerboards were carried out in a separate run to the localizer. A better method of locating the letter-relevant ROI (or VWFA) would be to make a real letter over checkerboard activation contrast from a single session as carried out by Cohen and colleagues (2002), rather than creating an exclusive mask.

An additional future analysis for a similar experimental paradigm would be to analyse the data collected by the eye-tracker, to ensure participants maintain fixation at all times. However, this was not deemed essential for this study since the classifier was trained on an independent localizer, during which participants had to perform a difficult task at fixation. In addition, the stimuli presented during the task session were presented only very briefly (200ms), making it difficult to prepare and make eye movements towards them.

As mentioned in the methods section, the pRF mapping did not cover the entire letter stimuli in the vertical dimension. This suggests that perhaps some of the letter-relevant ROI decoding is from voxels associated with the bottom corners of the letter stimuli, i.e. spatial locations usually associated with V1. The poor letter decoding results during imagery in this ROI might be an indication of this. However, the very significant classification when training on perception during the localizer and testing on the perception of the probe shows that letter shape can be decoded in this region, and is unlikely to be as a result of the perception of just the bottom extremities of the stimuli, since participants performed a difficult task at fixation for the duration of the localizer session.

Overall, it is likely that our early visual ROI includes voxels associated with grating orientation but also with the shape of the letter stimuli, since we used an inclusive mask to map where the spatial positions of the letter stimuli overlapped with the checkerboards. In this way, it is harder to conclusively say whether participants had "conceptual" representations of real letter stimuli or rather, representations of the global feature of letter shape.

5. Conclusion

From this MVPA study, we conclude that familiarity with a stimulus influences the classification of representations during imagery in the visual cortex. This effect should therefore be taken into account in any future studies using more complex stimuli in targeting the similarities and differences between visual mental imagery, working memory and perception. Our imagery results also corroborate the reverse hierarchy theory of perceptual learning, showing that representations of low-level features are no longer required when a higher-level representation has already been established. Lastly, our results also indicate that it is difficult to decode global representations of unfamiliar stimuli, while simultaneously showing better classification of the finer detail of these stimuli.

6. References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457–464.

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*(15), 1427-1431.

Binder, J.R., Medler, D.A., Westbury, C.F., Liebenthala, E., Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage, 33*(2), 739–748.

Bolger, D.J., Perfetti, C.A., Schneider, W. (2005). Cross-cultural effect on the brain revisited: universal structures plus writing system variation. *Human Brain Mapping, 25*, 92—104.

Braet, W., Wagemans, J., & Op de Beeck, H. P. (2012). The visual word form area is organized according to orthography. *NeuroImage, 59*(3), 2751–2759.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Christophel, T.B., Hebart, M.N., & Haynes, J.D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *Journal of Neuroscience, 32*, 12983–12989.

Cohen, L, Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A, & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*(Pt 2), 291–307.

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain, 125*(Pt 5), 1054–1069.

Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage, 39*(2), 647–660.

Flowers, D. L., Jones, K., Noble, K., VanMeter, J., Zeffiro, T. A, Wood, F. B., & Eden, G. F. (2004). Attention to single letters activates left extrastriate cortex. *NeuroImage, 21*(3), 829–839.

García-Pérez, M. A. (1998). Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. *Vision research, 38*(12), 1861–1881.

Harrison, S. A, & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632–635.

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience, 8*(5), 686–691.

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7*(7), 523–534.

Hughes, H. C., Fendrich, R., & Reuter-Lorenz, P. A. (1990). Global Versus Local Processing in the Absence of Low Spatial Frequencies. *Journal of Cognitive Neuroscience, 2*(3), 272–282.

James, K. H., James, T. W., Jobard, G., Wong, A. C. N., & Gauthier, I. (2005). Letter processing in the visual system: different activation patterns for single letters and strings. *Cognitive, Affective & Behavioral neuroscience, 5*(4), 452–466.

Jehee, J. F. M., Brady, D. K., & Tong, F. (2011). Attention improves encoding of task-relevant features in the human visual cortex. *The Journal of Neuroscience, 31*(22), 8210–9(?).

Kaas, A., Weigelt, S., Roebroeck, A., Kohler, A., & Muckli, L. (2010). Imagery of a moving object: the role of occipital cortex and human MT/V5+. *NeuroImage, 49*(1), 794–804.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8*(5), 679–685.

Klein, I., Paradis, a L., Poline, J. B., Kosslyn, S. M., & Le Bihan, D. (2000). Transient activity in the human calcarine cortex during visual-mental imagery: an event-related fMRI study. *Journal of Cognitive Neuroscience, 12* (Suppl 2), 15–23.

Knauff, M., Kassubek, J., Mulack, T., & Greenlee, M. W. (2000). Cortical activation evoked by visual mental imagery as measured by fMRI. *Neuroreport, 11*(18), 3957–3962.

Kok, P., Jehee, J. F. M., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron, 75*(2), 265–270.

Kosslyn, S.M., Ganis, G., & Thompson, W.L. (2003). Mental imagery: against the nihilistic hypothesis. *Trends in Cognitive Sciences (Regular Edition), 7*, 109-111.

KrishnaVeni, C.V., Sobha Rani, T. (2011). On the Classification of Imbalanced Datasets. *International Journal of Computer Science & Technology, 8491*, 145–148.

Likova, L. T. (2012). Drawing enhances cross-modal memory plasticity in the human brain: a case study in a totally blind adult. *Frontiers in human neuroscience, 6*.

Marzi, C. A, Mancini, F., Metitieri, T., & Savazzi, S. (2006). Retinal eccentricity effects on reaction time to imagined stimuli. *Neuropsychologia, 44*(8), 1489–1495.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353–383.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stiimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12*(6), 1013–1023.

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Podgorny, P., & Shephard, R.N. (1978). Functional Representations Common to Visual Perception and Imagination. *Journal of Experimental Psychology: Human Perception & Performance, 4*(1), 21-35.

Pylyshyn, Z. W. (2002). Mental imagery: in search of a theory. *The Behavioral and Brain Sciences, 25*(2), 157–82; discussion 182–237.

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: decoding category information during mental imagery. *NeuroImage, 50*(2), 818–25.

Riggall, A.C., & Postle, B.R. (2012). The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *Journal of Neuroscience, 32*, 12990–12998.

Serences, J.T., Ester, E.F., Vogel, E.K., & Awh, E. (2009). Stimulus- specific delay activity in human primary visual cortex. *Psychological Science, 20*, 207–214.

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex, 15*(10), 1570–1583.

Spillmann, L. (1999). From elements to perception: Local and global processing in visual neurons. *Perception, 28*(12), 1461–1492.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *NeuroImage, 56*(1), 330–344.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron, 55*(1), 143–156.

Exploring the Automaticity of Language-Perception Interactions

Erik L. Meijs[1]
Supervisors: Jolien C. Francken[1], Simon van Gaal[1,2], Floris P. de Lange[1]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2]Department of Psychology, University of Amsterdam, The Netherlands

Language has been shown to significantly alter how we perceive the world. This paper aims at finding out to what extent language and perception interact automatically. In our first experiment, we asked how automatically language changes perception in real-time. We showed participants visual motion stimuli, preceded by motion verbs that were congruent or incongruent with respect to this stimulus. In half of the trials, we backward masked the motion verbs to render them unaware. Language had a substantial effect on both reaction times and discrimination ability when participants were aware of the motion verbs. When participants were unaware of the verbs, language still changed perception, but only for stimuli presented in the right visual field. Possibly, the signal from unaware verbs is not strong enough to cause spreading activation from the left, language-dominant, hemisphere to the right hemisphere. Our second experiment used a 1-back task with motion and neutral verbs in the functional magnetic resonance imaging (fMRI) scanner. Within an exploratory set-up, we aimed at finding out whether this task, which is not specifically semantic in nature, activates sensory areas. We demonstrate that this is indeed the case: occipital areas (V3) react more strongly to motion verbs than to neutral verbs. Moreover, the semantic category to which the verb belonged could be decoded with above-chance accuracy from the voxels that show visual motion sensitivity, indicating that representations of visual motion and semantic motion may overlap. We tentatively conclude that language and perception are strongly related and that they interact automatically.

Keywords: embodied cognition, motion perception, awareness, semantics, priming, fMRI, MVPA

Corresponding author: Erik L. Meijs; E-mail: erik.meijs@donders.ru.nl

1. Introduction

Top-down factors (e.g. expectations, goals, world knowledge) have a big influence on how we perceive the world around us (Bar, 2009; Bar et al., 2006). A growing body of studies shows that language is one of these factors. Language, and in particular semantics, has been shown to interact with perceptual processes in many domains, for example: face perception (Landau, Aziz-Zadeh, & Ivry, 2010), auditory processing (Chen & Spence, 2011), color perception (Gilbert, Regier, Kay, & Ivry, 2006; Witzel, & Gegenfurtner, 2011) and motion perception (Meteyard, Bahrami, & Vigliocco, 2007; Meteyard, Zokaei, Bahrami, & Vigliocco, 2008). In this paper, we will focus on the latter, looking at the interaction between visual motion perception and motion language processing.

In 2007, Meteyard et al. demonstrated that detection of a threshold level visual motion stimulus was altered by the simultaneous auditory presentation of motion words over headphones. Congruency of visual motion and motion words affected the participant's internal decision criterion and their speed. In another study, Meteyard and colleagues inverted their paradigm and showed that motion word processing was changed by a concurrent salient visual motion stimulus (Meteyard et al., 2008).

Other studies have taken a different approach to investigate at the relationship between language and perception, looking at the involvement of sensory areas of the brain in semantic processing. Some have indicated that hMT+/V5, a visual area that is normally implicated in motion processing, is activated during the processing of motion semantics (Saygin, McCullough, Alac, & Emmorey, 2010). For example, one study showed that this area was significantly more active in response to motion sentences compared to other sentence types (Saygin et al., 2010). Interestingly, another study showed comparable effects for deaf people. Here, sign language with a motion content changed activation in hMT+/V5, independent of the movement caused by the signs themselves (McCullough, Saygin, Korpics, & Emmorey, 2012). It must be noted that there is debate regarding the exact locus of activations in such paradigms. Multiple studies find activation in areas adjacent to, but not overlapping with visual area hMT+/V5 (Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008; Desai, Binder, Conant, & Seidenberg, 2010; Humphreys, Newling, Jennings, & Gennari, 2013; Kable, Kan, Wilson, Thompson-Schill, & Chatterjee, 2005) and the role of sensory-motor cortices in semantic processing is still unclear (Hauk & Tschentscher, 2013). To summarize, there are strong indications that motion language and motion perception interact, possibly because there is overlap in the responsible brain regions. There is, however, no consensus about the nature of these interactions or their neural underpinnings (Klemfuss, Prinzmetal, & Ivry, 2012; Mahon & Caramazza, 2008; Willems & Casasanto, 2011).

Recent studies have brought new insights into the debate about language-perception interactions, demonstrating that such interactions may be more automatic than was previously thought. In these studies, detection of visual motion was altered (response times and internal decision criteria) by reading (Francken, Kok, Hagoort, & de Lange, n.d.) or hearing (Meteyard et al., 2007) motion words, notwithstanding participants were explicitly instructed that the words were irrelevant to the task and hence could be ignored. Interestingly, in the Francken et al. (n.d.) study linguistic interference on visual perception was only existent when the visual stimulus was presented in the right visual field. It is conceivable that because visual stimuli in the right visual field are processed first in the left – language dominant (Knecht et al., 2000) – hemisphere, linguistic influences are easier installed. Further evidence that suggests automatic interactions between language and perception comes from electrophysiological studies that show linguistic modulation of sensory-motor cortices at early latencies (Hauk & Pulvermüller, 2004), leaving almost no time for controlled processes to be causally involved.

Together with the behavioral data reported above, these results lead to the overall idea that the interaction between language and perception, or the recruitment of perceptual areas by semantic processes, happens in a relatively automatic fashion. Nevertheless, there are studies suggesting the opposite. For instance, a study by Waechter and colleagues (Waechter, Besner, Jennifer, & Stolz, 2011) indicated the necessity of attention for inter-modular interactions (but see also Relander, Rämä, & Kujala (2009)). Other studies have made similar claims, arguing that stimulus awareness is required (Batterink, Karns, Yamada, & Neville, 2010; Kang, Blake, & Woodman, 2011).

Because the current evidence appears to be conflicting, in this paper we will directly investigate the automaticity of language-perception interactions. Specifically, we will do this for motion language and visual motion stimuli, in two separate experiments

with entirely different approaches. First, we will ask to what extent language influences perception automatically, by looking whether this influence is dependent on awareness. Second, we aim to find out how automatic the recruitment of sensory-motor areas in a non-semantic (1-back) task with motion words is. To our knowledge, our study will be the first to explicitly study the automaticity of language-perception interactions. Therefore, we expect to add to the literature by giving new insights into the nature of the relationship between language and perception.

In the first experiment, we want to find out how automatic the influence of language on perception is. For that reason, we have adopted the paradigm by Francken et al. (n.d.) and changed it to match our research question. In this paradigm, a visual motion stimulus is preceded by congruent or incongruent motion verbs. Critically, to investigate the automaticity of the interaction we manipulate participants' awareness of the motion verbs, by adding backward masks after fifty percent of the verbs. For those motion verbs that are rendered unconscious by masking, we can ensure that processing occurs "automatically," without contributions of intentional, controlled processes.

The power of the unconscious has been shown by many (de Lange, van Gaal, Lamme, & Dehaene, 2011; Kiefer & Spitzer, 2000; Lamme & Roelfsema, 2000; van Gaal & Lamme, 2012). Nevertheless, we will make a few additional changes to the existing paradigm to augment our experimental power as much as possible, since the effect sizes for unconscious stimuli are often quite small (van Gaal, de Lange, & Cohen, 2012). First, in our study participants will attend to the words, which has been shown to be a necessary condition for the observation of unaware semantic effects (Cohen, Cavanagh, Chun, & Nakayama, 2012; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Kiefer, Adams, & Zovko, 2012; Naccache, Blandin, & Dehaene, 2002; Spruyt, De Houwer, Everaert, & Hermans, 2012; van den Bussche, Smets, Sasanguie, & Reynvoet, 2012). Second, we will use only the 20% words with the strongest motion association (pretested) from the existing study (Francken et al., n.d.), to make the (in)congruency between words and visual stimulation as extreme as possible. We expect that these changes will cause the congruency effect sizes for the aware condition to be larger than those in previous studies, where the stimuli were unattended. For the unaware condition, small effects are expected. Based on previously reported results, we also anticipate that our effects could be present only, or be bigger, for motion stimuli presented in the right visual field (Francken et al., n.d.; Gilbert et al., 2006). If language indeed influences perception in the absence of awareness, this would point at an automatic process. Consequently, higher-order explanations of language-perception interactions, such as imagery, would become rather unlikely.

Our second study is an exploratory functional magnetic resonance imaging (fMRI) study, in which we look to what extent a 1-back task activates (early) visual cortices that are also implicated in the processing of visual motion. This will allow us to make claims about the automaticity with which sensory areas are implicated in semantic processing, since the 1-back task is not necessarily semantic in nature. Accordingly, if we are able to find differential activation in visual areas for motion words compared to neutral words, we can assume that this originates from automatic activation of visual areas in response to the semantic content of these words. Crucially, the task will contain only verbs, to exclude possible word type confounds (Bedny et al., 2008). As in previous studies, we expect to find involvement of sensory areas such as hMT+/V5 (McCullough et al., 2012; Saygin et al., 2010), or neighboring regions such as posterior temporal cortex (Desai et al., 2010; Humphreys et al., 2013; Kable et al., 2005) during semantic processing.

Apart from looking at activation levels in visual cortex, we will go a step further and use decoding techniques in visual areas in order to find out more about the specific information that is available in these brain regions. Decoding analyses are capable of detecting differences on a much finer scale than univariate approaches, which can only capture fairly global activation differences. In our experiment, specifically, we will try to decode the semantic category of the verbs (upward, downward, neutral) we use in the 1-back task from voxels in the visual cortex that are also involved in visual motion processing.

We believe that this is the first study that attempts to decode specific semantic features from visual cortex. In the end, this will allow us to get a much more specific view on the semantic information (if any) that is available in visual areas. Based on other studies, we know that decoding of semantic categories is possible (Simanova, Hagoort, Oostenveld, & van Gerven, 2012). It is a matter of exploration, however, whether this is going to be possible in visual areas as well. Therefore, as said before, this study can be seen as an extensive exploratory study, aiming to find out more about the representation of semantic information in sensory

areas.

In sum, we will investigate the automaticity of language-perception interactions in two different studies. To foreshadow our conclusion, we will show that these interactions are relatively automatically and may not always require awareness. Furthermore, we demonstrate that motion words automatically and specifically activate visual areas, allowing for decoding of semantic features from occipital cortex.

2. Experiment 1: The automaticity of language-perception interactions

In experiment 1, we aim at finding out whether the direct, real-time, influence of language on perception is dependent on awareness of a linguistic stimulus or whether it is an automatic process. Therefore, we adapted the methods used by Franken et al. (n.d.). In this task, participants respond to visual motion after seeing word primes that are congruent, neutral or incongruent with respect to this motion. Critically, we manipulated awareness of the words by backward masking the verbs on half of the trials. Furthermore, in contrast to the existing study, participants attended the word primes to enhance effects on perception (Dehaene et al., 2006). We also changed the task participants performed from motion detection to a motion discrimination, so we can compare upward and downward motion within-subject. The goal of this experiment was to (1) replicate previous findings (Francken et al., n.d.) with attention to primes and (2) find out whether language-perception interactions happen automatically, that is, in the absence of prime awareness.

2.1 Methods

2.1.1 Participants

Twenty-eight healthy, right-handed participants (22 female, age 22.0 ± 3.1 years) with normal or corrected-to-normal vision took part in the experiment. All participants were native Dutch speakers and reported having no reading problems. One participant was excluded due to failure to comply with task instructions. The study was approved by the local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands) and all participants gave written informed consent according to the declaration of Helsinki. Compensation was 16 Euros, or an equivalent in course credit.

2.1.2 Stimuli

Stimuli were generated using the Psychophysics Toolbox (Brainard, 1997) for MATLAB (MathWorks, Natick, MA, US), and presented on a Samsung SyncMaster 940BF monitor (refresh rate: 60 Hz, resolution: 1024x768). We used a chin and forehead rest to ensure a constant viewing position within and between participants. Stimuli were presented in white (220 cd/m2) on a gray background (38 cd/m2).

Visual motion stimuli: The visual motion stimulus was a random-dot motion pattern consisting of white dots (speed: 13.8 deg/s, density: 2 dots/deg2) presented in an annulus aperture (radius: 5.5 deg) at the lower left or lower right quadrant of the screen (annulus center at 6.85 deg eccentricity vertically and horizontally). The stimulus lasted 12 presentation frames (200 ms). In the first frame, a random configuration of dots was presented within the annulus. Subsequently, in each frame the presentation was updated by replotting the dots: a certain percentage (see Procedure) of the dots was replotted consistently in one direction (upward or downward) to induce the motion percept. Dots moving outside of the annulus and other remaining dots were replotted at random location within the annulus.

Words and masks: There were three word categories: upward verbs (e.g. 'rise'), downward verbs (e.g. 'sink') and non-motion verbs (e.g. 'rest'). For each category, we picked five words based on their pre-tested association with the respective semantic category (Supplementary Table 1). We assured that the words of all categories were matched (all p's > .05) on lexical frequency (based on the CELEX database; celex.mpi.nl), word length (5-8 letters), number of syllables and word concreteness (using ratings obtained before the start of the experiment). A separate word list was used for the training part of the experiment (Supplementary Table 2). Masks were randomly generated combinations of consonants. They were 10 letters long so that they would optimally mask all words independent of word length. Both words and masks were presented at the center of the screen, using capital letters in a mono-spaced font ('Lucida typewriter').

2.1.3 Procedure

Participants performed a motion discrimination (upward vs. downward) task on a random-dot motion stimulus (200 ms) presented in either the left

visual field (LVF) or right visual field (RVF). They were instructed to respond as quickly and accurately as possible by pressing a button with the index finger or middle finger of their right hand. The response mapping was counterbalanced across participants. We instructed participants to keep fixating on the center of the screen throughout the trial. Task difficulty was matched across participants by thresholding the percentage of coherently moving dots (see below).

The visual motion stimulus was preceded by a word (33 ms), which in turn was preceded by a forward mask (50 ms). The semantic content of the word was upward, downward or not motion-related. Consequently, word meaning was congruent, neutral or incongruent with respect to the visual motion stimulus. Awareness of the word was manipulated by having either backward masks (2 x 33 ms; unaware condition) or a blank screen (67 ms; aware condition) after the word. A short inter-stimulus interval (ISI) of 17 ms was always presented after this (just before the motion stimulus). The complete sequence of events in a trial is illustrated in Figure 1.

In 10% of the trials the motion discrimination task was followed by an additional catch task. Here, participants indicated whether the word presented earlier in the trial was a motion verb (upward or downward category) or a non-motion verb. Catch trials were included for two reasons. First, they assured attention to the motion verbs, which is supposed to enhance the processing of the word in both the aware and unaware conditions (Naccache et al., 2002; Nestor, Behrmann, & Plaut, 2013; Norman, Heywood, & Kentridge, 2013; Spruyt et al., 2012). Second, catch trials could be used to estimate word awareness. For unaware trials, participants should not be able to perform above chance-level on the catch-task. Participants were instructed to always respond to the catch question and guess when they thought they did not know the answer.

The experiment consisted of two one-hour sessions on separate days within one week. We preferred to have two separate testing days because the study duration was too long to reliably test participants on a single day. A long study duration was essential to attain enough statistical power with our number of conditions and to allow sufficient time for instructions and training (see below). During each session there were 6 blocks of 105 trials (total of 1260 trials). After each block, participants got feedback about their performance on each of the tasks, followed by the opportunity to take a short break before going on to the next block. The inter-trial interval (ITI) was 1000-1166 ms.

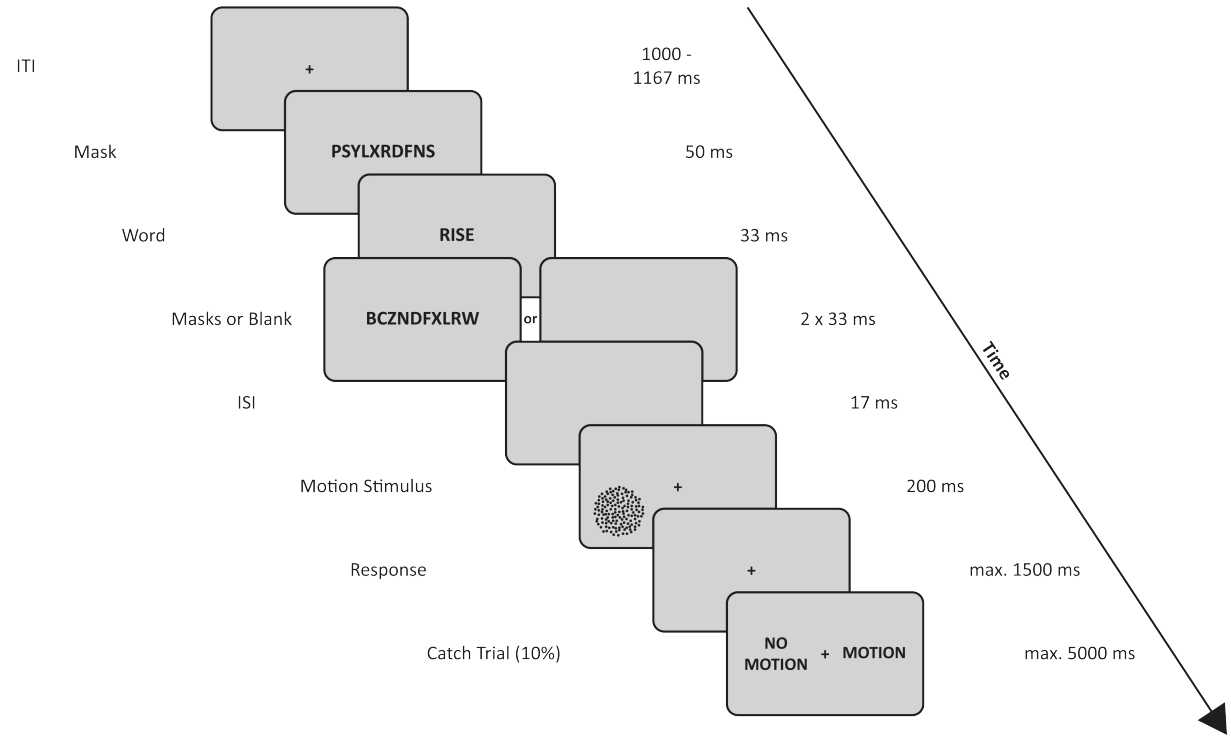


Fig. 1 Task design. Participants responded to a random-dot motion stimulus. This stimulus was preceded by a motion (upward, downward) or non-motion word. On half of the trials, this motion verb was backward masked in order to make participants unaware of its semantic content. In 10% of the trials, we included an additional catch task that assured attention to the verbs and that was used to get an estimate of word awareness.

A training (session 1: 120 trials; session 2: 90 trials) and thresholding (60 trials) phase took place before the start of the experiment at each session. The training phase was included to guarantee that participants understood the task instructions. During the training phase, participants were provided with trial-by-trial feedback for both the motion task and the catch trial task. To exclude training effects, we used a different word list in the training blocks. The first three training blocks in the first session had a fixed percentage of coherently moving dots (80%, 55% and 30% respectively). For the final two training blocks this percentage was adjusted based on the participant's performance on the previous block. The final coherence estimate was used as a starting point for the thresholding block. Here, we employed a Bayesian staircase procedure (Watson & Pelli, 1983) that thresholded the percentage of dots moving coherently in one direction such that task difficulty was matched (75% accuracy on the motion discrimination task) across participants. During the remainder of the experiment, the percentage of dots moving coherently was adjusted after each block to keep accuracy at a stable level. The end-point of the staircase procedure of the first session was taken as a starting point estimate for the training blocks in the second session.

At the end both sessions, participants performed a short posttest (75 trials), in which they performed the catch task only, with the goal of having an additional estimate of word awareness.

2.1.4 Behavioral analysis

Motion task: In our analyses, we used reaction times (RTs) and signal detection variables (MacMillan, 2002; Rahnev, Lau, & de Lange, 2011) to qualify both discrimination ability (d') and response bias (c). d' is a measure of a participant's discrimination ability, independent of any potential biases. A higher d' relates to a participant being better at discriminating upward from downward motion. c is an estimate of the internal response criterion a participant has. A negative criterion is related to a tendency to give an 'upward motion' response, while a positive criterion is related to the opposite tendency. Both measures are computed based on a ratio of hit and false alarm rates.

RTs were calculated based only on trials with a correct response, to rule out speed accuracy trade-offs (additional analyses showed that using all trials did not change the overall pattern of results). Furthermore, we excluded trials (1.6% of total) with an RT deviating more than three standard deviations (SD) from a participant's mean RT. For the remaining trials, we looked at the congruency (incongruent – congruent) effects. Congruent means that the visual motion direction matched the semantic category of the word that was presented. When visual motion and word meaning did not match, a trial was called incongruent. A repeated measures ANOVA with the factors congruency (congruent, neutral, incongruent), visual field (LVF, RVF) and awareness (aware, unaware) was performed for each of the dependent measures. Subsequently, we performed paired t-tests to directly compare the congruent en incongruent conditions separately for every visual field by awareness combination. Neutral trials were not included in this part of the analysis, since we did not have a specific hypothesis about them.

Catch trials and posttest: For the catch trials and posttest we looked at percentage correct and d'. Percentage correct was defined as the percentage of trials on which participants were able to indicate whether the word was a motion verb or not. d' was a measure of the unbiased discriminability of the words. Both measures were calculated and analyzed independently for catch trials and posttest trials. Trials with an RT deviating more than three SD's from a participant's mean (1.1% of catch trials, 1.5% of posttest trials) were excluded. The percentage correct and d' averages for the aware and unaware conditions were compared using paired t-tests. Subsequently, we used the percentages correct in binomial tests to determine on a single-participant basis whether performance on the catch task was better than chance (50% correct).

2.2 Results

2.2.1 Word awareness

We made two assumptions related to participants' awareness of the semantic content (motion, no motion) of the word that was presented in a trial. First, and most importantly, we assumed there would be a visibility difference between the unaware and aware conditions. We tested this premise on both the catch trials and the posttest data (see also Fig. 2). Paired t-tests indicated that indeed participants' catch task performance was better in the aware condition (catch trials: M = 86%, SE = 2.10; posttest: M = 92%, SE = 1.02) than in the unaware condition (catch trials: M = 55%, SE = 1.58; posttest: M = 58%, SE = 2.01) for both the catch trials ($t(26) = 14.16$, $p < .001$) and the posttest ($t(26) = 15.74$, $p < .001$). In catch trials, the average d' was higher (M = 2.18, SE = 0.17) for the aware condition than (M

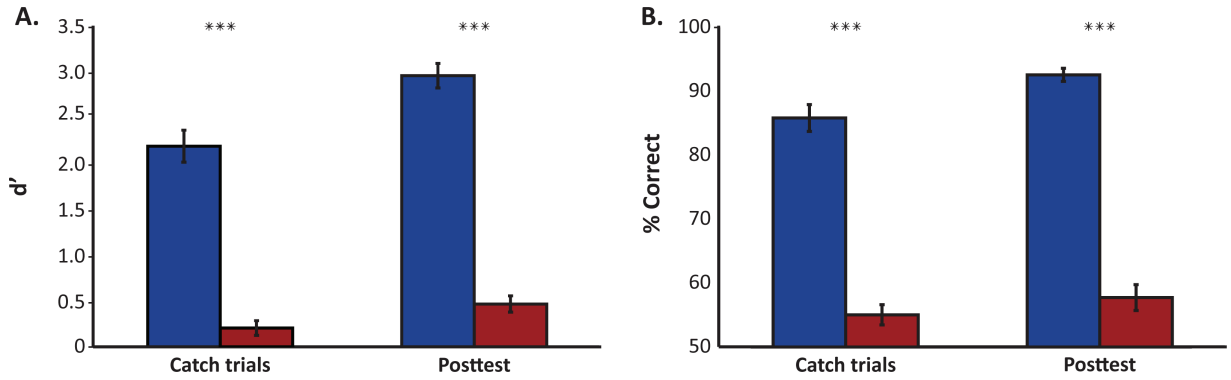


Fig. 2 Awareness levels. Average awareness levels of aware (blue) and unaware (red) conditions are given for **A.** d' and **B.** percentage correct in catch trials and the posttest (* = p < .05, ** = p < .01, *** = p < .001). Errorbars denote SE.

= 0.20, SE = .08) the unaware condition (t(26) = 12.83, p < .001). During the posttest, d' was also higher for aware (M = 2.95, SE = 0.13) than unaware (M = 0.45, SE = 0.09) trials (t(26) = 17.30, p < .001).

Secondly, we assumed that in the unaware condition participants' performance on the catch task is not above chance, since the masking should make it impossible for them to adequately judge the semantic category of a verb. Therefore, we analyzed the percentage correct on the catch trials and tested this against chance level performance. Binomial tests showed that 7 out of 27 participants were significantly (p < .05) better than chance on the unaware catch trials. Thus, for those participants the masking procedure was probably sufficiently strong to reduce visibility and make explicit semantic judgments about the words impossible. Whereas this may seem problematic, we managed to show using a number of control analyses (see Motion discrimination task results) that our behavioral results were most likely not compromised by this violation of assumption.

2.2.2 Motion discrimination task: RTs

To test whether motion language and visual motion perception interact on a behavioral level in the different conditions, we performed a repeated measures ANOVA with the factors congruency (congruent, neutral, incongruent), visual field (LVF, RVF) and awareness (aware, unaware). A significant awareness effect (F(1, 26) = 9.82, p < .01, η 2 = .274) was found, indicating people were on average faster on unaware trials (M = 480.20 ms, SE = 19.16) than on aware trials (M = 496.63 ms, SE = 19.98). There was also a significant effect of congruency (F(2, 25) = 30.38, p < .001, η2 = .708) showing that people were on average fastest in the congruent condition (M = 470.49 ms, SE = 19.03) and slowest in the incongruent condition (M = 508.91 ms, SE = 20.23), with the neutral condition lying in between (M = 485.84 ms, SE = 19.30). Then, a significant awareness x congruency interaction was found (F(2, 25) = 25.07, p < .001, η2 = .667), resulting from a bigger overall congruency effect in the aware condition (M = 64.25 ms, SE = 7.34) than in the unaware condition (M = 12.59 ms, SE = 4.67).

Planned paired t-tests gave a more detailed insight into how the congruency effects depend on awareness and visual field. Participants were significantly faster in congruent than incongruent trials for aware trials (LVF: t(26) = -7.82, p < .001; RVF: t(26) = -7.82, p < .001). Interestingly, for the unaware trials such a congruency effect was present only for motion stimuli presented in the RVF (t(26) = -2.94, p < .01) but not in the LVF (p > .5).

To control for the fact that some of the participants had higher-than-expected word category awareness for the unaware trials (see section on Word awareness), we performed two control analyses to check whether they were causing the unaware RVF effects we find here. First, we repeated the paired t-tests for the unaware conditions without these participants. This led to the same overall pattern of results as the total group results, with a significant unaware RVF (t(19) = -2.88, p < .01), but not LVF (p > .3), congruency effect.

Second, we performed a median split (median = 54% correct for unaware catch trials) on participants' word category awareness on unaware trials as measured by percentage correct. We then defined two groups of participants based on whether their awareness was above or below the median and performed a repeated measures ANOVA in each of the groups to look at the unaware RT effects in both visual fields. The only significant effect in these analyses was a significant visual field x congruency interaction (F(1, 13) = 6.68, p < .05) in the group

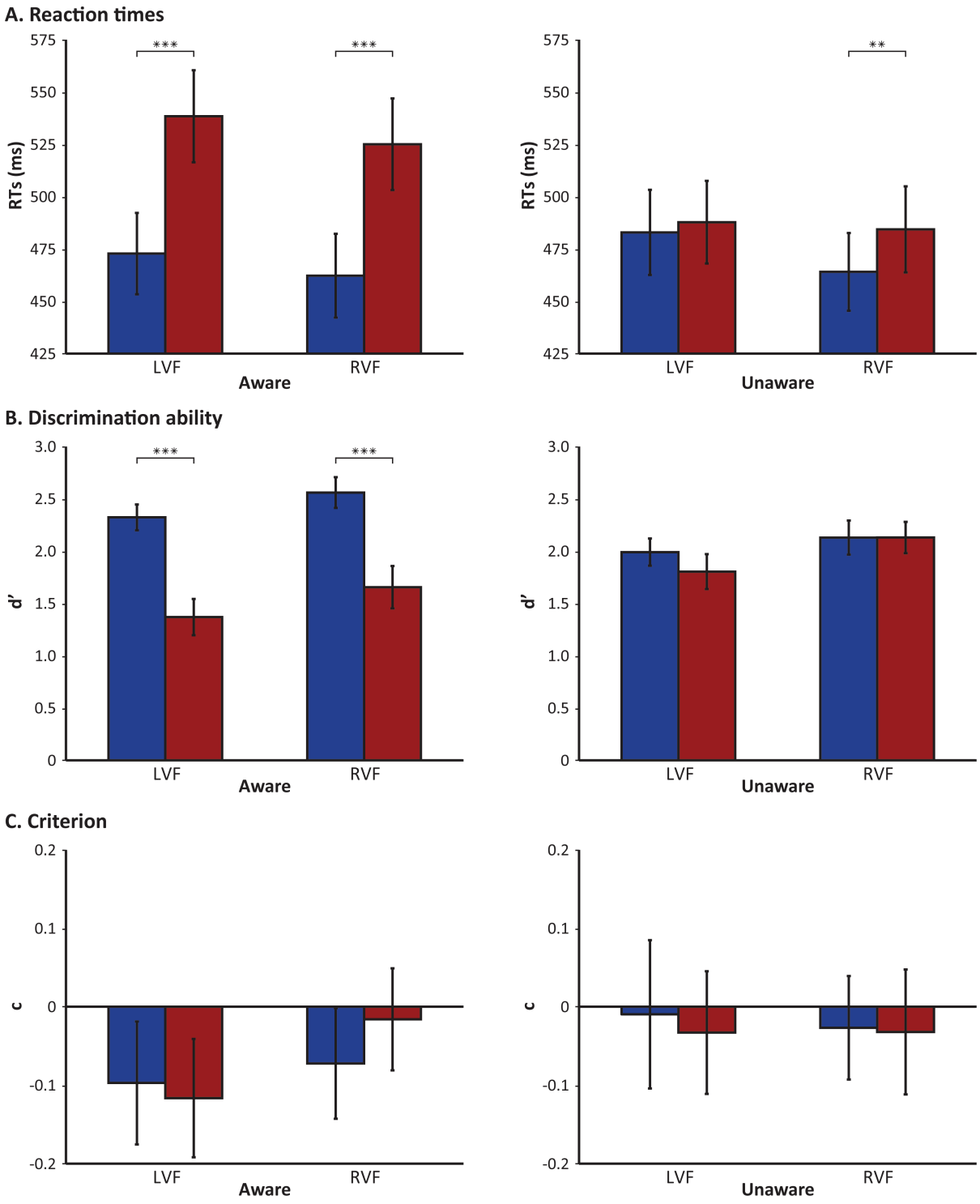


Fig. 3 Behavioral results of experiment 1. For each of the dependent measures, **A.** RTs, **B.** d' and **C.** c for the congruent (blue) and incongruent (red) conditions are given for all combinations of awareness (aware, unaware) x visual field (LVF, RVF). There were significant congruency effects in RTs and d' for both visual fields in the aware condition. Additionally, there was a significant RT congruency effect for the unaware condition, but only in the RVF (* = $p < .05$, ** = $p < .01$, *** = $p < .001$). Errorbars denote SE.

with low word awareness. Paired t-tests in this group showed a significant congruency effect in the RVF ($t(13) = -3.84$, $p < .01$) but not in the opposite visual field ($p > .4$). There were no significant effects at all in the high word awareness group (all $p > .05$). While these effects need to be interpreted with caution because of the small sample sizes after the median split, they do not support the idea that the unaware effects were caused by participants having too high word awareness.

2.2.3 Motion discrimination task: d' and c

As for the RTs, we performed repeated measures ANOVA's on the signal detection measures d' and c to look at the effects on motion discrimination ability and perceptual bias respectively. Discrimination ability was significantly changed by word-motion congruency ($F(2, 25) = 15.82$, $p < .001$, $\eta 2 = .56$). Performance was best in the congruent condition ($M = 2.25$, $SE = 0.11$), somewhat worse in the neutral condition ($M = 2.01$, $SE = 0.12$) and worst in the incongruent condition ($M = 1.74$, $SE = 0.15$). This congruency effects was dependent on awareness (congruency x awareness interaction $F(2, 25) = 35.75$, $p < .001$, $\eta 2 = .74$), resulting from a bigger overall congruency effect for aware trials ($M = -0.93$, $SE = 0.15$) than in unaware trials ($M = -0.09$, $SE = 0.07$). Paired t-tests indicated there were significant congruency effects only in the aware condition, regardless of the visual field in which the motion was presented (LVF: $t(26) = 6.06$, $p < .001$; RVF : $t(26) = 5.70$, $p < .001$).

For criterion, we observed slightly different results in the repeated measures ANOVA. The only significant effect ($F(1, 26) = 5.42$, $p < .05$, $\eta 2 = .17$) was the visual field x awareness interaction, caused by a numerically bigger awareness effect in the LVF ($M = 0.11$, $SE = 0.06$) than in the RVF ($M = 0.01$, $SE = 0.05$). Again, we used t-tests to shed light on the congruency effects under all circumstances. No significant congruency effects were found for any of the comparisons.

2.3 Discussion

In this first experiment, we wanted to find out to what extent linguistic primes change visual perception automatically. Therefore, in an existing paradigm (Francken et al., n.d.) we rendered part of the words unaware, to rule out controlled or intentional processes. For the consciously perceived words, we show that language and perception interact. There were large congruency effects for both reaction times and discrimination ability, showing people respond better and faster when the motion word is congruent with the visual motion stimulus. Thus, our aware condition replicates the findings in Francken et al. (n.d.). Yet, our effect sizes were bigger, most probably because attention to primes increases their effectiveness (Dehaene et al., 2006; Waechter et al., 2011).

In our unaware condition, we found a reaction time congruency effect between unaware words and motion stimuli that were presented in the RVF. One must keep in mind though that the unaware effects here are small and limited to reaction times, and hence require replication and further investigation. Effect lateralizations to the RVF are often taken as an indication of left-hemisphere, and thus linguistic, processing (Gilbert et al., 2006). However, this is not necessarily the case. It is also possible that lateralizations occur because the left-hemisphere is more concerned with categorical processes in general (Holmes & Wolff, 2012). Within the current design we cannot distinguish between these explanations. Moreover, the lateralization in particular must be interpreted with caution, since visual field did not significantly interact with the congruency effects.

From our finding of unconscious, automatic, language-perception interactions, we tentatively conclude that the sensory-motor involvement in language is not arising from post-semantic processes such as imagery or strategic changes (Hauk & Tschentscher, 2013). This is in line with the limited evidence on this topic, which suggest that sensory-motor areas involved in imagery do not overlap with the areas activated during semantic processes (Willems, Toni, Hagoort, & Casasanto, 2010). Besides, the complete absence of criterion effects in our study further indicates that people did not engage in strategic processing of the words in either of the awareness conditions. This contrasts with the criterion results in Francken et al. (n.d.), possibly because we changed from a motion detection task to a motion discrimination task, where it is more difficult to induce perceptual biases.

The results further contrast with current theories stating that awareness and attention are required for semantic effects to arise (Batterink et al., 2010; Kang et al., 2011; Waechter et al., 2011), but are in line with our expectations based on previous studies that showed effect in the absence of attention (Francken et al., n.d.; Meteyard et al., 2007). Such automatic influence of semantic activation on visual processing may be caused by spreading activation or automatic "cascading" of information (McQueen & Huettig, 2013) from a semantic to a more sensory level. Why the unaware effects were limited to reaction times and not present in the other behavioral measures remains unclear, but it is possible that these effects are operating on different time scales (van Ede, de Lange, & Maris, 2012) or have differing sensitivities.

One may doubt whether we are looking at truly semantic processes since our stimulus set was limited (Kang et al., 2011). However, in previous studies with five times as many words as our behavioral study, similar effects were observed (Francken

et al., n.d.). There, the number of primes was so extensive that non-semantic processing seems rather unlikely. Still, obviously we cannot prove at this point that participants in our study used similar strategies as the participants in that study. Response conflict explanations of the congruency effects are implausible because the semantic primes in our study were not coupled to a particular response.

It is difficult to establish the complete absence of word awareness in the unaware condition. It is possible that by using the catch trials to infer prime awareness we somewhat underestimated awareness of the primes because participants were switching from the motor task to the catch task (Vachon & Jolicœur, 2012). Nonetheless, we believe that basing our awareness estimate on catch trials was the best possible option, since it reflects what was happening in the actual experiment better than the posttest. Also, we have shown that participants were significantly worse in discriminating motion from neutral verbs in the unaware compared to the aware condition, confirming that there were at least quantitative difference in awareness between the two conditions. We found that in some participants word awareness may have been too high. Nevertheless, control analyses showed that the unaware effects were not only driven by the effects of those participants. Thus, we conclude that the differences in effect of aware and unaware primes arise from genuine differences in word visibility.

We would like to posit an explanation of the lateralized congruency effect in the unaware condition in light of contemporary theories on consciousness. These theories assume that only a sufficiently strong signal allows for "broadcasting" or feedback to other brain areas (Dehaene et al., 2006; Lamme & Roelfsema, 2000). Here, this would correspond to the spreading of the "linguistic" signal to other areas. As far as we know, this is a new way of explaining the hemispheric asymmetries, which is something that is often not done at all (e.g. Gilbert et al., 2006). This explanation would also account for non-lateralized effects in the aware condition. There, activation is strong enough to reach and change representations of visual motion in both hemispheres. More research on this topic is required to make a better model of connectivity and feedback processing within and between the brain regions that are involved. On the one hand, based on the embodied cognition literature (i.e. Saygin et al., 2010) it would be expected that interactions take place at an early, sensory-motor stage, since here the sensory areas are often assumed to be causally involved in the semantic processing. However, such activations are not consistently found (Hauk & Tschentscher, 2013; Pavan & Baggio, 2013; Willems & Casasanto, 2011), thus mechanisms at higher-order brain levels cannot be ruled out. Two different processes, e.g. visual and semantic, may meet at a later stage, interact there (Mahon & Caramazza, 2008) and subsequently send feedback to visual areas. This idea is supported by studies finding activations in areas close to early sensory-motor areas (Alink, Euler, Kriegeskorte, Singer, & Kohler, 2012; Bedny et al., 2008; Humphreys et al., 2013; Kable et al., 2005) instead of within those sensory-motor areas. What's more, using the same paradigm in a fMRI study, Francken et al. (n.d.) found that left hemisphere (language related) temporal areas, not visual areas, were most strongly related to the congruency effects they found. Importantly, such high-level interactions could still happen at a relatively high speed (Foxe & Simpson, 2002) and outside of awareness.

All in all, further research is vital to get a better understanding of when and why certain brain areas or processes are involved in language-perception interactions (Willems & Francken, 2012) and to what extent they are truly semantic. We are currently working on such experiments, employing fMRI to shed light on the neural mechanisms underlying language-perception interactions in both the aware and unaware conditions. We will present initial findings from this exploratory line of research in Experiment 2.

3. Experiment 2: Automatic activation of visual areas by motion words

While our first experiment was aimed at finding out how automatic language-perception interactions are, in our second experiment we look at the automaticity and specificity of visual area activations in response to motion words. To accomplish this, we used a language task (1-back task) in the fMRI scanner, in which we present neutral words and words implying upward or downward motion. A motion localizer was used to pinpoint areas involved in the processing of visual motion.

The goal of the experiment is two-fold. First, we want to see whether motion verbs (compared to non-motion verbs) in a 1-back task activate visual cortex. Since a 1-back task is not specifically semantic in nature, visual cortex involvement would be interpreted as automatic. We expect motion words to activate sensory area hMT+/V5 or directly surrounding areas (Bedny et al., 2008; Saygin et al.,

2010) more than neutral words. Second, to get a more detailed view on the semantic information contained within visual areas, we will use a multivariate classification approach. This approach is not only sensitive to global changes in activation, but also to fine-grained patterns of activation (Haynes & Rees, 2006). As such, it gives a better means to investigate the specificity of semantic information available in occipital areas. We will try to classify different semantic categories in regions of the visual cortex that are also implicated in visual motion processing. To investigate the specificity of representations even further, we will try to classify semantic motion after training a classifier on motion from the motion localizer. Together, these analyses will give a first idea of the representation of semantic motion information in visual areas, and whether these representations overlap with those of actual visual motion stimulation.

3.1 Methods

3.1.1 Participants

Twenty-six healthy, right-handed, native Dutch speaking participants (21 female, age 22.7 ± 2.9 years) with normal or corrected-to-normal vision and no reading problems took part in the experiment. All participants gave written informed consent, in accordance with the declaration of Helsinki and guidelines of the local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands). Compensation was 50-55 Euros, or an equivalent in course credit. One participant was excluded due to excessive head movement (> 5 mm). Another participant had to be excluded because analyses showed a deviant pattern (right-hemisphere dominance) of language lateralization, which would cause large variability in the imaging data. All data analyses were performed on the remaining 24 participants.

3.1.2 Stimuli

As in the behavioral study, stimuli were generated using the Psychophysics Toolbox (Brainard, 1997) for MATLAB (MathWorks, Natick, MA, US). They were presented using a rear-projection screen using an EIKI projector (refresh rate 60 Hz, 1024 × 768 resolution). Words and motion stimuli were presented in white (325 cd/m2) on a gray background (72 cd/m2).

1-back task. For the 1-back task, we used words from the same semantic categories as in the behavioral experiment: upward verbs, downward verbs and non-motion verbs. Here, both the original word list and the training list were used as experimental stimuli. This left us with 10 different words per category, divided over two different lists. We matched the categories on multiple characteristics, such that there were no significant differences on word length, lexical frequency (CELEX database, celex.mpi.nl), number of syllables and word concreteness (all p's > .05; (Supplementary Table 2). We also created random letter strings (6-8 letters) made up of consonants.

Motion localizer. The visual motion stimulus was a random-dot motion pattern consisting of white dots (speed: 6 deg/s, density: 2.5 dots/deg2, 100% coherence) presented in an annulus aperture at the vertical midline. Motion was presented either left, centrally or right on the screen. The eccentric motion (8.5 deg horizontal eccentricity) and central motion had an aperture radius of 7.5 and 9 degrees respectively. The central stimulus had a central aperture of 1 degree to allow for the fixation cross to be presented. The way of presenting dots was the same as reported in the behavioral experiment, except that there was a limited dot lifetime of 200 ms, to prevent crowding at the aperture border.

3.1.3 Procedure

Participants performed the motion localizer and word localizer during different scanning sessions. During those sessions we also collected data for other experiments (not reported here).

1-back task. The task consisted of 18 blocks (14 in the first participant) of 5 trials each. Per block, there was one trial for each condition (upward, downward, neutral, letter strings) and an additional fixation trial at the end of the block, to be used as a baseline condition. The order of the trials within a block was pseudo-randomized, with the exception of the fixation trial that was always the last trial of a block.

One trial (15 s) was built up of 25 presentations of a stimulus (300 ms) followed by a fixation cross (300 ms). Words were presented in the center of the screen using capital letters in a mono-spaced font ('lucida typewriter'). Participants performed a one-back task in which they indicated occasional repetitions of the same word. Trials were constructed such that they contained 1-4 repetitions at random moments. The last stimulus of a trial was never a repetition.

Between two trials there was an ITI of 1s. Halfway the task there was a short 1 minute break. Both word lists were used during half of the blocks,

the order of the lists was pseudo-randomized over blocks.

Motion localizer. The motion localizer consisted of 10 blocks of 7 trials with a 30 s break halfway the task. Motion presentation occurred in two directions (upward, downward) and at three different locations of the screen (left, center, right). Each combination of motion location and motion direction was present in every block, and counterbalanced across the trials in that block. The last trial of a block was always a fixation trial during which only a fixation cross was presented. Every trial lasted 16 s and was preceded by an ITI of 1 s. Participants were instructed to keep fixating, and press a button when the fixation cross changed color from white to dark grey. Such changed happened two or three times during a trial, at random intervals. A fixation change never occurred in the first 750 or last 1500 ms of a trial, to ensure participants had a realistic chance to spot the change and respond to it.

3.1.4 Image data acquisition

Functional MRI-images were acquired on a 3T Skyra MRI scanner (Siemens, Erlangen, Germany), with a T2*-weighted gradient-echo EPI sequence (TR = 2000 ms, TE = 30 ms, 29 transversal slices, distance factor 20%, voxel size 2x2x1.7 mm, flip angle = 80°, FOV = 192 mm). Whole-brain structural images were acquired with a T1-weighted rapid gradient-echo sequence (TR = 2300 ms, TE = 3.03 ms, voxel size 1x1x1 mm).

3.1.5 Image data preprocessing

We used SPM8 (http://www.fil.ion.ucl.ac.uk/spm; Wellcome Trust Centre for Neuroimaging, London, UK) for image preprocessing and univariate analyses. The first four volumes of each run were discarded to allow for scanner equilibration. All functional images were spatially realigned to the mean image and temporally aligned to the onset of the first slice of each volume. The structural image and functional images were co-registered. For the univariate analyses, images were also normalized to a standard T1 template in MNI space, using both linear and nonlinear parameters. Normalized images were smoothed with an isotropic Gaussian kernel of 8 mm full width at half maximum (FWHM). A high-pass filter (cut-off 128 s) was applied to remove low frequency signal drifts.

3.1.6 Univariate analyses

For each subject, data for each of the localizers was modeled using a block-wise approach within the general linear model (GLM) framework. Regressors representing each of the conditions were obtained by convolving the trials within the blocks with a canonical HRF. Head movement regressors (6) and one regressor modeling breaks and instruction screens were used as nuisance regressors.

We performed whole-brain statistical inference, using a cluster-level statistical test to assess clusters of significant activation. We used a family-wise error corrected cluster threshold of $p <. 05$, based on an auxiliary voxel threshold of $p < .001$ at the whole-brain level.

3.1.7 Multivoxel pattern analyses

Multivoxel pattern analyses (MVPA) were performed on the non-normalized, non-smoothed images. We defined multiple regions of interest (ROI) in which the MVPA's were performed. ROIs of left and right occipital cortex and an ROI of the whole (left + right) occipital cortex (OC) were defined based on combinations of standard masks of occipital areas (Brett, Anton, Valabregue, & Poline, 2002) within the MarsBar (Tzourio-Mazoyer et al., 2002) toolbox for SPM. We also defined hMT+/V5 ROI's in both hemispheres by contrasting LVF > RVF motion for the right hemisphere ROI and RVF > LVF motion for the left hemisphere ROI. Then, we took the group-level peak coordinates (MNI coordinates [46 70 4] for right hMT+/V5, [-46 70 8] for left hMT+/V5) closest to the reported location of this area in related studies (Bedny et al., 2008; Saygin et al., 2010) and drew a 10 mm sphere around it. Subsequently, re-normalization parameters were used to transform all group-level ROIs back to subject-space.

For all five ROI's, per subject we computed t-value maps based on the contrast of centrally presented visual motion compared to fixation. From the resulting statistical maps of the ROI's, we took a selection of the most active voxels. Various number of voxels selections were made: 10, 50, 100, 150, 200, 300, 400. The statistics reported in the Results section are based on a selection of 200 voxels. Still, to rule our spurious results we always ensured that similar results were obtained with other number of voxels selections (see also Supplementary Fig. 2 and Fig. 3).

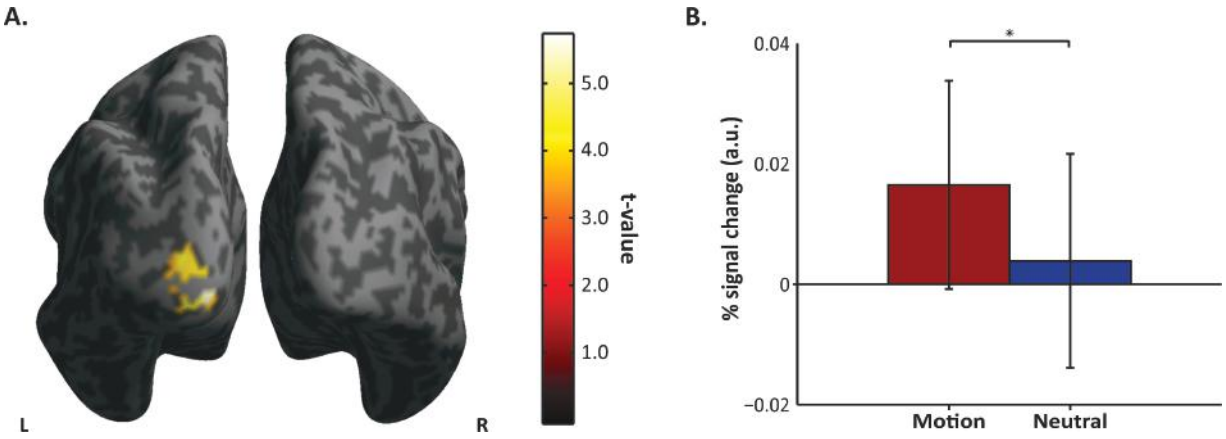


Fig. 4 Motion words activation in the 1-back task. **A.** Rear view of the brain, showing the significantly activated cluster (motion words > neutral words) in the visual cortex. **B.** % Signal change (a.u.) for motion (red) and neutral (blue) words within the activated cluster (* = p < .001). Errorbars denote SE.

Each trial was modeled by a separate GLM containing two regressors. One regressor included the selected trial, the other regressor contained all other trials (Mumford, Turner, Ashby, & Poldrack, 2012). Both regressors were convolved with a canonical HRF. This resulted in a pattern of voxel activations that was specific for the selected trial. The t-values obtained for each voxel were analyzed using MVPA classification methods (Beckett, Peirce, Sanchez-Panchuelo, Francis, & Schluppeck, 2012; Haynes & Rees, 2006; Kamitani & Tong, 2005, 2006).

For each MVPA, a linear Support Vector Machine (SVM) was trained to distinguish two categories of trials based on the activation patterns in the selected voxels. Two different methods were used to get a reliable, cross-validated, estimate of classifier accuracy: a blockwise leave-one-out strategy and a multiple sets strategy. For the blockwise leave-one-out strategy, useful in classifying within one task, the SVM was trained on relevant trials from all but one block and then tested on the left out block. This was repeated until each block had served as a test block, and performance was then averaged over blocks. The multiple sets strategy was used when training and testing on the 1-back task. Here, we used one stimulus set (one of the two word lists) for training of the SVM and the other one for testing. This was done twice, using both lists for training and a testing once, and then averaged of the two results. This way, classification based on specific low-level word characteristics can be ruled out, since these low-level characteristics differ between the two lists.

To test whether decoding accuracy was significant, we computed one-sided t-tests comparing the average accuracy level against chance level performance. Subsequently, post-hoc statistical tests were performed to compare classifier accuracy under different conditions.

3.2 Results

3.2.1 Univariate analysis

To find out whether motion words activate sensory areas in our 1-back task, we performed a univariate GLM (motion words > neutral words) on a whole-brain level. We found one significant cluster ($t(23) = 6.28$, $p(\text{cluster}) < .001$) in the left fusiform gyrus (MNI coordinates [-16 -92 -10]), which corresponds roughly to area V3 of the visual cortex (Rottschy et al., 2007). The same area was also activated by visual motion, as is shown by the extensive overlap of activations in Supplementary Figure 1. As expected, there were no significant activations for the contrast of neutral words > motion words

3.2.2 MVPA

We performed MVPA's on the 1-back task, the motion localizer and subsequently on a combination of the two, to find out more about motion information available in the visual cortex. Our main results will be based on the 1-back task data, which can provide insight into the semantic motion information that is represented within visual areas that are also implicated in visual motion. The motion localizer data is analyzed mostly for control purposes. Moreover, we wanted to make sure within-task classification was possible, since if this is not the case, it seems unlikely that we will get any results in the between-task analyses we are also planning to do, as voxel patterns are expected to differ more between-task than within-task.

Table 1. Overview of all MVPA's that are reported for Experiment 2. For the 1-back task, the "task" columns also indicate which word list(s) were used.

	Train classifier on:		Test classifier on:		ROI*	Classification accuracy sign.
	Task	Categories	Task	Categories		
1	Motion localizer	Up/down motion	Motion localizer	Up/down motion	lOC	Yes
					rOC	Yes
2	Motion localizer	Up/down motion	Motion localizer	Up/down motion	lMT+	No
					rMT+	No
3	Word localizer 1+2	Word category	Word localizer 1+2	Word category	OC	Yes
					lOC	Yes, except down
					rOC	Yes
4	Word localizer 1/2	Word category	Word localizer 2/1	Word category	OC	Yes
	Word localizer 2/2		Word localizer 1/2			
5	Motion localizer	Up/down motion	Word localizer 1+2	Upward / downward words	OC	No

*lOC = left occipital cortex, rOC = right occipital cortex, OC = occipital cortex, lMT+ = left hMT+/V5, rMT+ = right hMT+/V5

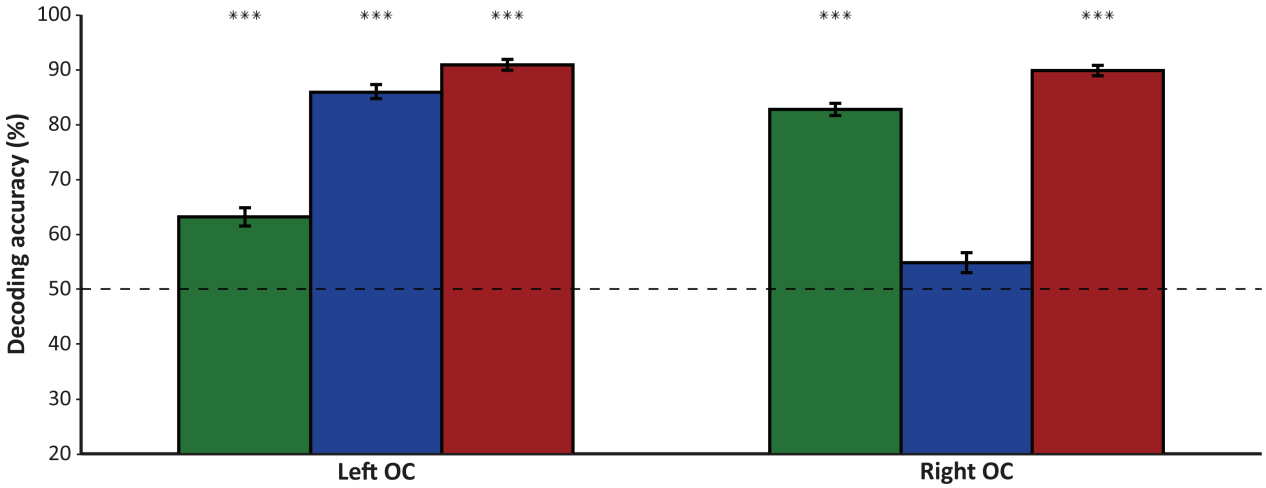


Fig. 5 Motion localizer MVPA: classification accuracy. Bars represent mean classification accuracy (% correctly identified) for both ROI's (left OC, right OC) based on a voxel selection of 200 voxels. Accuracy was tested against chance level (50%; dashed line). Independent of the location at which visual motion was presented (blue = LVF, green = RVF, red = central), classification was above chance level (* = $p < .05$, ** = $p < .01$, *** = $p < .001$), except for RVF motion decoding in the right OC. Errorbars denote SE.

Table 1 gives an overview of all classification analyses that were performed. Decoding results reported below are based on classification within the 200 voxels most sensitive to centrally presented motion within the designated ROI.

3.2.3 Motion localizer: Decoding visual motion direction

To establish whether motion direction could be classified within the motion localizer data, we ran MVPA's in both the left and right occipital cortex (OC) ROI's, separately for motion presented in the LVF, RVF and centrally (Fig. 4). For these analyses, we used the blockwise leave-one-out strategy. Central motion could be reliably decoded from both the lOC ($t(23) = 20.53$, $p < .001$) and rOC ($t(23) = 21.04$, $p < .001$. Motion in the contralateral visual field could be classified in both ROI's (lOC/RVF: $t(23) = 13.87$; roC/LVF: $t(23) = 14.65$, $p < .001$). However, ipsilateral motion could only be decoded in the LOC ($t(23) = 3.94$, $p < .001$) but not in the rOC ($p > .1$).

Since hMT+/V5 is often implicated as the main area responsible for motion perception, we expected motion classification to be possible in this ROI too.

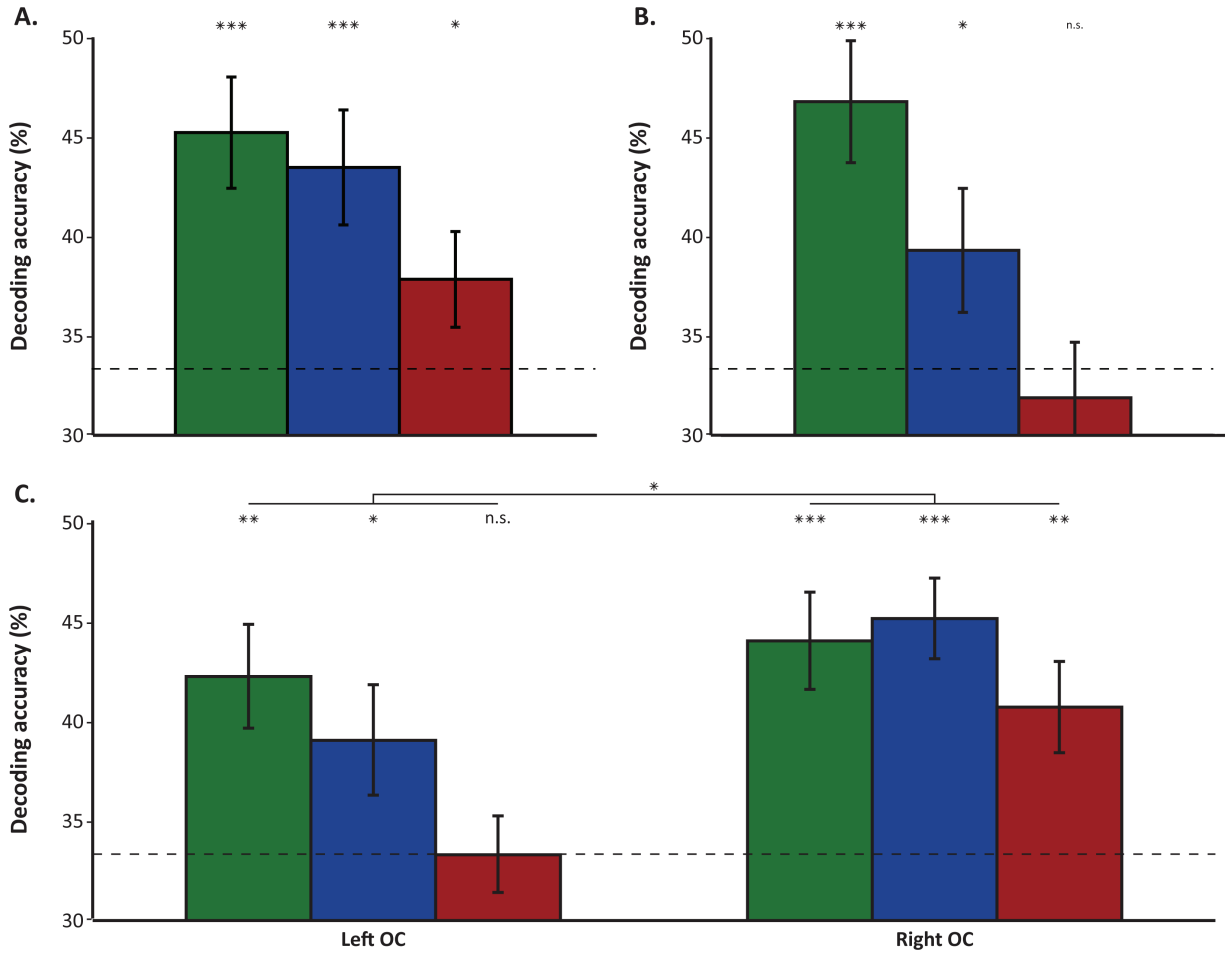


Fig. 6 1-back task: classification accuracies **A.** Bars depict decoding performance for each of the semantic categories, neutral (green), upward (blue) and downward (red) with the blockwise leave-one-out strategy in the OC ROI **B.** See A., but with the multiple-sets strategy. **C.** See A., but analysis is separately done for left and right OC, showing higher decoding performance in the right OC (* = $p < .05$, ** = $p < .01$, *** = $p < .001$). Errorbars denote SE.

Surprisingly enough, decoding performance did not surpass chance level for any of the analyses (all p's > .05) that were performed.

3.2.4 1-back task: Decoding semantic category

We wanted to find out whether we could decode the semantic category (neutral, up, down) of the words in the 1-back task from visual areas that were most responsive to visual motion stimulation. In order to do this, we trained three 2-way SVM's (neutral/up; neutral/down and up/down) and linearly combined the outcome weights to get estimates for each of the word categories (Fig. 6A).

Each of the word categories could be decoded from visual cortex (neutral: $t(23) = 4.25$, $p < .001$; up: $t(23) = 3.50$, $p < .001$; down: $t(23) = 1.87$, $p < .05$), indicating the availability of semantic information in visual areas. A repeated-measures ANOVA was performed to compare classification across the categories, giving a significant result of category effect ($F(2, 22) = 3.83$, $p < .05$, $\eta 2 = .26$), with highest classification accuracy for neutral words ($M = 0.452$, $SE = 0.028$), then for upward words ($M = 0.34$, $SE = 0.03$) and lowest for downward words ($M = 0.38$, $SE = 0.03$). As can be seen in Supplementary Figure 3, decoding accuracy was stable over different selections of number of voxels.

To rule out that decoding performance was driven by overall differences in activation, such as the univariate difference between motion and neutral words in the left OC, we performed two control analyses. First, for each participant we computed the average activation levels (betas) for every semantic category. Although the downward category seems to differ from the other two categories, paired t-tests showed this difference was not significant (all p's > .05). Additionally, this difference cannot explain the decoding results. If overall differences between the downward condition and the other conditions were driving the MVPA's, one would expect decoding for

downward words to be best, while we have shown above that it was worst.

Second, we decoded semantic word category separately in left and right OC separately and afterwards compared them (Fig. 6C). The classification of neutral words was significant in both ROI's (lOC: $t(23) = 3.41$, $p < .01$; rOC: $t(23) = 4.38$, $p < .001$), just as the classification of the upward semantic category (lOC: $t(23) = 2.06$, $p < .05$; rOC: $t(23) = 5.83$, $p < .001$). Decoding of downward words, however, was only possible in the right OC ($t(23) = 3.21$, $p < .01$). A repeated-measures ANOVA pointed out that overall decoding was significantly ($F(1,23) = 6.11$, $p < .05$, $\eta 2 = .21$) higher in the right OC ($M = 0.38$, $SE = 0.02$) than in the left OC ($M = 0.43$, $SE = 0.01$), making it even more unlikely that the left-lateralized univariate result was the main cause of overall classification ability.

To exclude the possibility that significant decoding performance was caused by differences in low-level visual features - even though the words from categories were matched - we performed the main analysis again with the multiple sets strategy (Fig. 6B). Decoding of upward ($t(23) = 4.38$, $p < .001$) and neutral ($t(23) = 1.91$, $p < .05$) words was still better than chance. It was not possible to classify downward words, thus for this category we cannot fully rule out the possibility that classification is (partially) dependent on low-level visual characteristics of the words we used.

3.2.5 Overlap between visual and semantic motion direction

Lastly, to find out whether visual motion and semantic motion are represented in a similar way in visual cortex, we ran a between-task classifier in which we tried to decode semantic motion direction (up, down) after training a classifier on upward and downward visual motion. If visual cortex is engaged during the processing of semantic motion in a similar way as during the processing of visual motion, classification accuracy would be high. However, this is not what we found when we did this analysis in the OC ROI. Classification performance did not exceed chance ($p > .5$).

3.3 Discussion

In this second experiment we took a different approach to the question of automaticity of language-perception interactions. Specifically, we employed a 1-back task containing motion words in the fMRI scanner, to find out whether this would automatically activate visual cortices. Furthermore, we wanted to explore the specificity of the semantic information that is available in occipital areas during the execution of this task.

First, our results show that a region in the left visual cortex, perhaps corresponding to area V3 (Rottschy et al., 2007) is significantly more active during the processing of motion words compared to the processing of neutral words. We were surprised that we did not find activation in area hMT+/V5, or the directly surrounding areas, since these are normally implicated in the processing of visual motion (Kamitani & Tong, 2005) or semantic motion (Bedny et al., 2008; Humphreys et al., 2013; Kable et al., 2005; Saygin et al., 2010). We do not know of any other study with similar findings. One important design difference that may relate to this difference is the way of processing the words. Previous studies had a specifically semantic task, while semantic word content could in principle be ignored in our study. It has been shown already that attended and ignored words may be processed differently (Ruz, Wolmetz, Tudela, & McCandliss, 2005). The type of semantic processing in our study may have differed from that in previous research. Nevertheless, we found differential activation based on semantic categories, implying there must have been semantic processing of some sort. This is in line with other findings of automatic semantic processing of words regardless of its relevance to task performance (Relander et al., 2009). Additionally, it is in agreement with a study by McQueen and Huettig (2013) that showed how exposure to printed words automatically leads to the retrieval of conceptual knowledge, possibly because information from one processing level automatically cascades to another level.

Finding that occipital areas (V3) are modulated by semantic motion is not wholly surprising. There is a large body of evidence linking parts of the visual cortex within or around V3 to the processing of visual motion (Beckett et al., 2012; Braddick et al., 2001; Gorbet, Wilkinson, & Wilson, 2013; Kamitani & Tong, 2006).

To go beyond the activation differences that are often reported, we used decoding analyses to look at the specific semantic information available within visual areas. We showed that visual cortex contains information about both the presence of semantic motion and the direction of this motion, as indicated by above-chance decoding for all semantic categories. Decoding ability could not be explained by univariate differences. It is not possible to fully rule out that decoding was partially based on low-level visual characteristics, but even when controlling

for this we could classify two out of three semantic categories with above-chance performance. As far as we know, this makes this the first study to show that the visual cortex contains information not only about the presence of semantic motion (Bedny et al., 2008; McCullough et al., 2012; Saygin et al., 2010), but also about semantic motion direction. Such informational content is a necessary requirement for a causal role of those visual areas in the processing of motion semantics. Indeed, it would seem unlikely that the visual cortex has a role in the processing of semantics if it does not contain the relevant information.

We do not have an explanation as to why decoding performance was higher in right than in left visual cortex. Perhaps, decoding was higher in the right hemisphere because it is implicated in the processing of concrete words (Lindell, 2006). This could also explain why decoding of neutral words is numerically higher in the left hemisphere, that is supposed to be more involved with processing abstract concepts. Another possible explanation for lateralization of decoding performance is that inter-individual localization differences were higher in the left hemisphere, resulting in better decoding performance for the right hemisphere (Simanova et al., 2012).

It is important to stress that the voxels included in these decoding analyses were selected based on their sensitivity to visual motion. Since we can decode the semantic categories (neutral, up, down) from those voxels, it seems that the same voxels contain information on both visual motion and semantic motion. This is in line with our results in Experiment 1, or similar results in other studies, where congruency effect between visual motion and semantic motion are reported (Francken et al., n.d.; Meteyard et al., 2007). Still, the representations were not similar enough to allow the classification of semantic motion after training on visual motion. This can probably be explained by the fact that decoding from one task to another is more difficult, since also task parameters and the form of stimulation differ substantially. Nevertheless, if motion semantics and visual motion were represented in exactly the same way, one would expect that decoding would be possible between the motion localizer and the 1-back task.

Alternative explanations of our findings may be possible. Decoding of semantic category could be based on the imagery of words. However, we do not believe this is likely. Word presentation and stimulus onset asynchrony (SOA) were short (both 300 ms), thus making imagery basically impossible. Furthermore, in Experiment 1 we ruled out controlled processes, such as imagery, as an explanation for the interaction between language and perception by showing congruency effects in the absence of awareness. Nevertheless, it would be good to find ways to rule out imagery even more. For example, future studies may use even shorter presentation durations than we did here.

Another probable explanation of both our univariate and decoding results is that they were caused by differences in imageability, which has been shown to significantly modulate area V3 (Desai et al., 2010). However, as discussed in the Methods, words were matched on concreteness, making this an implausible explanation of the results. Besides, left-hemisphere processing is often assumed to be more extensive for abstract compared to concrete concepts (Binder, Westbury, McKiernan, Possing, & Medler, 2005; Lindell, 2006). In our case, that means concrete (motion) words should result in lower activation than the more abstract neutral words, which is exactly opposite to what we observed. Also, the possibility to differentiate upward from downward words – both relatively concrete word categories – makes an explanation in terms of word concreteness or imageability unlikely.

In conclusion, we have found differential activation to different semantic categories in a 1-back task, suggesting automatic processing of semantic features. More specifically, we see that visual areas (V3) often implicated in the processing of visual motion are more active in response to motion words than to neutral words. Moreover, to the best of our knowledge, our study is the first to report the ability to decode semantic motion from the visual areas that are also implicated in visual motion processing. This hints at an overlap in the way visual motion and semantic motion are represented within visual cortex, and leaves the door open to a causal role of sensory areas in the processing of semantics. While at this point it is not possible to either prove or disprove a causal role for visual activations in the processing of semantics (specific activation of visual cortex in response to semantic content could in principle still result from feedback from higher areas), we show that one of the necessary requirements – specific information – is fulfilled. Neuroimaging studies with higher time resolution may give new insights into this issue.

4. General conclusion

We used two different approaches to investigate how automatic language and perception interact. In

our first experiment, we demonstrated that language can influence visual processing substantially, and that this is not limited to a controlled process. We found a lateralized congruency effect for reaction times in the RVF when words were masked, showing that even when people are unaware of the meaning of a word, it can change subsequent perceptual processes. We take the lateralization of the effect in this case to be caused by the weakness of the unaware signals (Dehaene et al., 2006; van Gaal & Lamme, 2012), which causes signals from language areas not to reach the right hemisphere. The automaticity of the effects most rules out post-semantic explanations, like imagery, of language-perception interactions.

Our second experiment demonstrated differential activation of visual areas to varying semantic categories in a 1-back task, which we assume is not necessarily semantic in nature. Next to indicating that some automatic semantic processing takes place, it shows that visual cortex relates not only to visual motion but also to semantic motion. Moreover, as far as we know we are the first to report detailed representations of semantic motion in visual areas. We do so by showing above chance decoding of upward, downward and neutral words, from voxels that are also modulated by visual motion stimulation. As noted before, this hints at a strong overlap in the way visual motion and semantic motion are represented in the brain.

In conclusion, we tentatively conclude that language and perception seem to be related in an automatic way. Language can change our view on the world, even when we are not aware of the language. Furthermore, language automatically activates sensory-motor cortices in a specific way that allows decoding of motion categories from visual cortex, suggesting there is overlap between the representation of visual and semantic motion. This could potentially - when reproducible - have big impact on the understanding of the neural implementation of semantics and the interactions between language and perception.

5. Acknowledgements

I would like to thank my supervisors for their guidance throughout the project. In particular, I would like to thank Jolien for her unlimited support, confidence and enthusiasm. I would like to extend my gratitude to the Prediction & Attention and Neurobiology of Language research groups for feedback on earlier versions of the project. Finally, I thank the rest of the DCCN staff for creating a pleasant and stimulating work environment throughout my internship.

6. References

Alink, A., Euler, F., Kriegeskorte, N., Singer, W., & Kohler, A. (2012). Auditory motion direction encoding in auditory cortex and high-level visual cortex. *Human Brain Mapping, 33*(4), 969–78. doi:10.1002/hbm.21263

Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 364*(1521), 1235–43. doi:10.1098/rstb.2008.0310

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., … Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America, 103*(2), 449–54. doi:10.1073/pnas.0507062103

Batterink, L., Karns, C. M., Yamada, Y., & Neville, H. (2010). The role of awareness in semantic and syntactic processing: an ERP attentional blink study. *Journal of Cognitive Neuroscience, 22*(11), 2514–29. doi:10.1162/jocn.2009.21361

Beckett, A., Peirce, J. W., Sanchez-Panchuelo, R.-M., Francis, S., & Schluppeck, D. (2012). Contribution of large scale biases in decoding of direction-of-motion from high-resolution fMRI data in human early visual cortex. *NeuroImage, 63*(3), 1623–32. doi:10.1016/j.neuroimage.2012.07.066

Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: the case of action verbs. *The Journal of Neuroscience, 28*(44), 11347–53. doi:10.1523/JNEUROSCI.3039-08.2008

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*(6), 905–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16021798

Braddick, O. J., O'Brien, J. M. D., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception*, *30*(1), 61–72. doi:10.1068/p3048

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176952

Brett, M., Anton, J.-L. L., Valabregue, R., & Poline, J.-B. B. (2002). Region of interest analysis using an SPM toolbox [abstract] Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. In *NeuroImage* (Vol. 16, p. abstract 497).

Chen, Y.-C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology. Human Perception and Performance, 37*(5), 1554–68. doi:10.1037/a0024329

Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences, 16*(8), 411–417. doi:10.1016/j.tics.2012.06.013

De Lange, F. P., van Gaal, S., Lamme, V. A. F., & Dehaene, S. (2011). How awareness changes the relative weights of evidence during human decision-making. *PLoS Biology, 9*(11), e1001203. doi:10.1371/journal.pbio.1001203

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204–11. doi:10.1016/j.tics.2006.03.007

Desai, R. H., Binder, J. R., Conant, L. L., & Seidenberg, M. S. (2010). Activation of sensory-motor areas in sentence comprehension. *Cerebral Cortex, 20*(2), 468–78. doi:10.1093/cercor/bhp115

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Experimental Brain Research, 142*(1), 139–50. doi:10.1007/s00221-001-0906-7

Francken, J. C., Kok, P., Hagoort, P., & de Lange, F. P. (n.d.). The behavioral and neural effects of language on motion perception.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America, 103*(2), 489–94. doi:10.1073/pnas.0509868103

Gorbet, D., Wilkinson, F., & Wilson, H. (2013). Visual regions V2, V3, and MT can discriminate between visual motion trajectories even when you can't. *Journal of Vision, 13*(9), 358–358. doi:10.1167/13.9.358

Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping, 21*(3), 191–201. doi:10.1002/hbm.10157

Hauk, O., & Tschentscher, N. (2013). The Body of Evidence: What Can Neuroscience Tell Us about Embodied Semantics? *Frontiers in Psychology, 4*(50), 1–14. doi:10.3389/fpsyg.2013.00050

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews. Neuroscience, 7*(7), 523–34. doi:10.1038/nrn1931

Holmes, K. J., & Wolff, P. (2012). Does categorical perception in the left hemisphere depend on language? *Journal of Experimental Psychology. General, 141*(3), 439–43. doi:10.1037/a0027289

Humphreys, G. F., Newling, K., Jennings, C., & Gennari, S. P. (2013). Motion and actions in language: semantic representations in occipito-temporal cortex. *Brain and Language, 125*(1), 94–105. doi:10.1016/j.bandl.2013.01.008

Kable, J. W., Kan, I. P., Wilson, A., Thompson-Schill, S. L., & Chatterjee, A. (2005). Conceptual representations of action in the lateral temporal cortex. *Journal of Cognitive Neuroscience, 17*(12), 1855–70. doi:10.1162/089892905775008625

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8*(5), 679–85. doi:10.1038/nn1444

Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology, 16*(11), 1096–102. doi:10.1016/j.cub.2006.04.003

Kang, M.-S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *The Journal of Neuroscience, 31*(38), 13535–45. doi:10.1523/JNEUROSCI.1691-11.2011

Kiefer, M., Adams, S. C., & Zovko, M. (2012). Attentional sensitization of unconscious visual processing: Top-down influences on masked priming. *Advances in Cognitive Psychology, 8*(1), 50–61. doi:10.2478/v10053-008-0102-4

Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *Neuroreport, 11*(11), 2401–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10943693

Klemfuss, N., Prinzmetal, W., & Ivry, R. B. (2012). How does language change perception: a cautionary note. *Frontiers in Psychology, 3*(March), 78. doi:10.3389/fpsyg.2012.00078

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., … Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain, 123 Pt 12*, 2512–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11099452

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23*(11), 571–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11074267

Landau, A. N., Aziz-Zadeh, L., & Ivry, R. B. (2010). The influence of language on perception: listening to sentences about faces affects the perception of faces. *The Journal of Neuroscience, 30*(45), 15254–61. doi:10.1523/JNEUROSCI.2046-10.2010

Lindell, A. K. (2006). In your right mind: right hemisphere contributions to language processing and production. *Neuropsychology Review, 16*(3), 131–48. doi:10.1007/s11065-006-9011-9

MacMillan, N. A. (2002). Signal Detection Theory. In H. Pashler (Ed.), *Stevens Handbook of Experimental Psychology: Vol. 4. Methodology* (3rd ed., pp. 43 – 90). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi:10.1002/0471214426

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. doi:10.1016/j.jphysparis.2008.03.004

McCullough, S., Saygin, A. P., Korpics, F., & Emmorey, K. (2012). Motion-sensitive cortex and motion semantics in American Sign Language. *NeuroImage, 63*(1), 111–8. doi:10.1016/j.neuroimage.2012.06.029

McQueen, J. M., & Huettig, F. (2013). Interference of spoken word recognition through phonological priming from visual objects and printed words. *Attention, Perception & Psychophysics*. doi:10.3758/s13414-013-0560-8

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: language affects low-level visual perception. *Psychological Science, 18*(11), 1007–13. doi:10.1111/j.1467-9280.2007.02016.x

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology, 18*(17), R732–R733. doi:10.1016/j.cub.2008.07.016

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage, 59*(3), 2636–43. doi:10.1016/j.neuroimage.2011.08.076

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science, 13*(5), 416–424. doi:10.1111/1467-9280.00474

Nestor, A., Behrmann, M., & Plaut, D. C. (2013). The neural basis of visual word form processing: a multivariate investigation. *Cerebral Cortex, 23*(7), 1673–84. doi:10.1093/cercor/bhs158

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychological Science, 24*(6), 836–43. doi:10.1177/0956797612461449

Pavan, A., & Baggio, G. (2013). Linguistic representations of motion do not depend on the visual motion system. *Psychological Science, 24*(2), 181–8. doi:10.1177/0956797612450882

Rahnev, D., Lau, H., & de Lange, F. P. (2011). Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain. *The Journal of Neuroscience, 31*(29), 10741–8. doi:10.1523/JNEUROSCI.1478-11.2011

Relander, K., Rämä, P., & Kujala, T. (2009). Word semantics is processed even without attentional effort. *Journal of Cognitive Neuroscience, 21*(8), 1511–22. doi:10.1162/jocn.2009.21127

Rottschy, C., Eickhoff, S. B., Schleicher, A., Mohlberg, H., Kujovic, M., Zilles, K., & Amunts, K. (2007). Ventral visual cortex in humans: cytoarchitectonic mapping of two extrastriate areas. *Human Brain Mapping, 28*(10), 1045–59. doi:10.1002/hbm.20348

Ruz, M., Wolmetz, M. E., Tudela, P., & McCandliss, B. D. (2005). Two brain pathways for attended and ignored words. *NeuroImage, 27*(4), 852–61. doi:10.1016/j.neuroimage.2005.05.031

Saygin, A. P., McCullough, S., Alac, M., & Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *Journal of Cognitive Neuroscience, 22*(11), 2480–90. doi:10.1162/jocn.2009.21388

Simanova, I., Hagoort, P., Oostenveld, R., & van Gerven, M. a J. (2012). Modality-Independent Decoding of Semantic Information from the Human Brain. *Cerebral Cortex*, 1–9. doi:10.1093/cercor/bhs324

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition*, 122(1), 91–5. doi:10.1016/j.cognition.2011.08.017

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15*(1), 273–89. doi:10.1006/nimg.2001.0978

Vachon, F., & Jolicœur, P. (2012). On the automaticity of semantic processing during task switching. *Journal of Cognitive Neuroscience, 24*(3), 611–26. doi:10.1162/jocn_a_00149

Van den Bussche, E., Smets, K., Sasanguie, D., & Reynvoet, B. (2012). The power of unconscious semantic processing: The effect of semantic relatedness between prime and target on subliminal priming. *Psychologica Belgica, 52*(1), 59–70.

Van Ede, F., de Lange, F. P., & Maris, E. (2012). Attentional cues affect accuracy and reaction time via different cognitive and neural processes. *The Journal of Neuroscience, 32*(30), 10408–12. doi:10.1523/JNEUROSCI.1337-12.2012

Van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in Human Neuroscience, 6*(May), 121. doi:10.3389/fnhum.2012.00121

Van Gaal, S., & Lamme, V. A. F. (2012). Unconscious high-level information processing: implication for neurobiological theories of consciousness. *The Neuroscientist, 18*(3), 287–301. doi:10.1177/1073858411404079

Waechter, S., Besner, D., Jennifer, A., & Stolz, J. A. (2011). Basic processes in reading: Spatial attention as a necessary preliminary to orthographic and semantic processing. *Visual Cognition, 19*(2), 171–202. doi:10.1080/13506285.2010.517228

Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–20. Retrieved from http://link.springer.com/article/10.3758/BF03202828

Willems, R. M., & Casasanto, D. (2011). Flexibility in embodied language understanding. *Frontiers in Psychology, 2*(June), 116. doi:10.3389/fpsyg.2011.00116

Willems, R. M., & Francken, J. C. (2012). Embodied cognition: taking the next step. *Frontiers in Psychology, 3*(December), 582. doi:10.3389/fpsyg.2012.00582

Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2010). Neural Dissociations between Action Verb Understanding and Motor Imagery. *Journal of Cognitive Neuroscience, 22*(10), 2387–2400. Retrieved from http://dx.doi.org/10.1162/jocn.2009.21386

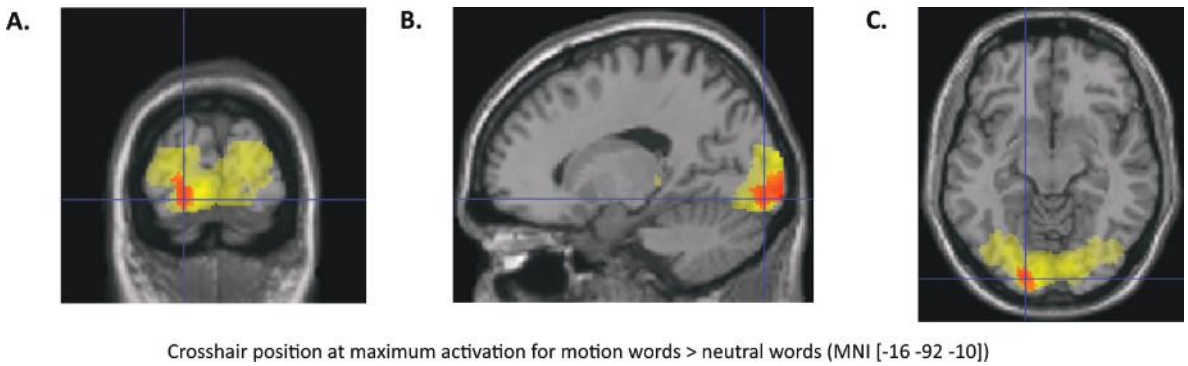
Witzel, C., & Gegenfurtner, K. R. (2011). Is there a lateralized category effect for color? *Journal of Vision*, *11*(12), 16. doi:10.1167/11.12.16

Supplementary Table 1. Characteristics of the words in word list 1 (Exp. 1 + Exp. 2). Given are frequency (CELEX database), word length (# letters), number of syllables and mean post-experiment collected ratings of word concreteness (0 = abstract to 5 = concrete) and word motion (negative = down, positive = up). Word motion ratings confirmed the word categorization that was made previous to the experiment.

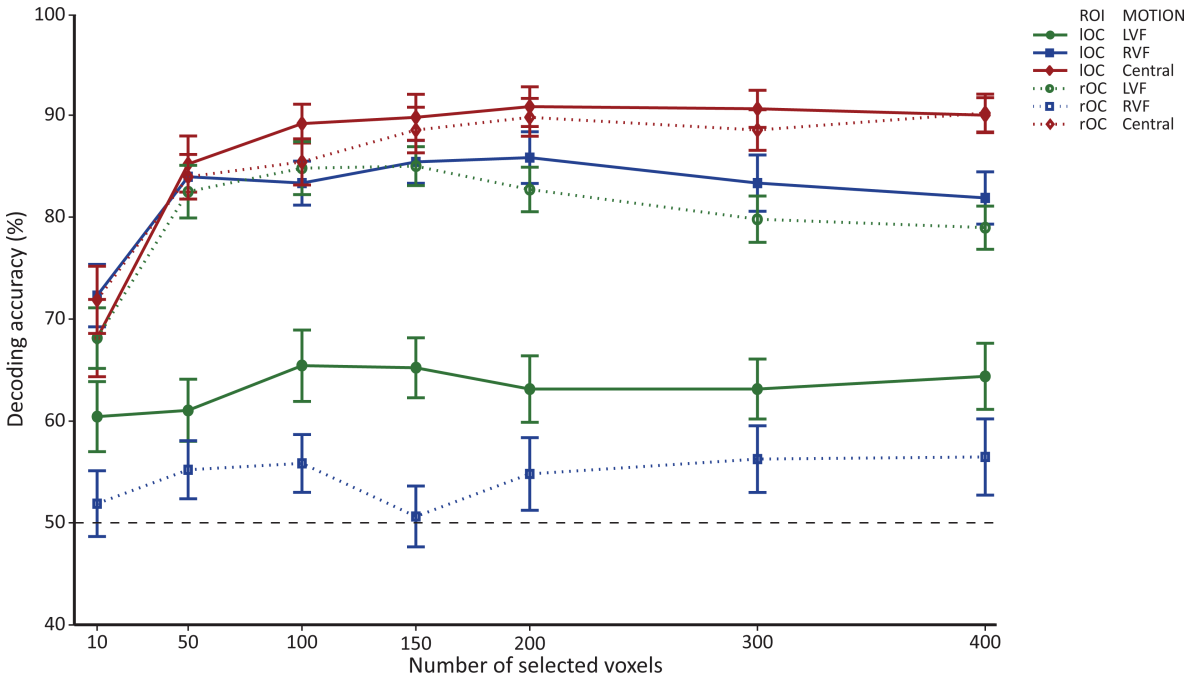
		Word characteristics			Ratings	
Category	**Word**	**Frequency**	**Length**	**Syllables**	**Concreteness**	**Direction**
Up	groeien	138	7	2	3.4	3.8
	klimmen	49	7	2	4.8	4.6
	opgaan	24	6	2	2.6	3.1
	opstaan	47	7	2	4.5	2.6
	stijgen	79	7	2	5.0	4.5
Down	dalen	55	5	2	4.4	-4.8
	duiken	46	6	2	4.4	-4.2
	neergaan	1	8	2	3.3	-4.4
	zakken	67	6	2	3.4	-1.9
	zinken	19	6	2	4.4	-3.6
Neutral	kosten	84	6	2	3.2	0.0
	wedden	9	6	2	3.4	0.2
	filmen	6	6	2	4.4	0.0
	ruilen	12	6	2	4.5	0.0
	rusten	59	6	2	3.1	0.0

Supplementary Table 2. Characteristics of the words in word list 2 (training list Exp. 1; experimental list Exp. 2). Given are frequency (CELEX database), word length (# letters), number of syllables and post-experiment collected ratings of word concreteness (0 = abstract to 5 = concrete) and mean word motion (negative = down, positive = up). Word motion ratings confirmed the word categorization that was made previous to the experiment.

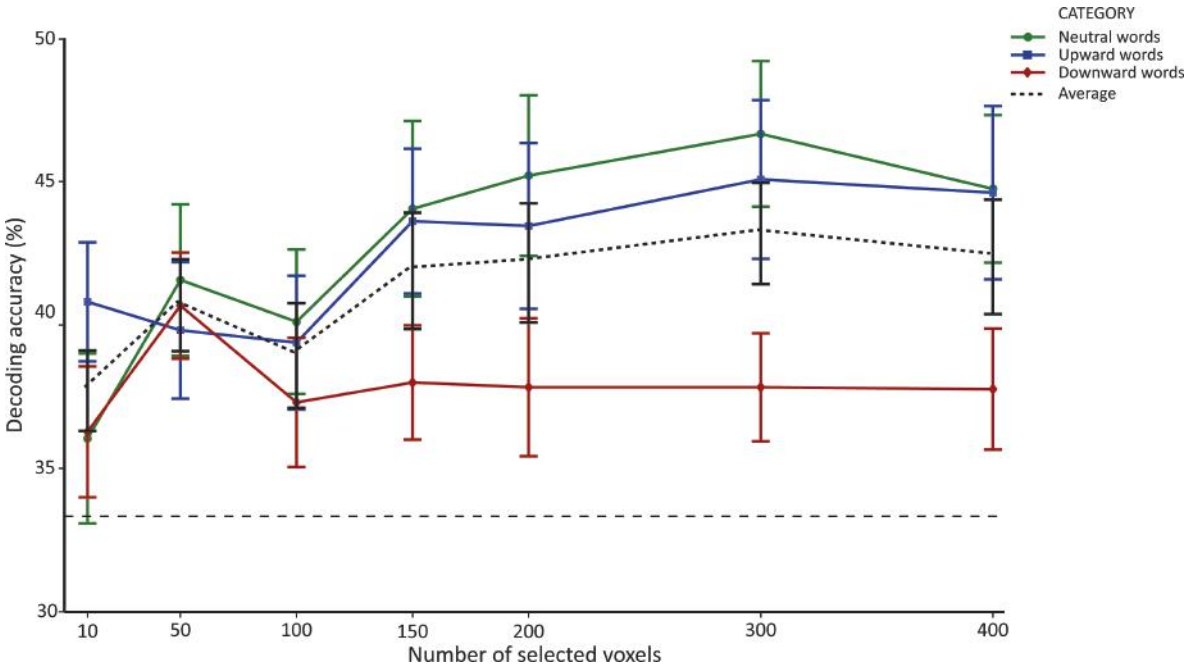
		Word characteristics			Ratings	
Category	**Word**	**Frequency**	**Length**	**Syllables**	**Concreteness**	**Direction**
Up	heffen	54	6	2	3.9	2.7
	opkomen	52	7	3	2.7	2.4
	stapelen	8	8	3	4.3	3.2
	strekken	46	8	2	3.6	2.2
	verhogen	30	8	3	4.0	3.4
Down	afnemen	37	7	3	3.1	-0.8
	landen	18	6	2	4.2	-2.3
	sneeuwen	6	8	2	4.9	-1.5
	verlagen	11	8	3	3.4	-3.3
	tuimelen	7	8	3	3.2	-2.1
Neutral	wensen	85	6	2	2.7	0.0
	bedaren	11	7	3	1.9	-0.5
	zweten	15	6	2	4.7	0.2
	gloeien	24	7	2	3.9	0.2
	rouwen	3	6	2	2.9	-0.1



Supplementary Fig. 1 Overlap of areas responsive to visual motion and semantic motion. The images show **A.** a coronal, **B.** sagittal and **C.** horizontal view of the brain respectively. Overlaid on the anatomical image are the t-values from the motion localizer (motion > fixation; thresholded at $p < .05$ FWE) and the 1-back task (motion words > neutral words; thresholded at $p < .001$). The area that responds to motion words compared to neutral words (red) largely overlaps (orange) with the area that responds to actual visual motion (yellow).



Supplementary Fig. 2 Motion classification accuracy over different numbers of voxel selections. The graph shows decoding performance (% correctly classified) for all ROI (left OC, right OC) x motion location (LVF, RVF, central) combinations. Independent of the number of voxels that was selected for the classifier, the same pattern emerges. Chance level performance is indicated by the black dashed line at 50%. Errorbars denote SE.



Supplementary Fig. 3 Semantic category classification accuracy over different numbers of voxel selections. The graph shows decoding performance (% correctly classified) for all ROI's (OC, left OC, right OC). Independent of the number of voxels that was selected for the classifier, the same pattern emerges (except for a selection of 10 voxels). Upward vs. neutral words (green) decoding is best, downward vs. neutral (blue) is intermediate and upward vs. downward (red) is worst. Chance level performance is indicated by the black dashed line at 50%. Errorbars denote SE.

Pointing for me or for you? Kinematics and Neural Correlates of Communicative Pointing

Anke Murillo Oosterwijk[1]
Supervisors: Ivan Toni[1], Arjen Stolk[1], Miriam de Boer[1], Lennart Verhagen[1]

[1] *Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*

Humans have a unique ability to convey complex messages by means of simple pointing movements intended to change the mental state of an addressee. However, the primate visuomotor system has evolved to control instrumental actions that are designed to change the physical environment. Thus, at some level of organisation, the control of communicative actions needs to interface with the control of instrumental actions. The nature of this interface is heavily debated. A question that remains is to what extent representations of the communicative goal interact with the motor system. Do they specify additional constraints that are executed in an instrumental fashion, or are they flexibly modulating low-level features of the action plan? We address this issue by comparing the kinematics of communicative and instrumental pointing actions in the context of a cooperative game. We found that the presence of communicative intent elicited adjustments in movement trajectories and end-points that were adapted to the communicative content and the viewpoint of the addressee. This shows that specifics of the communicative goal have access to fine-tuned motor control at multiple levels of organization, incorporating knowledge of the object, the context and the addressee into an action plan. In an accompanying pilot functional magnetic resonance imaging (fMRI) experiment we show the feasibility of the experimental setup that allows real-life pointing in the MR-environment, opening up the possibility of determining the neural correlates of communicative pointing using fMRI. We show that an occipital-parieto-frontal network is differentially activated depending on the type of message that is being conveyed by a communicative pointing action.

Keywords: communication, kinematics, social cognition

Corresponding author: Anke Murillo Oosterwijk; E-mail: a.murillooosterwijk@donders.ru.nl

1. Introduction

Intentional communication is a hallmark of human cognition. The power of our communicative abilities is exemplified by pointing. Pointing is a biomechanically seemingly simple movement by which we can express complex messages: for example, a pointing movement towards an apple can mean "can I eat this apple?" or "here is the apple you are looking for". On the phenomenological level the pointing movement closely resembles a movement shared with other primates: the instrumental reach-to-point movement. In fact, it has been suggested that the evolution of our communicative abilities originates from adaptations in our action repertoire (Tomasello, 2008), with a prominent role for pointing (Behne, Liszkowski, Carpenter & Tomasello, 2012; Liszkowski et al., 2004; Tomasello, 2006). But while the primate visuomotor system has evolved to control instrumental actions that change the physical environment, like regasping an apple to eat it, communicative actions differ critically on the intentional level: they are intended to change the mental state of an addressee, such as referring to the apple as a suggestion to eat. Because both communicative and instrumental actions use the same movement apparatus, the control of communicative actions needs to interface with the control of instrumental actions at some level of motor organization. A dominant view proposes that communicative and instrumental actions can be directly mapped onto one another (Rizzolatti & Arbib, 1998). This view predicts that the kinematics and neural substrates of instrumental and communicative actions are identical since both types of action are produced and understood by means of a visuomotor mirror system. Here, addressees are believed to understand the communicative value of an action by virtue of automatic priming between interlocutors (Garrod & Pickering, 2004; Glenberg & Gallese, 2012), or brain-to-brain coupling (Hasson et al., 2012). However, theoretical accounts of human communication, as well as some recent empirical evidence, suggests that planning communicative actions requires additional computations that go beyond situational sensory evidence and lie outside of the visuomotor system (Stolk, Hunnius, Bekkering & Toni, 2013). We aim to understand how, and at what level of organization, higher-order conceptual representations of the communicative goal interact with low-level motor control processes. We test two hypotheses.

The first hypothesis posits that communicative computations specify additional movement constraints that are implemented in an instrumental action plan to signal communicative intent. This view is in line with the notion that the communicative faculty, can top-down instruct 'lower-level' systems (Carruthers, 2002) such as the motor system, and that conceptual processing can inform but does not penetrate motor control (Goodale & Milner, 1992; Milner & Goodale, 2008). Accordingly, planning a communicative action is not substantially different from planning an instrumental action, but an action could be supplemented with an ostensive tag to aid the addressee. In a study by Peeters and colleagues (2013) participants pointed slower and longer when their pointing action was informative for an addressee observing their movements. Such a deliberately slow movement during pointing, especially near the end, could both mark and enhance the opportunity for the addressee to accumulate sensory evidence needed to infer the content. Importantly, such adaptations do not carry the communicative content itself.

The second hypothesis posits that communicative computations have access to motor planning and can adaptively influence action execution to carry communicative content. This could aid the addressee, not only in allowing sufficient opportunity to decode the communicative message, but also in inferring the communicative intention to be the one most probable given the form of the action (Baker, Saxe & Tenenbaum, 2009; Blokpoel et al., 2012; Brennan & Clark, 1996; van Rooij et al., 2011). Accordingly, the planning of instrumental and communicative actions is substantially different and communicative intent, message content and specifics of the addressee could potentially influence action execution in specific ways. This influence could take place at a relatively high level of motor organization, for example by specifying a spatially adjusted end-point target for the movement, or at a lower level as well, affecting the whole trajectory of the movement beyond the requirements for the end location. In line with this hypothesis, Cleret de Langavant and colleagues (2011) found a slightly different movement trajectory and end point distribution when participants pointed with the intention to signal a referent object to an addressee.

We test these two hypotheses by means of two interrelated experiments. First, in a behavioural experiment participants played a cooperative game that required two players to jointly map three coloured tokens to three geometric objects on a screen. We measured several kinematic parameters including reaction times, spatial and temporal dynamics, and end-point locations, while putting

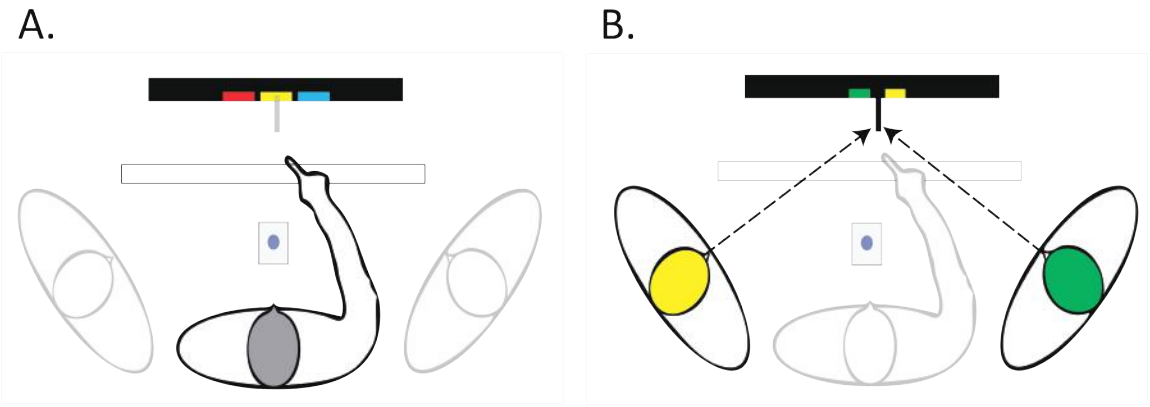


Fig. 1 The experimental setting of experiment 1. **A.** top-view from the setup depicting the communicator tokens on the bottom computer screen. The IR-frame was located within pointing distance of the communicator who was seated in the middle (gray head). **B.** top-view from the setup depicting the goal tokens on the bottom computer screen. The left addressee was unable to see the goal-token when on the right of the wall and vice versa for the right addressee.

different communicative demands on the pointing actions. We manipulated three features of the action goal: the action intention, either communicative or instrumental, the end-point location of the pointing movement, either referring to the left, middle or right token and the addressee location, positioned either on the left or right side of the communicator. This allowed us to reveal several components of the triadic interaction between communicator, object and addressee and see whether kinematic adjustments are general to a communicative context or also specific to the location of a communicative partner. After quantifying the effects of the manipulations on the pointing kinematics we expect to be able to dissociate at which organizational level communicative and instrumental pointing actions relate to one another. Second, in a pilot functional MRI experiment we implemented a modified version of the cooperative game used in the kinematic experiment. The main challenge in using functional magnetic resonance imaging (fMRI) to study communicative pointing is to acquire a reliable measurement of brain activation while participants move their arm and interact with an addressee. Thus, we will focus on the feasibility of employing the experimental task used in the behavioural experiment in an fMRI setting. A direct comparison of activation evoked by either instrumental or communicative actions lies outside the scope of this pilot fMRI experiment. However, fMRI has been successfully employed in the past to study the neural substrate of planning instrumental actions based on different object features (Verhagen, Dijkerman, Grol & Toni, 2008). Therefore, we investigate whether the cortical circuitry responsible for planning a communicative pointing action is influenced by the type of stimulus feature that is relevant for communication.

2. Methods

2.1 Experiment 1: behavioral kinematic study

2.1.1 Participants

Fifteen right-handed participants (9 female, age 17-29) were recruited for the role of communicator. For every communicator, two additional participants were recruited to play the addressees. Nine of the communicators had played the role of addressee previously. Two communicators were excluded from analysis because they talked during the experiment or did not comply with the task requirements (resulting in more than 40% of incorrect trials in one or more conditions). All participants gave written informed consent according to the institutional guidelines of the local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands), and were either given financial compensation or credits towards completing a course requirement.

2.1.2 Experimental setup

Participants were seated at a table facing two vertically aligned computer screens (Fig. 1A). The top of the bottom screen (22 inch) was positioned at eye level of the communicator; the upper screen was suspended immediately above the bottom screen. An infrared position detection frame, the 'IR-frame' (TMDtouch iMac Zorro Masck 21.5 inch 16:9, USB connected), was positioned at a distance of 17 centimetres from the bottom screen. This frame detected the point of entrance of the index finger in the communicator's pointing movements.

Communicators were always positioned in the middle, immediately in front of the frame. Their heads were stabilized by means of a chin rest. In between pointing movements their right index finger rested on a capacitive sensor, the 'home-key'. The frame was positioned within comfortable pointing distance, which varied from 35 to 40 centimetres from the home-key, depending on the participant's arm length. Stimuli were presented on the bottom screen outside of reaching distance, at 52 to 57 centimetres from the home-key. Flanking the communicator were the two addressees (Fig. 1B). Their bodies were oriented towards the centre of the screens at an angle of 40 degrees with respect to the communicator. Head rotation and eye contact was prevented by means of the chin rest.

2.1.3 Experimental task

Three squares, each homogeneously coloured, labelled 'communicator tokens' (4x4 cm, Fig. 1A), were presented, always horizontally aligned, on the bottom. Three circles each with two colours, labelled 'addressee tokens', were presented on the upper screen occupying three of nine possible prefixed locations on a 3x3 grid. A 'goal token', which was an exact copy of one of the addressee tokens, was exclusively presented to the communicator; a wall blocked this token from the addressee's view (Fig. 1B). All tokens were presented against a black background.

The communicator and addressee were instructed that their common goal of each trial was to ensure that the addressee selected the correct addressee token. At the beginning of each trial, only the communicator was informed as to which addressee token was correct as indicated by the goal-token. The communicator could only convey this information by means of pointing to one of the three single coloured communicator tokens, left, middle or right. Consequentially, the communicator had to provide the addressee with the colour information that would allow the addressee to pick the correct addressee token. Consider the stimuli in Figure 2. The goal token denotes the red-blue addressee token. Since yellow can unequivocally refer to the yellow-magenta addressee token, he can either point out the red or the blue communicator token. In this situation, pointing to the blue communicator token would be most informative, since red is ambiguous and might refer to red-blue or red-green addressee token. Hence, pointing out the blue token would lead the addressee to select the correct red-blue addressee token. Each trial could be solved by mapping the

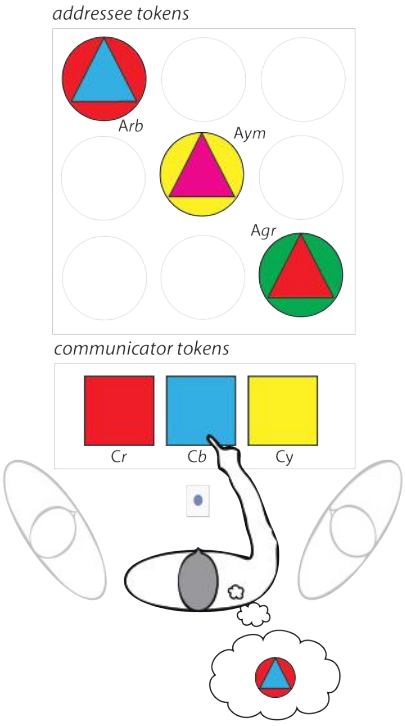


Fig. 2 Stimuli of experiment 1. Addressee tokens consisted of colored circles and triangles, here red-blue (Arb), yellow-magenta (Aym) and green-red (Agr). The communicator tokens consisted of single colored squares, here red (Cr), blue (Cb) and yellow (Cy). The goal-token was presented as privileged information to the communicator.

three communicator tokens to the addressee tokens in a similar way. The colours used in each trial were randomly selected from a set of six colours. Importantly, there was no consistent relationship between the colours and colour configurations to prevent the possibility of a fixed stimulus-response mapping. In 90% of the trials both of the colours of the goal-token recurred in the communicator tokens. In 10% of the trials, only one of the colours of the goal-token recurred in the communicator tokens.

2.1.4 Experimental design

The action manipulation consisted of how the addressee was provided with the colour information required to solve a trial. In the communicative condition, addressees were dependent on observing the pointing movement of the communicator towards one of the three communicator tokens. The communicator, knowing this dependency, was required to make a pointing movement for the addressee to see. In the instrumental condition, the addressee was visually provided with a 'match-token' (a copy of one of the addressee tokens) and no inference on the pointing movement was necessary. Importantly, the match-token presented was

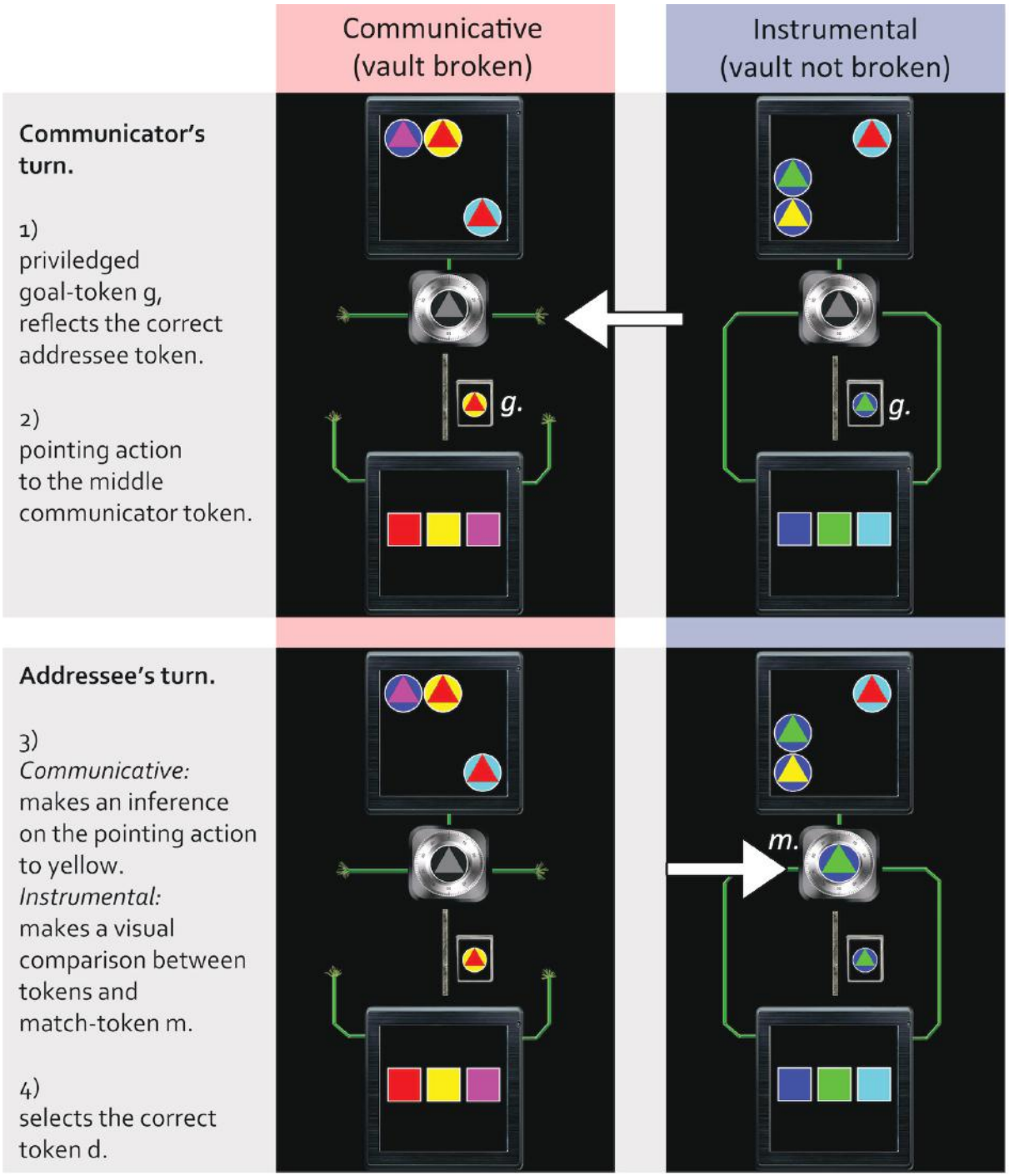


Fig. 3 The experimental manipulation of action. In both trial-examples displayed here, the game was played by the left addressee with the goal-figure (g) on the right side of the wall. Two phases of the trails are distinguished, the turn of the communicator consisted of stimuli presentation and pointing action, the turn of the addressee consisted of either an inference on the pointing movement or a visual comparison with the match-token (m) and the selection of an addressee token. Arrows indicate the crucial differences between the communicative and instrumental condition; the broken wire of the broken vault and the goal-token of the addressee.

conditional and depended on the pointing movement of the communicator. To help the participants understand and memorize the task rules, and to make the goal-oriented nature of the instrumental pointing movements self-evident, the game was embedded in a story where the joint goal of the participants was to open a highly secured vault by entering a correct sequence of addressee tokens. In the instrumental condition (Fig. 3), the computer system of the vault was working properly and the pointing movements could be detected by the infrared frame which made computer decoding possible. Hence, when the communicator pointed to a communicator token and returned to the home-key, the match token was

displayed on the upper screen for the addressee to see. The addressee simply had to confirm this token by selecting the identical addressee token. In the communicative condition however, the electronic wiring of the vault was broken (Fig. 3). Since now the detection mechanism of the vault did not work, the addressee was required to infer the correct addressee token from the pointing movement of the communicator towards one of the tokens. The addressee manipulation involved the spatial position of co-players that alternated in playing the game as addressee. The third factor, token, was the end location of the pointing movement manipulated by means of the three possible referent-tokens mentioned, either left, middle or right.

2.1.5 Experimental procedure

At the start of the experiment, communicators were familiarised with the IR-frame and performed two practice sessions of the game. In the first practice session, the participant that would be playing as communicator in the real experiment adopted the role of addressee. Right before the experiment started, participants switched positions and the communicator could engage in the second practice session as communicator. Communicators were instructed to move their finger through the IR-frame (so that pointing coordinates could be detected for online decoding) and return to the home-key after pointing. All communicative trials were started with a button-push from the addressee. Instrumental trials were started after a variable inter-trial-interval (1.4 to 3.4 seconds). After finishing the pointing movement and returning to the home-key, a cursor appeared next to the addressee tokens on the upper screen. Addressees could select one token via a mouse-click and were instructed to start moving after the cursor appeared. After an addressee token was selected, feedback was given in the form of a thick red (incorrect) or green (correct) line surrounding the stimuli.

The experiment consisted of three sessions of twenty minutes, 240 trials in blocks of ten. At the start of each trial one token was pseudo-randomly assigned as correct, which co-player was addressee was switched every block and the type of action changed every other block. Before each block of ten trials, participants were explicitly informed in which action and addressee conditions they would be for the upcoming block. Feedback on performance was given after each block: correct feedback if all ten trials were correct; incorrect feedback if an error was made in one or more of the ten trials.

2.1.6 Data-acquisition and analysis

Using an electromagnetic tracking system (LIBERTY, Polhemus) the position and orientation of four sensors were sampled at 250 Hz. One sensor was placed on the distal phalanx of the index finger on the right side just below the nail (labelled 'index'). Two of the sensors were placed on the hand, one on top of the second metacarpophalangeal joint (labelled 'hand-radial') and the other on top of the fifth metacarpophalangeal joint (labelled 'hand-ulnar). The last sensor was positioned on the distal side of the corpus radii (labelled 'wrist').

Kinematic data were analyzed using MATLAB (MathWorks, Natick, MA, USA). A detailed specification of the data processing and analysis procedures can be found in Verhagen et al. (2012); here we describe only the most important specifications. The spatial time series were low-pass filtered at 15 Hz using a sixth-order Butterworth filter. We took particular care to use a robust estimate of the onset and offset of the different movements (Schot, Brenner & Smeets, 2010), by combining both distance and velocity parameters. The pre-movement phase, the 'selection time' was defined as the time from goal-token presentation to release of the home-key. 'Movement time' was defined as the time from movement onset to end-point, not to be confused with the whole pointing movement duration. Movement onset was the point in time with the index finger at the largest distance from the tokens, just after the velocity of the index finger had increased above 0.1 m/s along the mid-sagittal plane while still in proximity of the home-key. The 'holding time' was the duration of the index finger being closest to a referent-token while decelerating to a velocity below 0.1 m/s. The 'return time' was defined as the time from the end-point back to the home-key, again decelerating below 0.1 m/s.

We describe the pointing movements using common kinematic parameters: trajectory lengths (TL), peak velocity (PV), time-to-peak velocity (tPV) as of movement onset and the relative time to PV (rtPV) as a fraction of movement time (Jeannerod, 1984). In addition, we describe the 'spatial dynamics' and 'temporal dynamics' of the pointing trajectories. To describe the spatial dynamics, 300 samples along the trajectory were isolated at equivalent spatial intervals considering three dimensions of Euclidean space. Averaging trajectories in space, regardless of time, is not trivial and requires an iterative algorithm to converge upon a solution (Liu & Todorov, 2007). To describe the temporal dynamics, a similar procedure was taken but now using 300 equivalent

time intervals as percentage of the trial-specific movement time. For every sample we calculated the coordinates on x, y and z axis: horizontal displacement (lateral), depth (forward) and altitude (height) respectively. We calculated a group average in order to offer a graphical representation of spatial and temporal dynamics, but statistical inference was drawn considering the inter-subject variability.

2.1.7 Statistical inference

Statistical inference was drawn using the SPSS 16.0 software package. We excluded incorrect trials that were caused by the communicator pointing to a wrong token (5.0%), or by the addressee selecting the wrong token (5.8 %). In addition, we excluded the first two trials and trials where no movement was made at all or too late (3.75%). Finally, trials where the main kinematic parameters deviated from the first or third quartile by more than three interquartile ranges (6.4%) where excluded. In total, 84.0% of all trials survived the exclusion criteria and entered further analysis.

All parameters were found to commit to the assumption of normality. Trials were averaged for each experimental condition, and the resulting means were entered into a univariate repeated-measures ANOVA testing for main effects and interaction effects between conditions within subjects. Statistical inference of the spatial and temporal dynamics is not possible by means of similar univariate or multivariate approaches, because of the inherent dependencies between data points neighbouring in space or time. Therefore, we chose to use non-parametric cluster-based permutation statistics (Maris & Oosterveld, 2007). We obtained an accurate Monte Carlo estimate of the true p-value by comparing the cluster statistics of interest to the distribution of cluster statistics calculated from 10,000 random permutations of the conditions. The trajectory variability was calculated on the basis of the spatial trajectories. For each of 300 points along the trajectory, the instantaneous gradient in three dimensions was used as a normal vector to describe a plane perpendicular to the movement direction. For all trajectories the intersection with this plane was calculated using linear interpolation resulting in a 2D distribution of points (oriented in 3D). The confidence ellipse was calculated on the basis of an Eigen-decomposition of the covariance matrix of this set of points and scaled using a $\chi 2$ distribution to match a 95% confidence interval in 2D (McIntyre et al., 1998). The critical parameter describing the variability of the trajectories along the movement is the area of the confidence ellipses.

2.2 Experiment 2: functional neuroimaging study

2.2.1 Participants

Five participants (2 females, age 19-23) were recruited to take part in the pilot-fMRI experiment. None of them had participated in experiment 1. For every participant, one co-player was recruited. One of the participants was excluded from the analysis because of excessive head movements during the movement phase of the task. All participants were right handed and had normal or corrected-to-normal vision. Participants gave written informed consent according to institutional guidelines of the local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands) and were given financial compensation.

2.2.2 Experimental setup

The general procedure and apparatus were similar to the ones used in experiment 1, except for a few adaptations required by the imaging procedure (Fig. 4). The participant lay in supine position with his or her head tilted forward, allowing a clear vision on the stimuli. The right elbow was fixated to the body in order to avoid movement of the upper arm. The stimuli were presented on an iPad (Apple iPad 4 32GB, USB connected) that was fixated onto the scanner bench just above the feet of the communicator. The iPad was held by a custom-made wooden frame with a little wooden panel that functioned as the wall obstructing the vision of the addressee on the goal-token. The iPad was connected as a secondary screen to a pc in the control room outside the scanner environment via an optical USB 1.0 connection, using the 'Air Display' application (Avatron Software Inc). This setup caused a temporal delay in displaying the stimuli on the iPad of ~100 milliseconds ±30 milliseconds. Only one co-player was present during the task, standing on the right side of the communicator. He or she stood in upright position besides the bed next to the scanner bore. Both players could see the addressee and communicator tokens on the iPad. The addressee's vision on the goal-token was occluded by the wall. The addressee could see the arm and hand of the communicator but they could not make eye contact because the head of

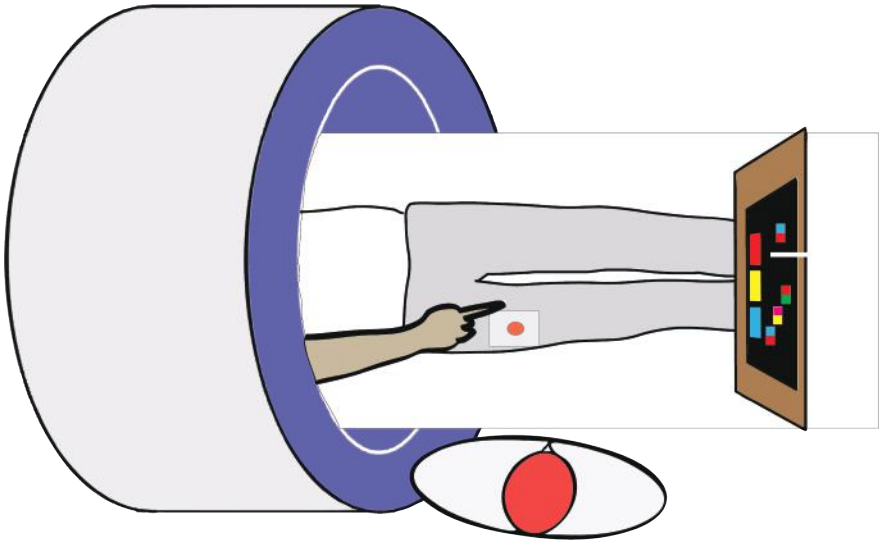


Fig. 4 The experimental setting of experiment 2. The communicator lies in the scanning device with the addressee standing on the right hand side. Both the communicator and the addressee had a clear vision on the iPad that was positioned right above the communicator's legs. Vision of the addressee on the goal-token was obstructed by a wall (white). The home-key was attached on the communicator's upper right leg.

the communicator was encapsulated by the bore of the scanner. The addressee was instructed to stay in upright position and not to bend over to make eye contact with the participant during the experiment. The communicator was instructed not to move the hand except during action execution. At all other times the right index finger rested on the home-key.

2.2.3 Experimental task and design

The experimental task was similar to that of experiment 1, with a few modifications in the stimuli appearance and task strategy. Again, all six shapes were presented to both the communicator and the addressee. A goal-token, identical to one of the addressee tokens, was exclusively seen by the communicator.

Similarly to experiment 1, the game required participants to map three single-coloured tokens to three dual-coloured tokens. In contrast, the mapping could be made by coding the stimuli in one of two ways: either based on overlapping token-colours, 'identity condition' or on overlapping token-locations, 'spatial condition'. In the identity condition (Fig. 5A) tokens never overlapped spatially; that is, they never occupied the same location on the 3x3 grid. The strategy to solve identity condition was the same as in experiment 1. The spatial condition required participants to pay attention to the overlapping positions of the tokens, since none of the addressee tokens overlapped in colour with those of the communicator (Fig. 5B). The rule given to the participants was that one token was always located on the exact same position on the grid, whereas the other two were minimally shifted. In the example displayed in Figure 5B, the best fit for the upper left token (Cul) is the upper left token of the addressee (Aul), since it is on the exact same location on the 3x3 grid. The best fit for the middle-right token of the communicator (Cmr) is the upper-right token of the addressee (Aur) since it is located one position higher on the grid. A similar rationale can be given for the bottom-middle token (Cbm); its best fit is the bottom-right token (Arb). Each trial could be solved in a similar manner. In order for the spatial condition to work, the communicator tokens in this experiment were located on a 3x3 grid, instead of being horizontally aligned like in experiment 1. This resulted in pointing actions made towards one of nine possible locations. Since the instrumental condition was part of this experiment, the validation of an instrumental pointing action with a cover story was unnecessary and not implemented in this experiment.

2.2.4 Experimental procedure

The experimental procedure was similar to the procedure in experiment 1. Communicator and addressee were jointly familiarized with the general procedure and the strategy of the game and participants practiced with the experimental

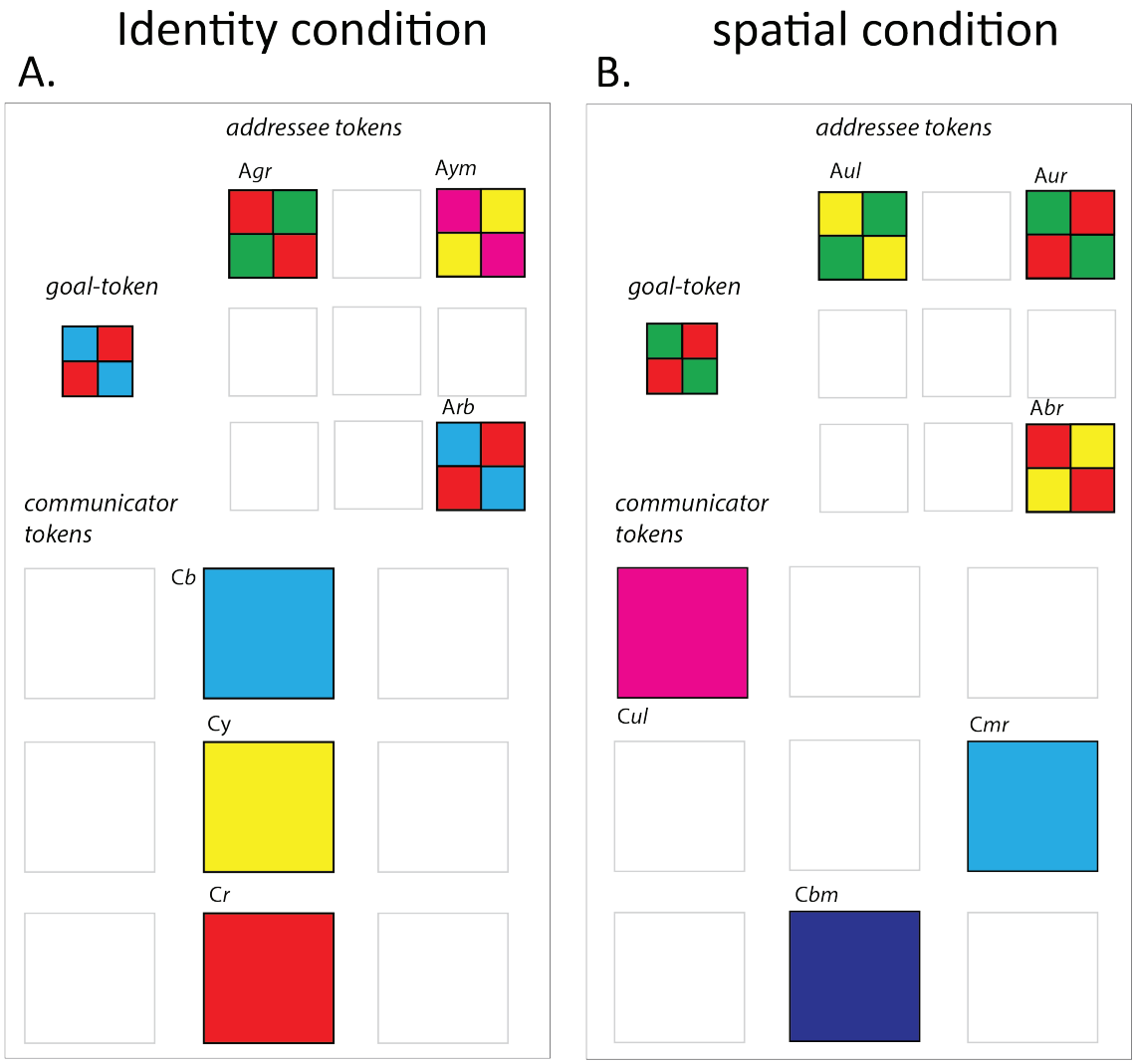


Fig. 5 Stimuli of experiment 2. Stimuli as displayed on the iPad in two conditions of coding. **A.** The identity condition with green-red (Agr), yellow-magenta (Aym) and red-blue (Arb) addressee tokens, red (Cr), yellow (Cy) and blue (Cb) communicator tokens. **B.** The spatial condition with upper-left (Aul), upper-right (Aur) and bottom-right (Abr) addressee tokens, upper-left (Cul), middle-right (Cmr) and bottom-middle (Cbm) communicator tokens.

setup in the scanning environment. At first the communicator was encouraged to verbalise his pointing intentions, thus shortening the learning period. The practice session continued without verbalising until performance was at ceiling level. At trial onset all stimuli were sequentially presented on the iPad starting with the communicator tokens, followed by those of the addressee and the goal-token (0.5 seconds in between). The rest of the trial sequence was equivalent to that of experiment 1. The communicator was free to point towards a token and return to the home-key. Following the pointing movement of the communicator, the addressee had to push one of three buttons that corresponded to the three addressee tokens. Afterwards, feedback was given (lasting one second) on the iPad in the form of the words 'correct' or 'incorrect'. There was a variable inter-trial interval that was randomly determined to be between three and seven seconds.

The experiment consisted of two sessions of 20 minutes in which participants completed on average 120 trials. Identity and spatial conditions were alternated in blocks of 12 trials. Before each block, participants were informed in which coding condition they would be for the upcoming trials. As in experiment 1, a trial was incorrect only if the addressee selected a shape other than the goal-token presented to the communicator. Unlike in experiment 1, no feedback on the performance of an entire block was given.

2.2.5 fMRI data acquisition and analysis

Communicators were scanned using a 3 Tesla Skyra scanner (Siemens, Erlangen, Germany). Blood-oxygen-level-dependent (BOLD) sensitive functional images were acquired using a single shot gradient echo planar imaging (EPI) sequence (TR/TE 2s, 42 transversal slices, interleaved acquisition, voxel size 3.3 x 3.3 x 3.0 mm3). Structural images were acquired using a MP-RAGE sequence (TR/TE 2s, voxel size 1 x 1 x 1 mm3).

The images were preprocessed and statistically analyzed using SPM8 (Statistical Parameteric Mapping; http://www.fil.ion.ucl.ac.uk/spm). Preprocessing of the functional scans included spatial realignment (rigid body transformations using sinc interpolation algorithm), slice-time correction, co-registration (of functional and anatomical images, after prior co-registration of both image types to an MNI template), reslicing (1.5 x 1.5 x 1.5mm), spatial normalization (to MNI space), and spatial smoothing (isotropic 8mm FWHM Gaussian kernel). Each anatomical image was segmented into three different tissue compartments (grey matter, white matter, cerebral spinal fluid) after the skull had been stripped (using the brain-extraction-tool implemented in FSL, FMRIB centre, Oxford University). Mean signals in the latter two compartments and head movement related parameters (translations and rotations to the mean as obtained during the spatial realignment, and their first derivatives) were entered as regressors of no interest in first-level fMRI analyses (Verhagen, Dijkerman, Grol & Toni, 2008). Furthermore, the data was high-pass filtered (cut-off 128 s), and temporal auto-correlation was modelled as a first-order auto-regressive process using a restricted maximum-likelihood estimation.

Three epochs were considered for the first level fMRI analyses: selection phase from stimulus presentation to movement onset (home-key release), movement phase from movement onset to return to the home-key, and a feedback phase after token selection by the addressee. Each epoch time-series was convolved with a canonical hemodynamic response function and used as a regressor of interest in the SPM multiple regression analysis. Following our hypothesized difference in neural activity, we modelled the identity and the spatial trials for selection and movement phases in separate models.

2.2.6 Statistical model and inference

Of main interest were the differences in planning and movement related neural activity between the identity and spatial conditions. To summarize the results over subjects, a fixed-effect model was constructed that included the two sessions of all four participants. We report the results of this fixed-effects analysis, with inferences drawn at the cluster level, corrected for multiple comparisons over the whole brain using family-wise error correction ($p < 0.05$). The commonalities in neural activity between the selection and the movement phase were tested by means of conjunction analyses (Nichols et al., 2005), which tested whether there was significant overlap in identity and spatial conditions. Differential activity between identity and spatial coding was tested by means of contrast analyses.

3. Results

3.1 Experiment 1: Kinematic study

Participants were engaged in the task and communicated effectively, solving on average 92.2% of the trials correctly. The remaining trials were excluded from further analysis either because of an error (5.0%) or due to an incorrect movement (2.8%). The number of error trials did not differ significantly between conditions. In the following sections, we will report on the effects observed in general movement parameters of the pointing actions made by the communicator. Several phases of a pointing movement were distinguished: the selection phase (the pre-movement time from stimuli presentation to initiation of the movement), the movement phase (the time from initiation of the movement to the end-point holding position), the holding phase (the time the end-point holding position was maintained) and the return phase (the time from end-point holding position to the return to the home-key). In addition, we will report on the effects observed in the end-point of the pointing movement and the temporal dynamics of the movement trajectory. A full overview of the results is given in Table 1 and 2.

3.1.1 Similarities in general movement parameters

The movement time (during which the hand was transported to the target) was not significantly different between the communicative and instrumental pointing movements ($F(1,12)$, $p = .085$, Fig. 6). The trajectories of both actions towards the three possible token locations appeared very similar at the scale of the whole movement (~50 cm, Fig. 7). Not surprisingly, in both communicative

Table 1. Statistical probabilities (F-statistics) of experimental manipulations on all parameters considered in this experiment (for the index-finger). We considered the three main effects of action, addressee and token. In addition, interaction effects of these factors are reported. Statistical inference was performed using univariate repeated measures ANOVAs, with 1 degree of freedom (2 degrees for all token interaction effects) and 12 (24) for the residual. In addition, all three axes were considered in a multivariate end-point analysis (MEP) with 3 degrees of freedom and 10 for the residual. F-statistics below 1 are indicated (<1) and are not reported with a p-value. ST = selection times; MT = (forward) movement times; HT = holding times; RT = return times; RTA = reaction times of addressees; TL = trajectory length of forward movement; PV = peak velocity of forward movement; TPV = time-to-peak velocity of forward movement ; RTPV = relative time-to-peak velocity of forward movement. EPx = end-point measured in the x-axis; EPy = end-point measured in the y-axis; iEPz = end-point measured in the z-axis.

Epoch	Anatomical region	Hemisphere	p-value	Cluster size	Local maximum
Selection					
Identity	Inferior Frontal Gyrus	Left	.000	500	-50, 30, 20
	Middle Frontal Gyrus	Left	.009	139	-32, 12, 48
	Inferior Parietal Sulcus (post.)	Left	.000	272	-30, -68, 40
	Inferior Occipital Gyrus	Right	.002	180	34, -90, -10
Spatial	Precentral Sulcus	Left	.000	461	-18 ,-8, 52
	Precentral Sulcus	Right	.000	357	30, -10, 50
	Middle Occipital Gyrus	Left	.000	221	-38, -80, 28
	Middle Occipital Gyrus	Right	.000	544	48, -80, 12
	Precuneus	Right	.000	556	10, -52, 46
	Precuneus	Right	.013	129	16, -56, 18
Movement					
Identity	Middle Cingulate Gyrus	Right (ext. Left)	.003	167	6, 8, 42
Spatial	Superior Parietal Lobule	Left	.002	184	-36, -54, 58
	Inferior Parietal Lobule	Right	.024	114	34, -44, 38
	Precuneus	Right (ext. Left)	.000	1264	2, -52, 58
	Middle Occipital Gyrus	Right	.000	347	40, -80, 8
	Cerebellum	Right	.015	126	28, -64, -24

and instrumental pointing movement, the finger converged in the direction of the indented referent token ($F(2,24)= 1629.66$, $p = .000$). Furthermore, the trajectory of the communicator's index finger was longer for movements towards tokens positioned further on the right in both action types ($F(1,12) = 99.14$, $p = .000$). This likely reflects the fact that the communicator used the right arm to point and the concomitant limitations on shoulder-joint rotation. In line with this, the return time was slower ($F(2,24) = 14.21$, $p = .000$) and the peak velocity was reached, both in absolute and relatively terms, at a later point in time for tokens positioned further on the right (absolute: $F(1,12) = 8.75$, $p = .001$; relative: $F(1,12) = 10.02$, $p = .001$). The addressees' reaction times were not significantly different between action conditions, but both the left and right addressees were slower with the token positioned further to the left ($F(2,24) = 17.19$, $p = .000$). Taken together, these effects suggest that the action execution of communicative pointing movement is comparable to the execution of instrumental pointing movements. However, important differences were observed in other general movement parameters.

3.1.2 Differences in general movement parameters

The selection time (time before movement onset) was longer in the communicative condition than in the instrumental condition ($F(1,12) = 12.14$, $p = .005$, Fig. 6). Importantly, the holding time (time the finger was held near the token) was almost twice as long in the communicative condition than in the instrumental condition ($F(1,12) = 16.49$, $p = .002$, Fig. 6). Interestingly, this effect was strongest for the middle token (interaction of action x token, $F(2,24) = 3.77$, $p = .038$). This token was flanked by the two other tokens, and therefore the most difficult to dissociate. The return time (during which the hand was transported back to the home-key) was also longer in the communicative condition than

Table 2. Means of parameters considered in this experiment. We considered the three main effects of action, addressee and token. A. Means per condition that revealed a significant main effect (table 1). B. Means per condition that revealed a significant interaction effect (table 1). No means of the multivariate analyses are included here. ST = selection times; MT = (forward) movement times; HT = holding times; RT = return times; RTA = reaction times of addressees; TL = trajectory length of forward movement; TPV = time-to-peak velocity of forward movement ; RTPV = relative time-to-peak velocity of forward movement. EPx = end-point measured in the x-axis; EPy = end-point measured in the y-axis; iEPz = end-point measured in the z-axis.

	action	**addressee**	**token**	**action x addressee**	**action x token**	**addressee x token**	**action x A. x T.**
ST	0.005 (12.14)	0.179 (2.00)	0.105 (2.49)	(<1)	0.072 (2.93)	0.277 (1.35)	(<1)
MT	0.085 (3.61)	0.179 (1.86)	0.322 (1.19)	0.107 (3.03)	(<1)	0.051 (3.37)	0.206 (1.69)
HT	0.002 (16.49)	(<1)	0.006 (6.28)	0.088 (3.46)	0.038 (3.77)	0.123 (2.29)	(<1)
RT	0.004 (12.89)	0.138 (2.53)	0.000 (14.21)	0.322 (1.02)	0.229 (1.57)	0.368 (1.04)	(<1)
RTA	(<1)	(<1)	0.000 (17.19)	(<1)	0.078 (2.84)	(<1)	(<1)
TL	(<1)	(<1)	0.000 (99.14)	(<1)	(<1)	(<1)	(<1)
PV	0.0278 (1.29)	(<1)	(<1)	(<1)	(<1)	(<1)	(<1)
TPV	0.262 (1.39)	0.088 (3.45)	0.001 (8.75)	(<1)	(<1)	0.243 (1.50)	0.232 (1.55)
RTPV	0.114 (1.91)	0.008 (9.98)	0.001 (10.02)	(<1)	(<1)	0.168 (1.92)	(<1)
EPx	0.150 (2.361)	0.002 (14.82)	0.000 (1629.66)	0.024 (6.69)	0.001 (10.29)	(<1)	(<1)
EPy	(<1)	(<1)	0.000 (237.49)	(<1)	0.112 (2.40)	0.179 (1.74)	(<1)
EPz	0.088 (3.44)	(<1)	0.000 (56.57)	(<1)	(<1)	0.329 (1.17)	0.156 (2.00)

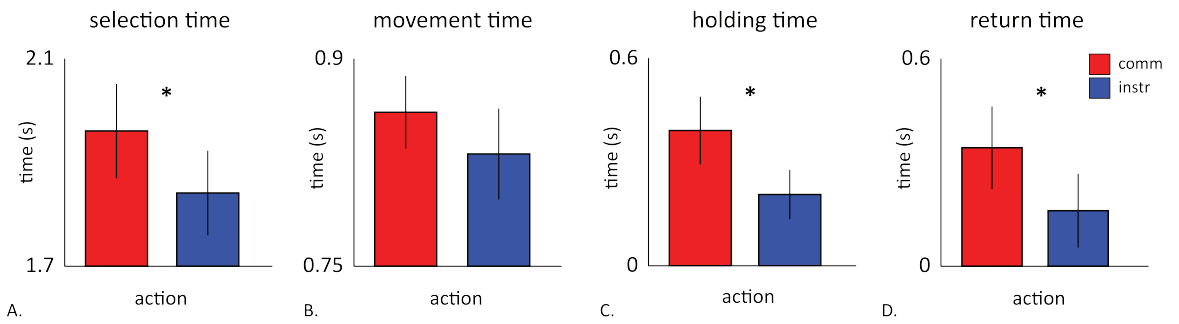


Fig. 6 General movement parameters of experiment 1. Graphical representation of the mean values and standard errors of the reaction times for the communicative and instrumental conditions. ST = selection time; MT = movement time; HT = holding time; RT = return time. (* = significant at p = .05).

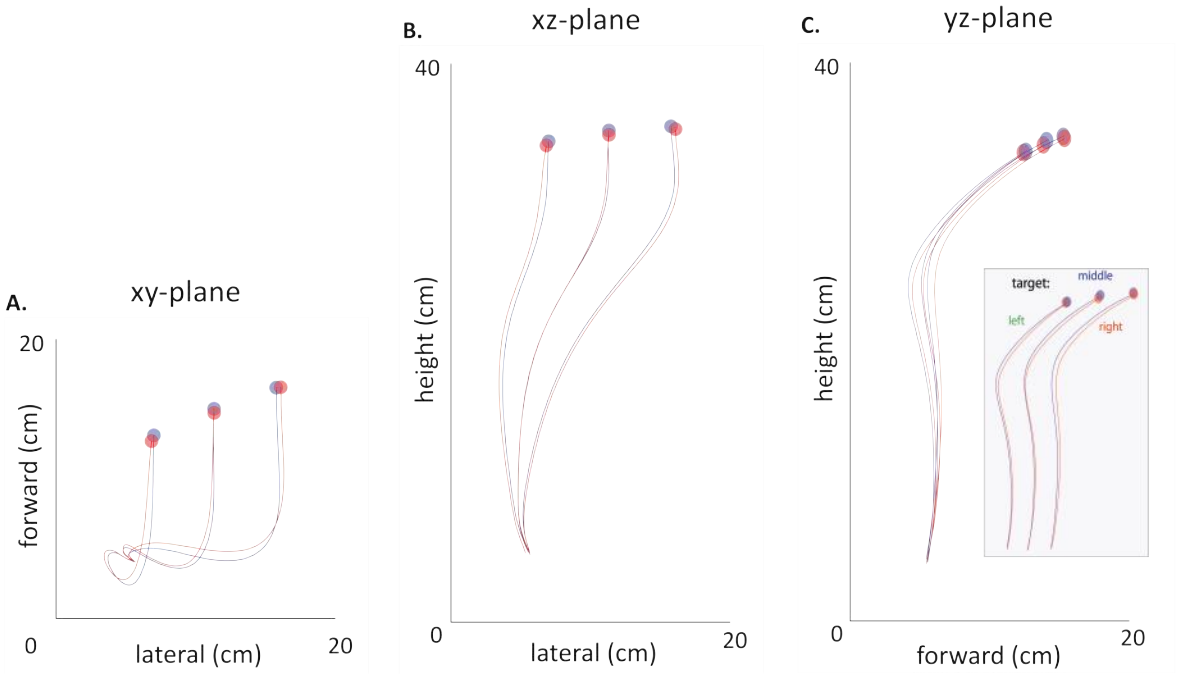


Fig. 7 The spatial dynamics of movement trajectories. The mean trajectories for the communicative (red) and instrumental (blue) conditions are illustrated in the **A.** xy-plane. **B.** xz-plane and **C.** yz-plane .

in the instrumental condition ($F(1,12) = 12.89$, $p = .004$, Fig. 6), indicative of either a generally slower movement or an after-effect of the deceleration in the holding phase. Hence, these adjustments could reflect an ostensive tag in the form of general slowness of movement execution or a prolonged holding time.

3.1.3 Pointing end positions

Following the design of our setup with the tokens configured side-by-side horizontally we focused the analysis of the end-point on their lateralization (along the x-axis).We observed two effects that show that planning communicative actions is different from planning instrumental actions. First, the positions of the index finger at the end of the pointing movements discriminated more strongly between the three tokens in the communicative condition than in the instrumental condition. Namely, the end-points of the communicative pointing movements on the two outer two tokens were more outwards of the centre (action x token, $F(2,24) = 10.29$, $p = .001$, Figure 8). Second, pointing end positions when interacting with the left addressee were significantly different from pointing movements made when interacting with the right addressee. Namely, end-points converged towards the left when interacting with left addressees and towards the right when interacting with right addressees ($F(2,24) = 14.83$, $p = .002$, Figure 9). Importantly, this effect on the end-point was driven by the communicative condition: pointing actions ended significantly more towards the location of the addressee in the communicative condition than in the instrumental condition (action x addressee, $F(1,12) = 6.69$, $p = .024$). These effects

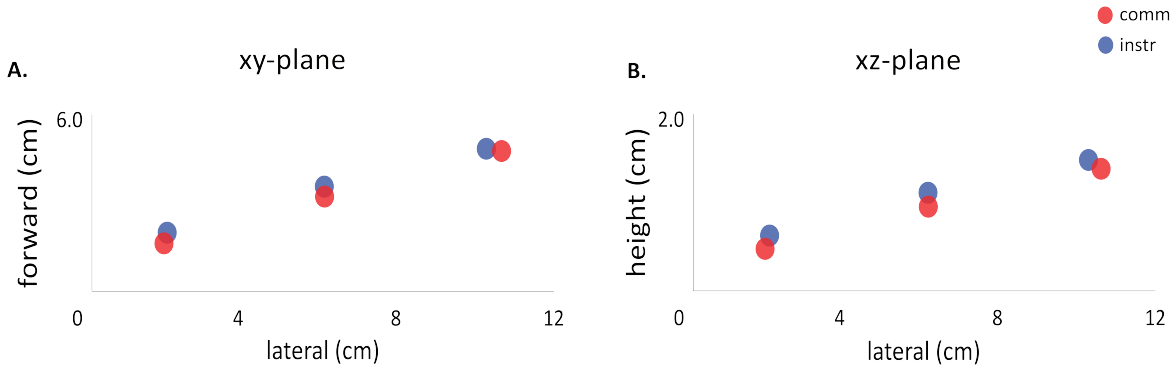


Fig. 8 Averaged end-points for communicative and instrumental conditions in the xy and the xz plane.

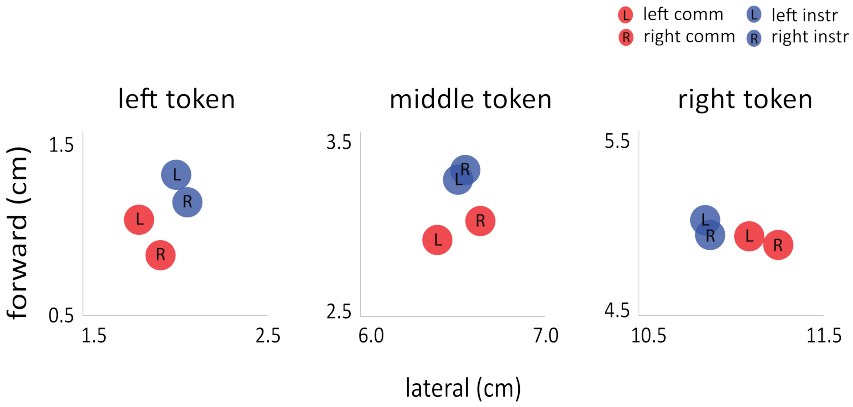


Fig. 9 Mean end-points for communicative and instrumental conditions per addressee in the xy-plane.

suggest that, when communicative intent is present, the targeted end position of the movement is adjusted taking into account both the characteristics of the message (emphasizing the communicatively relevant distinction between the referent-objects), and the characteristics of the audience (adjusting to the location of the addressee).

3.1.4 Pointing trajectory dynamics

Specific effects of action and addressee were observed in the trajectory at the scale of the token (~1 cm). First, we observed that not only the end positions but also the trajectory of pointing movements towards the left and right tokens developed more laterally in the communicative condition than in the instrumental condition (Figure 10). Critically, these adjustments arose already very early during movement execution, from 10% of movement duration and pertained until completion of the whole movement ($p = .003$). Moreover, we found a main effect of addressee, with the lateral development of pointing actions for the left addressee converging more towards the left, and those for the right addressee vice versa, from 83% of movement duration onwards ($p = .045$).

In addition to these effects in the lateral direction (in line with the token configuration), we observed a main effect of action and addressee on the forward development of pointing movements (Figure 11). First, pointing movements in the communicative condition developed more forward than those in the instrumental condition from 52% of movement duration onward ($p = .004$). Second, trajectories for the right addressee developed more forward than those for the left addressee, but only in the communicative condition, from 33% until 53% of the movements ($p = .049$). Congruently, the peak-velocity of the forward movement was reached relatively later when pointing for the right addressee than for the left addressee ($F(1,12) = 9.98$, $p = .008$), irrespective of the presence of communicative intent.

Finally, there was a higher variability in communicative pointing trajectories than in instrumental pointing trajectories reaching significance at two phases of the movement, between 43% and 61% ($p = .006$) and 67% and 100% of movement duration ($p = .044$, Figure 12). Taken together, these observations suggest that the communicative intent and communicative context of a pointing movement, including specifics

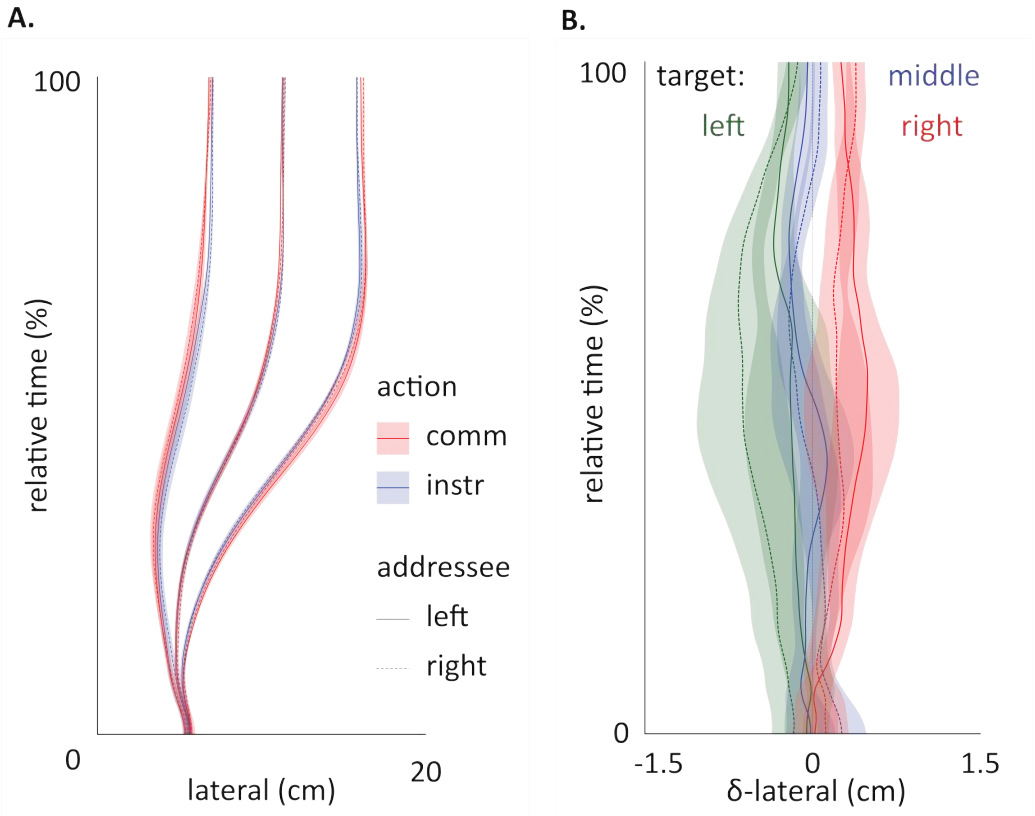


Fig. 10 The lateral temporal dynamics of movement trajectories. **A.** The mean lateral temporal development for the communicative and instrumental, and the left and right addressee conditions. **B.** The mean contrasted temporal development for the communicative and instrumental conditions per token.

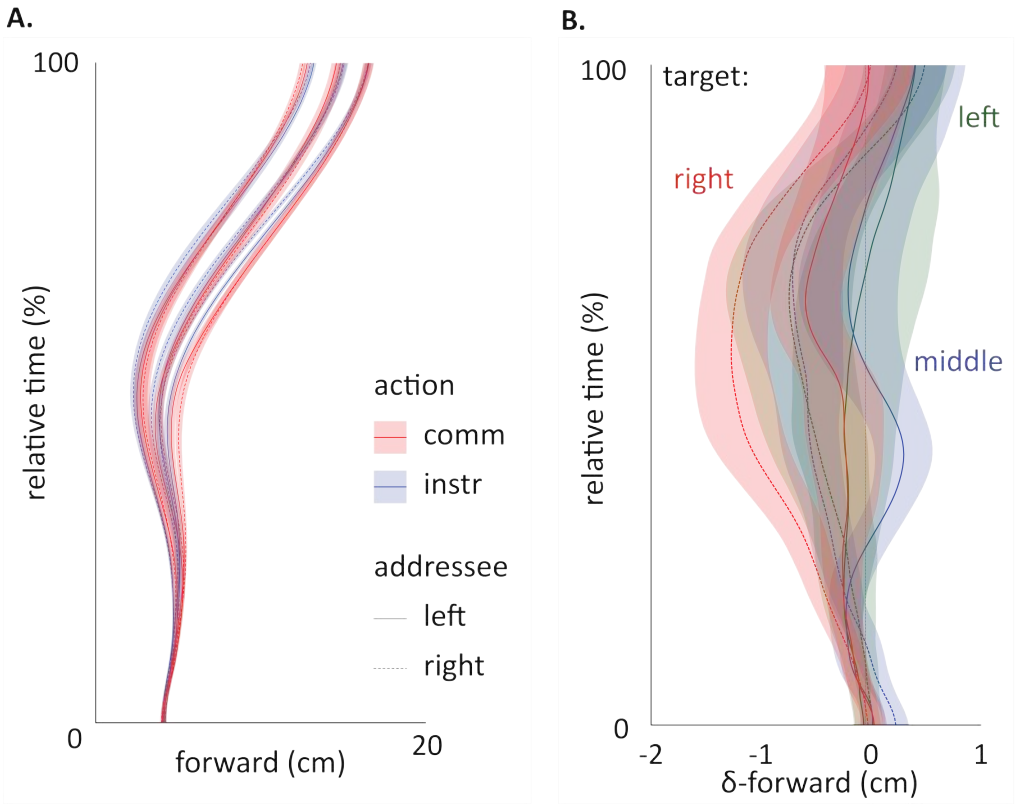


Fig. 11 The forward temporal dynamics of movement trajectories. **A.** The mean forward temporal development for the communicative and instrumental, and the left and right addressee conditions. **B.** The mean contrasted forward temporal development for the communicative and instrumental conditions per token.

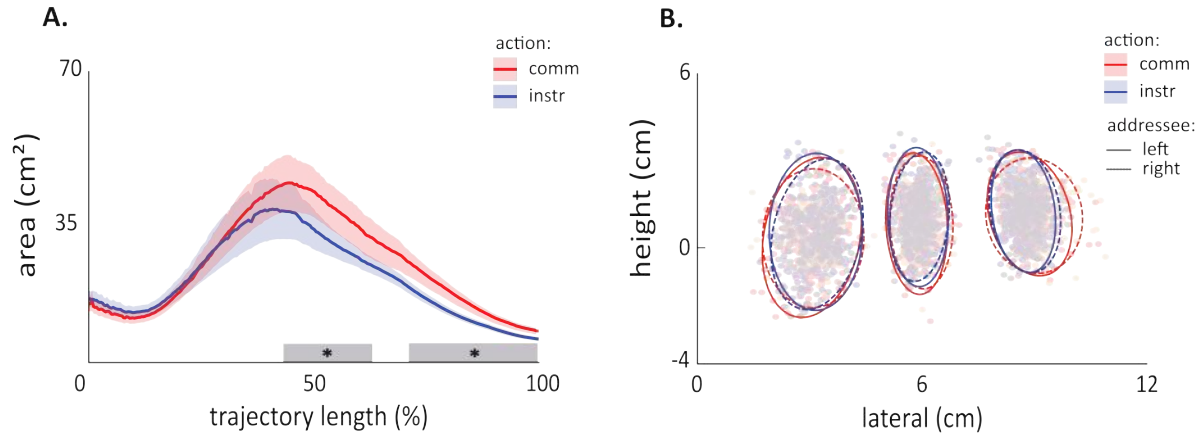


Fig. 12 The variability in spatial dynamics of trajectory and end-points. **A.** Shows the trajectory dynamics of the area of the 95% confidence ellipses for the communicative and instrumental pointing conditions. **B.** Shows the end-points and area of the 95% confidence ellipses for the communicative and instrumental pointing conditions for left and right addressee.

of the object and addressee locations, not only affect the selection of the pointing end positions, but also pervasively penetrates the planning of the movement as a whole.

3.2 Experiment 2: functional neuroimaging study

Participants solved on average 88.6% of the trials, which showed participants were actively engaged in the task and that the addressees were able to infer which token the communicator was pointing at. The duration of the selection phase was on average 4.25 seconds (SE = 1.88 s), and was slightly longer in the identity condition (M = 4.56 s, SE = 2.33 s) than in the spatial condition (M = 3.93 s, SE = 1.42 s). The duration of the movement phase was on average 2.32 seconds (SE = 0.69 s), and was similar in the identity (M = 2.39 s, SE = 0.85 s) and spatial conditions (M = 2.25 s, SE = 0.53 s).

3.2.1 BOLD amplitude during token selection

Selecting and planning a pointing movement elicited bilateral activation in occipital, parietal and frontal brain areas in both coding conditions (Figure 13A). On top of these general effects a differential neural activation pattern was observed for the identity and spatial conditions (Figure 13B-C, Table 3). Compared to the spatial condition, the identity condition elicited stronger neural activation in a left lateralized parietal-frontal network, encompassing both the inferior posterior parietal sulcus (Figure 13G) and the inferior and middle frontal gyri (Figure 13F), and in one area in the right hemisphere, the right inferior occipital gyrus (Figure 13D). The left middle frontal gyrus showed a deactivation in all other epochs of the task (Figure 13H): the selection in the spatial condition, the movement phase, and in the feedback phase of both the correct and incorrect trials. The other differentially activated brain areas showed baseline-activation in the movement phase and a strong reactivation in especially the incorrect feedback phase.

The spatial condition elicited stronger activation in a bilateral network including the pre-central sulci (Figure 13E), the middle occipital gyri (Figure 13D), and the right precuneus (Figure 13B). The left precentral sulci and the middle occipital gyri were strongly reactivated during the feedback phase, especially in incorrect trials (Figure 13I). Left and right precentral gyri showed a strong activation during the subsequent movement phase (Figure 14A). The right precuneus showed a similar strong activation during movement, which was specific for the spatial condition (Figure 14D-G).

3.2.2 BOLD amplitude during pointing movement

The motor cortex and cerebellum were strongly activated during movement execution in both the identity and spatial conditions, predominantly in the left hemisphere (figure 14A-B). Activation in the middle cingulate gyrus was stronger in the identity condition than in the spatial condition (Figure 14C), predominantly in the right hemisphere, but extending into the left. This area showed a deactivation during feedback (Figure 14F). A large parietal network, consisting of superior parietal left, inferior parietal right and the precuneus was strongly activated during the spatial condition (Figure 14B). In addition, the

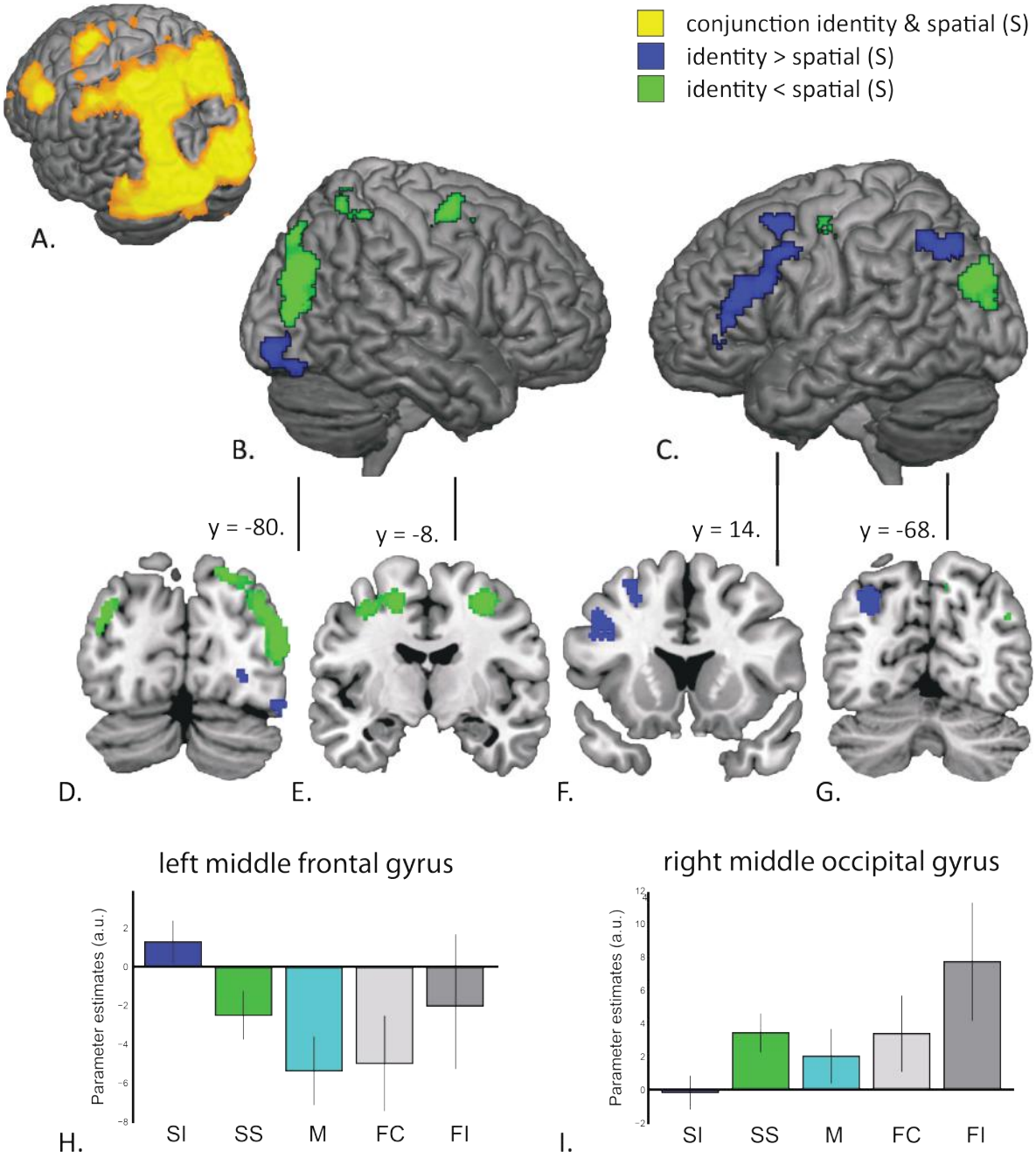


Fig. 13 Brain regions involved in the selection phase. Clusters (> 90 voxels) shown were significant at the voxel level at an uncorrected threshold of $p < 0.001$. Contrast maps were implicitly masked at $p < 0.05$, whole brain FWE corrected: identity > spatial (blue) with the activation of selection in identity trials, spatial > identity (green) with the activation of selection in spatial trials. Bar plots of parameter estimates are given for the peak-voxel reported in table 3, for the selection identity (SI), selection spatial (SS), movement (M, both identity and spatial), correct feedback (CF) and incorrect feedback (IF) epochs. **A.** The conjunction shows an occipital-frontal-parietal network of basic neural activation in identity and spatial trials. **B.** Right lateralized differential activation. **C.** Left lateralized differential activation. **D.** Coronal view of middle occipital areas (green) and the right inferior occipital area (blue). **E.** Coronal view of the precentral sulci. **F.** Coronal view of left middle and inferior frontal areas. **G.** Coronal view of the left inferior parietal sulcus. **H.** Parameter estimates of the left middle frontal gyrus in the five epochs. **I.** Parameter estimates of the right middle occipital gyrus in the five epochs.

right middle occipital gyrus and the cerebellum in the right hemisphere showed strong activation in the spatial condition. All aforementioned areas, with the exception of the precuneus, showed a strong reactivation following feedback, especially in incorrect trials. Activation of the right precuneus, as well as the left superior parietal gyrus, occurred specifically in the movement phase in spatial trials (Fig. 14G). Activation of the middle occipital gyrus was more pronounced in the selection phase of spatial trials.

Table 3. differential neural activation evoked during identity and spatial trials. MNI-stereotactic coordinates of the local maxima of regions showing an effect during selection and movement phases. Cluster size is given in number of voxels. The p-values represent the FWE cluster-level corrected values.

A		Action		Addressee		Token		
		Communicative	Instrumental	Left	Right	Left	Middle	Right
ST	ms	1961 (91)	1841 (82)					
MT	ms	862 (26)	831 (33)					
HT	ms	392 (98)	206 (72)			258 (86)	294 (92)	276 (83)
RT	ms	736 (30)	690 (27)			725 (26)	716 (28)	203 (28)
RTA	ms	1316 (33)	1307 (42)			1342 (35)	1312 (42)	1283 (34)
TL	cm					39.6 (1.2)	41.4 (1.1)	43.9 (1.2)
TPV	cm/s					112 (19)	90 (16)	88 (16)
RTPV	%			11.5 (2.1)	12.1 (2.1)	10.6 (2.0)	11.4 (2.1)	14.0 (2.3)
EPx	mm			63.14 (0.90)	64.50 (0.96)	19.48 (1.55)	63.26 (0.87)	109.37 (0.92)
EPy	mm					13.66 (6.62)	33.99 (6.02)	51.30 (6.10)
EPz	mm					4.53 (3.34)	11.19 (2.60)	15.22 (2.49)

B		Action x addressee				Action x token						Addressee x token					
		Communicative		Instrumental		Communicative			Instrumental			Left A			Right A		
		Left A	Right A	Left A	Right A	Left	Middle	Right	Left	Mid-dle	Right	Left	Middle	Right	Left	Middle	Right
MT	ms											848 (34)	868 (28)	838 (33)	841 (26)	840 (29)	837 (29)
HT	ms					364 (96)	413 (103)	401 (98)	198 (68)	206 (72)	217 (76)						
EPx	*mm*	62.55 (0.95)	65.15 (0.96)	63.40 (0.88)	64.18 (0.93)	18.64 (1.66)	63.45 (0.95)	110.90 (1.08)	20.06 (1.62)	63.27 (0.97)	107.85 (0.79)						

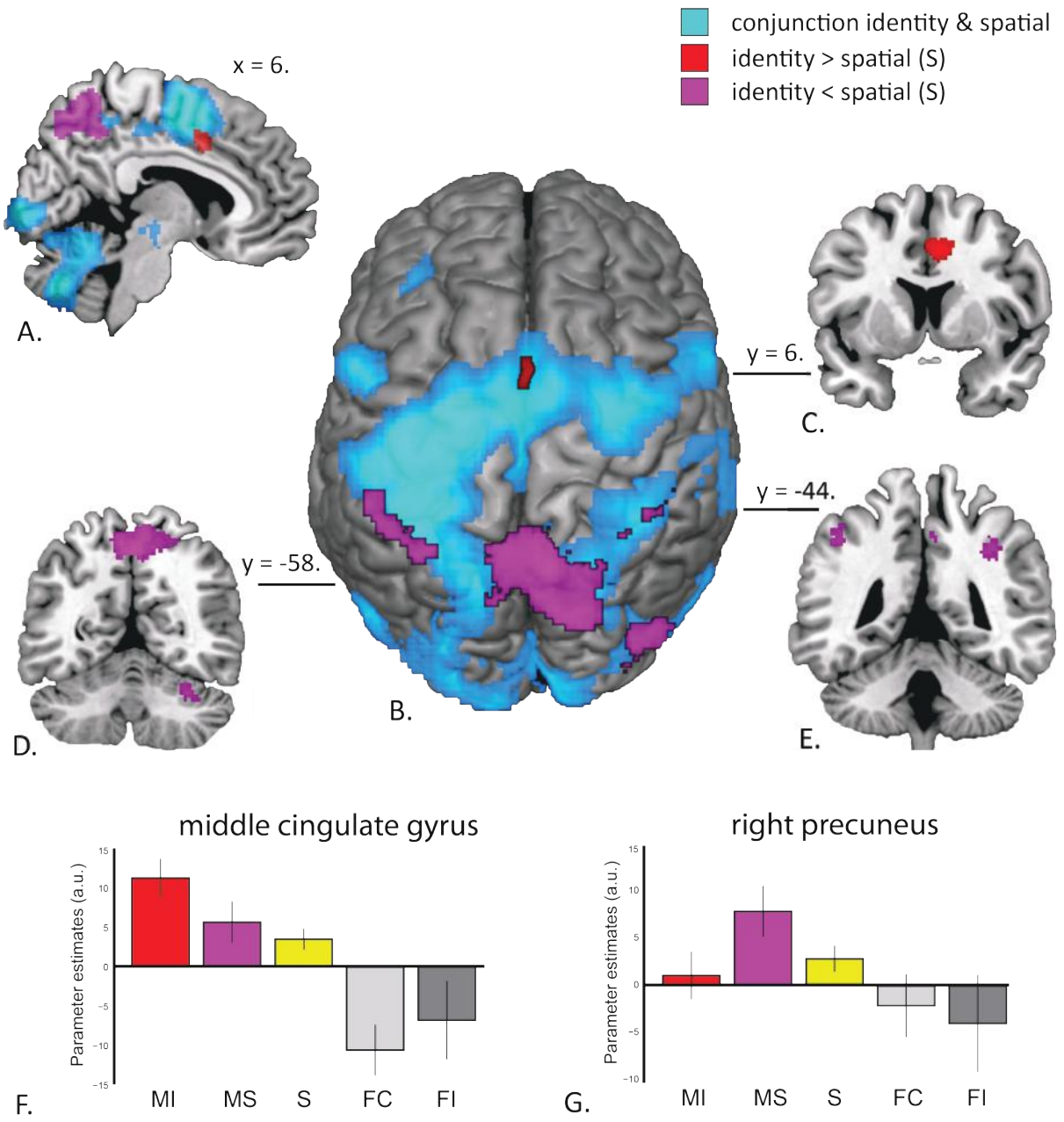


Fig. 14 Brain regions involved in the movement phase. Clusters (> 90 voxels) shown were significant at the voxel level at an uncorrected threshold of $p < 0.001$. Contrast maps were implicitly masked at $p < 0.05$, whole brain FWE corrected: identity > spatial (red) with the activation of movement in identity trials, spatial > identity (purple) with the activation of movement in spatial trials. Bar plots of parameter estimates are given for the peak-voxel reported in table 3, for the movement identity (MI), movement spatial (MS), selection (S, both identity and spatial), correct feedback (CF) and incorrect feedback (IF) epochs. **A.** Sagittal view showing conjunction and contrast maps. The conjunction shows basic neural activation in a parietal-occipital network and cerebellum in both identity and spatial trials, with specific activity in precuneus (purple) and cingulate gyrus (red). **B.** Top view of conjunction and contrast maps. **C.** Coronal view of the middle cingulate gyrus. **D.** Coronal view of the right precuneus. **E.** Coronal view of superior and inferior parietal areas. **F.** Parameter estimates of the left middle cingulate gyrus. **G.** Parameter estimates of the right precuneus.

4. Discussion

We investigated how human communicative abilities interface with the planning and execution of pointing movements by means of two interrelated approaches. In a behavioural experiment, we investigated the kinematic characteristics of pointing actions to learn to what extent a communicative intention can influence parameters of action planning. Whereas an influential account of human communicative abilities suggest there is no substantial difference in the planning of communicative and instrumental actions (Rizzolatti & Arbib, 1998), we observed clear differences in

action execution. In line with our first hypothesis (Carruthers, 2002; Goodale & Milner, 1992), we found evidence suggesting that communicative abilities can top-down instruct an instrumental action plan to mark the communicative intent for the addressee. However, the influence of computations supporting communication was not limited to this high-level instruction, but pervasively interacted with low-level parameters of action planning, which was in line with our second hypothesis (Baker, Saxe & Tenenbaum, 2009; Blokpoel et al., 2012; Brennan & Clark, 1996; van Rooij et al., 2011). Spatial and temporal movement analyses revealed that features relevant for communication, namely the presence of communicative intent (action manipulation), the location of the referent object (token manipulation), and the location of the addressee (addressee manipulation), are tightly integrated in the planning of pointing movements. Below we will further elaborate on our findings and their relation to different accounts of the interaction between communicative and instrumental faculties in the human brain. In a complementary pilot fMRI experiment we aimed to develop and test an experimental setup that allows investigation of the neural correlates of planning communicative pointing movements. Ultimately, the goal of this experiment is to reveal the neural mechanisms underlying the instrumental and communicative nature of pointing, but also to investigate whether these neural mechanisms are differentially recruited based on the coding of the communicative message. In this report only the latter issue is discussed. The results reveal a neural network that is used for planning communicative actions and is lateralized depending on the coding of the message, either referring to the colour property, or the location property of an object.

4.1 Kinematics of communicative pointing

In the kinematic experiment the perceptual input, the required motor output, the feedback on performance and the collaborative social setting were all kept constant. Critically, the presence, locations and engagement of addressees in the task, and the dependencies between the communicator and the addressees were also matched between conditions. This ruled out the possibility that the adjustments in kinematic parameters were caused by a low-level perceptual differences. Furthermore, addressee reaction times and error-rates were not found to be different between conditions, which confirmed that the two types of pointing actions were also matched for their social consequences. The trajectory length and return time were longer, and the peak velocity was reached later, for tokens positioned further on the right. Together with the finding that the movement time was not significantly different between the conditions, this indicates that the low-level motor control of communicative and instrumental pointing movements is comparable at the scale of the whole movement (~30 cm). This confirms the assumption that both actions are implemented by the same movement apparatus and motor system. However, important effects on a smaller scale (~1 cm) show that planning a communicative action does not merely entail planning a goal-directed instrumental action.

4.1.1 Marking communicative intent

Pointing with communicative intent led to several general adaptations in movement planning and execution compared to instrumental pointing. First, we observed a longer selection time (the time from stimuli presentation to movement initiation) for communicative pointing actions than for instrumental pointing actions. This either reflects an increased cognitive load in planning the communicative pointing action, or a conscious effort to provide the addressee with sufficient time to process and prepare the problem presented. Second, we observed a prolonged holding time (near the referent-token) in the communicative condition. Note that the instrumental consequence (the visual information presented to the addressee) of the action could not have driven this difference, since it occurred only upon arrival on the home-key. Such a prolonged holding time could serve as an ostensive tag to aid the addressee. Essentially, it both marks and extends the time the addressee has to collect the sensory evidence needed to infer the intended message. These findings are in line with previous observations (Peeters et al., 2013). Interestingly, the enhanced holding time was most pronounced for the middle token, the one most difficult to dissociate for the addressee. This highlights that the ostensive tag can be adaptively modified depending on the difficulty of the message. A study by Newman-Norlund and colleagues (2009) showed that communicators rested longer on the communicatively relevant part of an action sequence when communicating with a presumed child addressee than with an adult. Similarly to the effect seen in the current study, they were using time as a tool to place emphasis on target information when communication was thought to

be more difficult. Together these findings suggest that the communicative intention of an action can influence motor control on an instructional level, by marking the communicative relevant features of the action but without being specific to the content of the message. These findings seem to support our first hypothesis, namely that a higher-order conceptual representation (the communicative intention) can top-down instruct instrumental action planning to mark and enhance the communicative intent to ensure successful communication.

4.1.2 Conveying communicative content

Our second hypothesis suggests a further integration between the communicative and motor system, namely that actions are tailored to convey communicative content by having access to several control parameters of the motor plan. Recent evidence supports this notion. Cleret de Langavant and colleagues (2011) found that there was a difference in the direction of the confidence ellipsoid of the end-point variability between communicative and non-communicative pointing movements. In the current study we were not able to replicate this result. This could be a consequence of the relatively strict constraints under which the effect on end-point variability was originally observed: exclusively for the object positioned in the middle. For pointing action towards four objects surrounding the middle one, the ellipse orientations did not reveal significant differences between conditions and were, moreover, not consequent in their direction. We did however observe several effects that could dissociate at what level communicative computations could interact with control parameters of motor planning.

We observed adjustments of the targeted end point of pointing movements specific to the communicative content and addressee location. First, communicators emphasised communicatively relevant features of the intended token: pointing more outward either left or right for the respective tokens in the communicative condition than in the instrumental condition. This reflects a specific end-point adjustment that helped to convey the content of the message, namely that the referent-token with its correct colour information and not a neighbour token is intended. Second, we observed shifted end points towards the location of the addressee in the communicative condition, which indicates that the viewpoint of the addressee was consistently considered when making communicative pointing actions towards all three of the tokens. These findings show that the motor control of communicative actions is adaptive to several features of the communicative goal, incorporating information about the message content and the viewpoint of the addressee.

This conclusion is further supported by effects seen in the development of movement trajectories that preceded the end-point location and, again, were specific to the communicative content and the location of the addressee. First, as compared to instrumental pointing movements, communicative pointing movements developed more laterally already from 10% of movement duration onwards, and more forward from 52% onwards. This indicates that the presence of communicative intent pervasively affected action planning from very early on, right after movement initiation. These adjustments were specific to the message content because it could potentially help the addressee in discriminating the intended referent-token from distracting tokens even before reaching the end point. We also found a higher variability in communicative pointing trajectories in the middle and final parts of the movement. These observations are in line with previous observations that show that communicative actions are planned as a function of the partner's recognition. Sartori and colleagues (2009) and Becchio and colleagues (2008) demonstrated that different social intentions, either lifting or moving an object or showing or passing it to a partner, affected the trajectories of grasping movements from early on (more trajectory deviation and earlier grasp preparation). Second, pointing trajectories converged more towards the location of the addressee, halfway through the forward development from 33% until 53% of movement duration, and at the end of the lateral development, from 83% onwards. As can be seen in Figure 8B, it seems likely that the location of the addressee is represented from early on but only reached significance later. Compared to the instrumental movements, the lateral trajectories of the communicative pointing movements developed contralaterally to the spatial location of the addressee in the middle of the trajectory, and converged towards the location of the addressee only at the very end. This suggests that communicative pointing movements were shaped by the spatial location of the addressee throughout the whole movement.

These findings show that not only the targeted end location of the movement is adjusted, but also the trajectory of the movement is influenced by a specific communicative intention from early on. Moreover, we observed adjustments to the location of the addressee, which shows that specific spatial information of the audience viewpoint can

be integrated in the action plan. Altogether, these findings show that the conceptual representation of a communicative intention, and information of several features of the communicative goal, can penetrate the motor system and not only adjust the targeted end point, but pervasively affect action execution from early on.

4.2 Neural correlates of communicative pointing

We tested the feasibility of real-life communicative pointing interactions in the magnetic resonance imaging environment. We implemented a novel setup which allowed pointing movements to be executed in the magnetic resonance environment. We were able to present stimuli on an iPad screen so that participants could directly point at objects in front of them. Importantly, this method did not influence the homogeneity of the magnetic field in the bore of the scanner nor signal-to-noise-ratio of the echo-planar-imaging BOLD-signal measurements. Also, in spite of inevitable ongoing head movements related to executing a pointing movement with the finger, hand and lower arm, we were successful in retrieving neural images of considerable quality. Analysis showed promising results for a subsequent fMRI experiment using a similar setup and, additionally, incorporating an instrumental pointing action like the one assessed in the kinematic experiment.

We investigated neural activation patterns of a communicator that occurred during communicative pointing actions where the referent of the movement was coded by making use of either the colour or location features of the tokens. Whereas selecting tokens based on colour properties activated a left lateralized middle lateral prefrontal and posterior parietal circuit, selecting tokens based on spatial properties activated the pre-central gyri, middle occipital areas, and a right lateralized precuneus node. Strikingly, both networks were re-activated following feedback of incorrect selections by the co-player. Given the overlap in sensory inputs and motor outputs between conditions, the recruitment of these networks was likely to be elicited by a difference in stimulus encoding in accordance with the task strategy. The lateral prefrontal activation in the identity condition suggests encoding in a more abstract manner than encoding in the spatial condition. The lateral prefrontal cortex is known to be critical for the conditional selection of behaviourally relevant stimuli (Rushworth et al., 2005), independent of sensory modality (Passingham & Wise, 2012). An alternative, but given the task constraints less likely explanation, could be that the participants were using verbal encoding or rehearsal to select the target. The bilateral parieto-occipital cortex and dorsal premotor cortex activation suggest processing of spatial representations; it was seen both during action selection and movement. Moreover, parieto-occipital cortex and dorsal premotor cortex are strongly connected in non-human and human primates and together critical for spatial sensori-motor transformations (Tomassini et al., 2007). Taken together, these observations indicate that the identity condition recruited brain areas related to encoding the enduring characteristics of an object, whereas the spatial condition, recruited brain areas related to encoding the current spatial location of objects. These findings illustrate the high sensitivity of the current setup and design to elicit task specific activations. It remains to be tested if these different sources of information are processed in a similar or distinct way depending on movement intention, either communicative or instrumental. An interaction between areas involved in token selection, as shown here, and areas specifically activated in communicative but not instrumental pointing would show how different sources of information are integrated in either a mind-oriented or a goal-oriented action plan.

5. Conclusion

We investigated at what level of motor organization, the planning of communicative pointing actions differs from the well-studied instrumental counterparts. We found evidence supporting both hypotheses proposed. First, we found evidence for the presence of an ostensive tag reflecting increased communicative intent mediated by a longer holding phase. This shows that higher-order information of the communicative goal can access lower-level motor control on an instruction-level. Second, we found evidence suggesting that a communicative intention can pervasively affect the low-level implementation of a pointing gesture to serve the communicative content and adapt to the viewpoint of the audience. This suggests that the communicative faculty of an action is highly integrated with the motor system, adaptively influencing action execution at multiple levels of motor control. As such, pointing actions are not solely goal-oriented towards an object, but they are 'mind-oriented' in nature, sensitive to the conceptual and pragmatic knowledge of the communicative goal.

6. References

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*: 329-49.

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). The case of Dr. Jekyll and Mr. Hyde: A kinematic study on social intention. *Consciousness and Cognition, 17*: 557–64.

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology, 30*, 359-75.

Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., & van Rooij, I. (2012). Recipient design in human communication: simple heuristics or perspective taking? *Frontiers Human Neuroscience, 6*: 253.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*:1482-93.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and brain sciences 25*: 657-726.

Cleret de Langavant, L., Remy, P., Trinkler, I., McIntyre, J., Dupoux, E., Berthoz, A., & Bachoud-Lévi A. C. (2011). Behavioral and neural correlates of communication via pointing. *PLoS One, 15*(6), e17719.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Science, 8*: 8-11.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex 48*: 905-22.

Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20-5.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to- brain coupling: a mechanism for creating and sharing a social world. *Trends in Cognitive Sciences 16*: 114-21.

Jeannerod, M. (1984). The timing of natural prehension movements. *Journal Motor Behavior, 16*, 235-54.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*, 297–307.

Liu, D., & Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *The Journal of Neuroscience, 27*, 9354-68.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *The Journal of Neuroscience Methods, 164*, 177-90.

McIntyre, J., Stratta, F., & Lacquaniti, F. (1997). Viewer-centered frame of reference for pointing to memorized targets in three-dimensional space. *Journal of Neurophysiology, 78*(3), 1601-18.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia, 46*(3), 774-85.

Newman-Norlund, S. E., Noordzij, M. L., Newman-Norlund, R. D., Volman, I. A. C., de Ruiter, J., Hagoort, P., & Toni, I. (2009). Recipient design in tacit-communication. *Cognition 111*: 46–54.

Nichols, T. E., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*, 653-60.

Passingham, R. E., & Wise, S. P. (2012). The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight. Oxford: Oxford University Press.

Peeters, D., Chu, M., Holler, J., Özyürek, A., & Hagoort, P. (2013). Getting to the point: The influence of communicative intent on the kinematics of pointing gestures. *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*: 1127-1132.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in neuroscience 21*: 5: 188-94.

Rushworth, M. F., Buckley, M. J., Gough, P. M., Alexander, I. H., Kyriazis, D., McDonald, K. R., & Passingham, R. E. (2005). Attentional selection and action selection in the ventral and orbital prefrontal cortex. *The Journal of Neuroscience 14*, 25(50), 11628-36.

Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Conscious Cognition, 18*(3), 766-72.

Schot, W. D., Brenner, E., & Smeets, J. B. (2010). Robust movement segmentation by combining multiple sources of information. *The Journal of Neuroscience Methods 187*, 147-55.

Stolk, A., Hunnius, S., Bekkering, H., & Toni, I. (2013). Early social experience predicts referential communicative adjustments in five-year-old children. *PLoS One, 8*: e72667.

Tomasello, M. (2006). Why don't apes point? In N. J. Enfield, & S. C. Levinson (Eds.), *Roots of human sociality: culture, cognition and interaction* (pp. 506-524). Oxford: Berg.

Tomasello, M. (2008). Origins of Human Communication. Cambridge, MA/London: The MIT Press.

Tomassini, V., Jbabdi, S., Klein, J.C., Behrens, T.E.J., Pozzilli, C., Matthews, P.M., Rushworth, M. F. S., & Johansen-Berg, H. (2007). Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *The Journal of Neuroscience, 27*(38), 10259-69.

Van Rooij, I., Kwisthout, J., Blokpoel, M., Szymanik, J., Wareham, T., & Toni, I. (2011). Intentional communication: Computationally easy or difficult? *Frontiers in Human Neuroscience, 5*: 52.

Verhagen, L., Dijkerman, H. C., Grol, M. J., & Toni, I. (2008). Perceptuo-motor interactions during prehension movements. *The Journal of Neuroscience 28*(18), 4726-35.

Abstracts

Proceedings of the Master's Programme Cognitive Neuroscience is committed to publishing all submitted theses. Given the number of submissions we select certain articles under the recommendation of the editors for our printed edition. To interested readers, we have provided the abstracts of all other articles of which the full versions are available on our website: www.ru.nl/master/cns/journal.

Directional Asymmetries in Vowel Perception

Nadine P.W.D. de Rue, Alejandrina Cristia, Paula Fikkert, Sho Tsuji

Directional asymmetries in vowel perception are attributed a key role in language acquisition. Therefore, it is of the outmost importance to understand why such asymmetries exist. In the theoretical portion of this thesis, we first provide an overview of current hypotheses. One of them suggests that directional asymmetries in vowel perception are caused by natural discontinuities in the general auditory system. If this hypothesis is correct, asymmetries should be elicited by linguistic stimuli as well as by non-linguistic analogues that preserve crucial acoustic information. In the empirical section, we conducted two perceptual tasks to verify this prediction. We tested discrimination of spoken vowel contrasts and their nonspeech analogues in adult Dutch listeners in two nonnative vowel pairs. All hypotheses predicted more efficient discrimination in linguistic stimuli when the vowel changed in peripheral direction than when the vowel changed towards the center of vowel space; but only the general auditory hypothesis predicted this bias to extend to non-linguistic analogues. In fact, results showed no evidence of perceptual asymmetries in speech discrimination. We conclude that asymmetries are weak in adults, leading to small effect sizes. Regarding the non-linguistic analogues, results were promising as adults exhibited discrimination performance above chance level. Thus, further work with larger adult participant samples and/or with infant samples exploiting the same non-linguistic analogues bears promise to answer the question of the etiology of vowel asymmetries.

What you see is What You've Learned: Reward Biases Visuospatial Attention

Sebastiaan den Boer, Johanna Zumer, Sean Fallon, Ole Jensen

Reward can have a large impact on behavior. This could be due to a modulation of selective attention, but could also be explained by an increase in effort. The present study disentangles these two possible explanation, and investigates whether reward may directly affect attentional control by creating a perceptual bias. By orthogonalizing reward from task performance, we show that this is indeed the case. We show that reward creates a non-volitional perceptual bias that increases the amount of attention that is allocated towards a previously rewarded stimulus. This effect is reflected in reaction times and also in modulations of posterior alpha oscillations.

Do Alpha Oscillations Provide a Mechanism for Prioritizing Salient Unattended Stimuli?

Yağmur Güçlütürk, Mathilde Bonnefond, Ole Jensen

Recently, it has been hypothesized that alpha oscillations allow the representations of competing unattended stimuli to be ordered in the visual cortex with respect to their saliency levels. According to this framework, alpha oscillations and gamma oscillations (>30 Hz) phase- locked to alpha oscillations produce a temporal phase code for saliency. Furthermore, power of alpha oscillations that increases with reduced attention, determines the number of unattended stimuli that can be processed, such that a higher alpha power would result in fewer representations of stimuli to activate, and vice versa. In this study we wanted to answer the question: 'does alpha activity provide a mechanism for prioritizing and ordering unattended visual inputs with respect to the saliency levels of each input?' To this end, we performed three experiments in which, we have taken advantage of the well-studied attentional phenomenon called the prior entry effect. In particular, the experiments involved visual temporal order judgement tasks with and without a spatial attention component, and an auditory task. We used magnetoenchephalography (MEG) to measure oscillatory activity in order to test our hypothesis. The analysis of the behavioural and MEG data revealed a number of differences between the task conditions, which suggests a stronger prioritization for higher alpha power in the occipital brain regions, in line with the suggested mechanism for prioritization and ordering of salient unattended stimuli. Our results warrant further research on the suggested framework.

Studying the Acquisition of Emotion Words Over Time: a Longitudinal Behavioral and Electrophysiological Study

Renée Middelburg, Agnes Sianipar, Ton Dijkstra

This paper investigated the sensitivity of second language learners to the emotional content of words. In addition, the potentially helpful role of cognates in accessing the emotional contents of words was examined. Thirty-two young adult German native speakers participated in four Dutch lexical decision experiments, combining reaction time (RT) measures with EEG recordings over their first six months of learning Dutch (L2). The learners already reacted differently to emotion words after only one to two weeks of language learning. Especially positive words seemed to have a special status in the earliest stages of second language acquisition: RTs to these words were faster, fewer errors were made on them, and a more positive-going Late Positive Complex was found in response to these words. Arousal of the words itself did not affect the ERPs, but it did influence the valence response. In sum, from the earliest stages of second language acquisition onwards, emotional valence played a role in L2 processing, and, as has been found previously, learning a new language was a fast process; i.e., processing words in a new language up to the semantic and emotion levels already occurs at a very early stage of learning. We did not find strong support for a facilitating role of cognates, neither in the ERP patterns, nor in the behavioral data.

Linearity of Spin-echo and Gradient-echo BOLD

Nils Müller, David Norris, Markus Barth

The blood-oxygenation level dependent (BOLD) contrast is one of the most important means of measuring brain activity. For the analysis of most functional magnetic resonance imaging data it is assumed that the BOLD contrast behaves in a linear fashion. Recent literature suggests differences between gradient-echo (GE) and spin-echo (SE) based sequences in terms of SE showing a more linear response at short inter stimulus intervals (ISI). In case of densely spaced events this could result in spin echo being superior to GE despite its reduced sensitivity. We investigated this differential linearity in the visual cortex for SE and GE at 7 Tesla. We used a dual echo sequence, acquiring both GE and SE simultaneously, and a paired-stimulus paradigm. Based on our results we cannot conclude whether there is a difference in linearity. This could be a result of the unexpectedly low temporal signal to noise ratio (tSNR) of SE compared to GE.

The Effects of an Enriched Environment on the Intrinsic Properties and Neuromodulations of CA1 Pyramidal Cells

María Jesús Valero Aracama, Motoharu Yoshida, Wim Scheenen

It has long been established that living in enriched environments (EE) improves learning and memory. While enhancement of hippocampal plasticity by EE has also been widely reported, the modification of intrinsic cellular properties and their interaction with neuromodulators are not well understood. Using whole cell patch clamp techniques in the hippocampal CA1 slice preparation from mice housed in EE and control environments (CE), I tested the modification of intrinsic cellular properties of CA1 pyramidal cells by an EE. The EE manipulation increased the excitability and the sag amplitude and its ratio in response to a hyperpolarizing current pulse. Additionally, I tested the acute effects of four neuromodulators that are crucial for learning and memory: acetylcholine (ACh) agonist, noradrenaline (NA), serotonin (5-HT) and brain-derived neurotrophic factor (BDNF) agonist. EE-cells seem to have a higher sensitivity to NA and 5-HT in reducing the sag amplitude and ratio. The sag response has been reported as a measure for the amount of hyperpolarization-activated cation current (Ih) in the cell. Since the Ih is known to be modulated by cyclic adenosine monophosphate (cAMP) and NA and 5-HT modulate cAMP levels, these results suggest a regulation of cAMP actions in EEs. In addition, I tested the ability of CA1 cells to support persistent firing (PF) after a short current pulse as a neural mark of working memory. PF was observed with cholinergic and TrkB receptor agonists, but not with NA or 5-HT suggesting a role of ACh and BDNF in PF. Since previous studies found an increment of BDNF in the hippocampi from EE-animals, EE might improve learning by facilitating PF through increasing BDNF concentration in hippocampus. In summary, my data suggests that EE increases the intrinsic excitability and the Ih, modifies the cellular sensitivity to NA and 5-HT, and it probably increases PF through a BDNF concentration increment.

MicroRNA-338 in Neuronal Growth and Development: Investigating the Role of microRNA-338 in on the Morphological Development of Cortical Neurons in vitro

Mark van Kessel, Aron Kos, Armaz Aschrafi

Post-transcriptional regulation of gene expression is an important factor healthy neuronal development. One of the regulatory mechanisms that has received a lot of attention recently is miRNAs, small non-coding pieces of RNA that influence the translation of messenger RNA. In this study we specifically looked at miRNA-338-3p to establish its role in neuronal development in vitro. We show that expression levels for miRNA-338 vary with region and developmental time points and, that at specific time points, there are gross morphological differences in terms of dendritic structure as well as dendritic spine head diameter and spine length. Spine density and synaptogenesis on the other hand were unaffected. These results indicate miRNA-338 might increase somatic projections and a role for miRNA-338 in synapse maturation or synapse maintenance.

Institutes associated with the Master's Programme Cognitive Neuroscience

Donders Institute for Brain, Cognition and Behaviour:
Centre for Cognitive Neuroimaging
Kapittelweg 29
6525 EN Nijmegen

P.O. Box 9101
6500 HB Nijmegen
www.ru.nl/neuroimaging/

Donders Institute for Brain, Cognition and Behaviour:
Centre for Neuroscience
Geert Grooteplein Noord 21, hp 126
6525 EZ Nijmegen

P.O. Box 9101
6500 HE Nijmegen
www.ru.nl/neuroscience

Donders Institute for Brain, Cognition and Behaviour:
Centre for Cognition
Montessorilaan 3
6525 HR Nijmegen

P.O. Box 9104
6500 HB Nijmegen
www.ru.nl/cognition/

MAX-PLANCK-GESELLSCHAFT

Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD Nijmegen

P.O. Box 310
6500 AH Nijmegen
http://www.mpi.nl

IN·DEI·NOMINE·FELICITER

Nijmegen Centre for Molecular Life Sciences
Geert Grooteplein 28
6525 GA Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.ncmls.nl

Universitair Medisch Centrum St Radboud
Geert Grooteplein-Zuid 10
6525 GA Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.umcn.nl/

Baby Research Center
Montessorilaan 10
6525 HD Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://babyresearchcenter.nl