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

Programme Director: **Ardi Roelofs**
Journal Logo: **Claudia Lüttke**
Cover Image: **Layout Team**
Photo Editor-in-Chief: Provided by **Karita Ojala**

Contact Information:
Journal CNS
Radboud University
Postbus 9104
6500 HE Nijmegen
The Netherlands
nijmegencns@gmail.com

Table of Contents

From the Editor-in-Chief

Dear reader,

What you have in your hands is the second issue of *Proceedings of the Master's Programme Cognitive Neuroscience* for the academic year 2015-2016. Earlier this year, we have already published one issue full of interesting neuroscientific articles from the students of the research Master's programme.

For the current issue, the journal team selected articles especially focusing on perception, action and cognitive control. A wide variety of neuroimaging techniques have been used in these studies, including electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), as well as the brain stimulation method transcranial magnetic stimulation (TMS). The articles touch upon such matters as how stability in visual perception could arise from the influences of previously seen stimuli, choosing which hand to use for action in an environment requiring whole body movement, the compensational visual strategies of Parkinson's disease patients on movement initiation, and biases influencing approach and avoidance behaviour with regard to reward and punishment. Moreover, this issue includes an excellent article on the influences of the first learned language on processing of a second language. All in all, these articles show the high quality of research done in the Master's programme.

The publication of this journal issue has been made possible by enthusiastic researchers in training who put in a lot of time and effort. As the student journal is an extracurricular activity, students work on it voluntarily and next to their studies and other responsibilities. Moreover, the reviewers of the articles deserve a sincere thank you for evaluating the articles and making it possible for the journal team to make an informed decision on which articles to publish. I would also like to thank the authors for their work and for going through the revision process, improving their articles according to feedback received from the reviewers.

I wish you a pleasant read and give a final thank you to everyone involved in the making of this issue!

Nijmegen, July 2016

Karita Ojala

Editor-in-Chief

To Smooth or not to Smooth: Investigating the Role of Serial Dependence in Stabilising Visual Perception

Matthias Fritsche[1]
Supervisors: Pim Mostert[1], Floris de Lange[1]

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

Visual input is often noisy and discontinuous due to blinks, saccades, movements and many other factors. Yet visual perception is characterised by remarkable stability. A mechanism that has been proposed to mediate such perceptual stability is serial dependence (Fischer & Whitney, 2014). By using both previous and current input to generate a percept at the current moment, the brain could capitalise on the stability of the physical environment in order to stabilise perception. In the current study, we investigated two potential properties of serial dependence, which would further support its proposed role in mediating perceptual stability. First, we investigated whether serial dependence selectively stabilises percepts of the same object and ceases when a different object is perceived. Second, we investigated whether the degree to which the visual system leverages previous input to stabilise perception depends on the sensory uncertainty associated with previous and current input. Probing serial dependence in orientation perception, we found no evidence for object selectivity and only partial evidence for a sensible weighting of previous and current input according to sensory uncertainty – serial dependence was stronger when the uncertainty associated with the current stimulus was high, especially when the uncertainty associated with the previous stimulus was also high. Unexpectedly, we discovered an intriguing temporal dynamic of serial dependence: while the current percept was biased towards very recent input (~4 seconds ago), it was biased away from more remote input (15 to 40 seconds ago), possibly mirroring temporal dynamics of the physical environment. We conclude that serial dependence could in principle serve to stabilise perception, but appears to lack properties which would allow to do so in an optimal manner.

Keywords: perceptual stabilisation, serial dependence, priming, orientation perception

---

Corresponding author: Matthias Fritsche; **E-mail:** m.fritsche@donders.ru.nl

1. Introduction

In day-to-day life our visual systems are confronted with a persistent problem: the visual input that gives rise to our percepts is highly unstable. Visual input is continually disrupted by internal and external events such as blinks, saccades, movements and temporary occlusions. Moreover, the visual system, like any biological system, is inherently noisy. Yet, despite of all these factors of instability, human visual perception appears to be remarkably stable. However, so far it has remained an open question of how this perceptual stability is achieved.

In general, it is widely assumed that the brain exploits statistical regularities of the world in order to optimise perception (Barlow, 1961; Simoncelli & Olshausen, 2001). One important property of the world is that it remains relatively stable over short timescales. This property could be exploited by the brain in order to stabilise perception. In particular, given the continuity in the physical environment, the brain could assume that the current visual input should be similar to recently experienced visual input. Indeed, such a mechanism was recently put forward in the form of serial dependence (Fischer & Whitney, 2014). Serial dependence refers to the phenomenon that the subjective percept of a stimulus is quantitatively biased towards stimuli seen in the recent past. For instance, when viewing oriented Gabor patches, the perceived orientation of the currently presented Gabor patch is systematically biased towards orientations of stimuli seen up to 15 seconds ago. Effectively, subsequent percepts appear more similar to each other than the underlying physical stimuli. Besides orientation perception, serial dependence has been demonstrated in other domains, such as numerosity and face perception (Cicchini, Anobile, & Burr, 2014; Liberman, Fischer, & Whitney, 2014). Therefore, smoothing perception through serial dependence could, in principle, mediate perceptual stability.

The goal of the current study was to investigate whether serial dependence exhibits properties, which would further support its proposed role in stabilizing perception. In particular, we focused on two predictions derived from the framework of perceptual stabilisation.

The first prediction is that, if serial dependence serves to stabilise perception, it should optimally operate on the level of objects, since objects tend to be stable and continuously present from one moment to the next. Consequently, the percept of an object should be biased to previous percepts of the same object, mediating perceptual stability. However, serial dependence between subsequent stimuli should cease, when these clearly represent different objects. In these cases, where a true change in the physical environment is apparent, smoothing perception would be detrimental, as it would obstruct the detection of this change. Therefore, we hypothesised that serial dependence should be object selective.

The second prediction states that the degree to which the visual system leverages input of the recent past to smooth perception should rely both on the quality of previous input and the quality of current input. For instance, imagine a night scene of a street with a flickering streetlight. In periods of illumination you might have a good perception of objects around you. In periods of partial darkness, on the other hand, visual information is of lower quality. For the visual system it would be a good strategy to use information from a period of illumination to stabilise perception during darker periods, but not the other way round. Consequently, the visual system should weight sensory information of the past and present according to their quality or associated uncertainty when generating the current percept. Notably, this idea is in line with popular Bayesian models of perception that postulate that our brains combine inherently noisy and ambiguous sensory information with prior knowledge about the world in a probabilistically optimal manner (Fiser, Berkes, Orbán, & Lengyel, 2010; Vilares & Kording, 2011; Yuille & Kersten, 2006). These models have been used to explain biases in orientation perception (Girshick, Landy, & Simoncelli, 2011; Wei & Stocker, 2015) and adaptation effects (Stocker & Simoncelli, 2006b). Bayesian principles could therefore lie at the heart of the neural computations leading to serial dependence.

We conducted two experiments in order to test the predictions formulated above. In both experiments we used orientation-judgment tasks to probe serial dependence. Participants viewed a series of randomly oriented stimuli and had to report the perceived orientation of each stimulus by adjusting a response bar. Subsequently, we studied the influence of previously presented stimuli on the perceived orientation of the current stimulus. In Experiment 1, we tested whether serial dependence in orientation perception is object selective. Our stimuli consisted of oriented Gabor patches. We manipulated the identities of the Gabor patches by varying their spatial frequency between two levels from trial to trial. We found no evidence for object selectivity. Instead, our results suggest a dynamic weighting of previous and current visual input according to amount of task relevant information, which varied between the different spatial frequency levels. This finding was in line with our second prediction. In Experiment 2, we tested whether serial dependence takes sensory uncertainty into account when generating the current percept. Our stimuli were Fourier filtered noise patches, which allowed for a fine control of orientation information, or noise, in each stimulus. Sensory uncertainty was manipulated by varying the amount of stimulus noise from trial to trial. We found that serial dependence is stronger when the noise in the current stimulus is high – especially so if the noise in the previous stimulus was also high. This finding partially contradicts our hypothesis of dynamic weighting of input by associated stimulus uncertainty. Strikingly, in both experiments we found a shift in the polarity of serial dependence for stimuli presented further in the past. That is, while the current percept was biased towards stimuli seen just before, it was also biased away from stimuli presented further back in time. In summary, we found only partial evidence for properties, which would support a functional role of serial dependence in stabilizing perception. However, our results point to an interesting, previously unknown temporal dynamic of serial dependence.

2. Experiment 1 – is serial dependence object selective?

In Experiment 1, we tested whether serial dependence in orientation perception is object selective. The hypothesis of object selectivity was derived from the idea that, if serial dependence would serve to stabilise perception, it should operate optimally on the object level. This is because objects tend to be constant, rarely changing from one moment to the next. The idea of serial dependence operating on the object level generates two predictions. First, serial dependence should bias the percept of an object towards previous percepts of the same object. Second, if properties of an object clearly change, indicating that a change in the identity of the object occurred, serial dependence should cease, since continued perceptual smoothing would obstruct the detection of a true change in the environment. Together, these two predictions form the object selectivity hypothesis.

Generally, the object selectivity hypothesis can be tested by measuring serial dependence between successively presented stimuli, while manipulating the object identity of these stimuli. According to the hypothesis, the percept of the current stimulus should be biased towards the one of the previous stimulus, if both stimuli are perceived as being the same object. Conversely, this bias should be reduced or absent if two successive stimuli are perceived as being different objects.

In the current experiment, we adopted the Fischer & Whitney paradigm, in which participants viewed a series of randomly oriented Gabor patches and had to report the perceived orientation of each Gabor by adjusting a response bar. Crucially, Gabor patches in our experiment were either of low or high spatial frequency, and spatial frequency could vary from trial to trial. Consequently, a Gabor on the current trial could be preceded by a Gabor with the same or with a different spatial frequency presented on the previous trial. We expected to find a systematic bias of perceived orientation towards the orientation of the previous Gabor, if previous and current Gabor were of the same spatial frequency. We further hypothesised that, in case of object selectivity, serial dependence in perceived orientation should cease, or be reduced, if current and previous Gabors were of different spatial frequencies, indicating a change in object identity. We chose to manipulate object identity via changes in stimulus spatial frequency since abrupt changes in spatial frequency rarely occur within the same object over short timescales and thus are a good indicator for changes of object identity.

2.1 Methods

2.1.1 Participants

Twenty-three participants (15 female, age 19–29 years), with normal or corrected-to-normal visual acuity, participated in this experiment. All participants provided informed written consent prior to the start of the experiment. The study was approved by the Radboud University Institutional Review Board.

2.1.2 Stimuli & design

The sequence of events within each trial is illustrated in Figure 1A. Throughout the entire experiment a central fixation dot (white, 0.6° visual angle diameter) was presented on a mid-grey background. Participants were instructed to maintain fixation. First, a randomly oriented Gabor patch (windowed sine wave grating) was presented at 6.5°

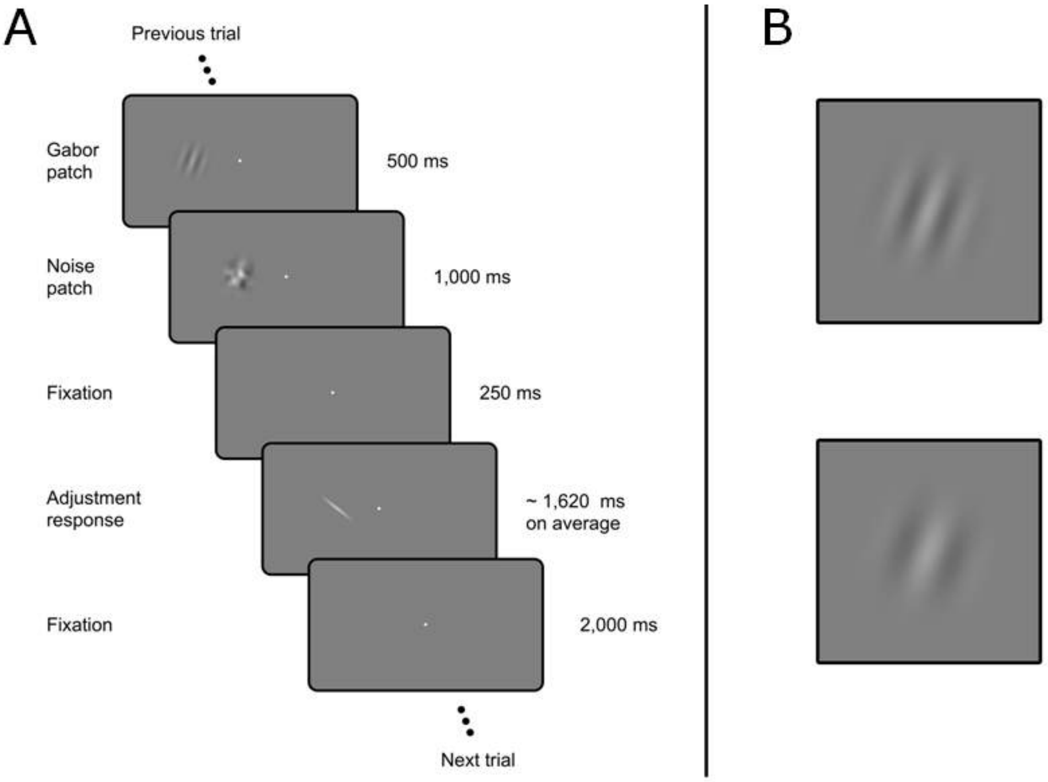

Fig. 1 A. Sequence of events within each trial of Experiment 1. Participants viewed a high or low spatial frequency Gabor stimulus either to the left or right of fixation (separate, interleaved blocks) and subsequently reported the perceived orientation of each Gabor by adjusting the orientation of a response bar. **B.** Examples of a high spatial frequency stimulus (upper panel) and a low spatial frequency stimulus (lower panel).

eccentricity (left or right of fixation dot, alternating in separate interleaved blocks). The Gabor patch could either have a low spatial frequency (0.33 cycles/°) or high spatial frequency (0.5 cycles/°; see Fig. 1B). The patches were windowed by a Gaussian envelope (1.5° *SD*) and had 25% Michelson contrast. After 500 ms the Gabor patch was replaced by a noise patch, presented for 1000 ms, to minimise effects of visual afterimages. The noise patches consisted of white noise, smoothed with a 0.5° *SD* Gaussian kernel and windowed by a Gaussian envelope (1.5° *SD*). A response bar (0.3° wide white bar windowed by a 0.8° Gaussian envelope) appeared 25 ms after offset of the noise patch at the same location as the Gabor and noise patches. The bar's initial orientation was determined randomly in each trial and it could be rotated clockwise and anti-clockwise using the right and left arrow keys of a standard keyboard. The participants' task was to adjust the response bar such that it matched the orientation of the Gabor patch seen just before. Once adjusted to the desired position, the response was submitted by pressing the space bar. The response was followed by a 2 second inter-trial-interval. Each participant completed a series of 820 trials, divided into 10 blocks. The sequences of presented stimuli were temporally counterbalanced with respect to the combinations of spatial frequencies of Gabor patches on trial n and $n-1$. That is, stimuli of low (high) spatial frequency were preceded equally often by stimuli of low and high spatial frequencies. The stimulus sequences were pseudo-randomised for each participant.

Stimuli were generated with the Psychophysics Toolbox for MATLAB (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) and were displayed on a 24" flat panel display (resolution: 1920 x 1080, refresh rate: 60 Hz). Participants viewed the stimuli from a distance of 53 cm in a dimly lit room.

2.1.3 Data analysis

Outlier correction

In the first step of data analysis we excluded those trials in which the response error (shortest angular distance between stimulus orientation and response orientation) was further than three circular standard deviations away from the participant's mean response error. This was done in order to exclude trials on which the participant gave random responses due to blinks or attentional lapses during stimulus presentation as well as due to inadvertent responses. After that, response errors were demeaned for each participant to remove general clockwise or anti-clockwise response biases.

Measuring serial dependence

The general procedure for assessing perceptual serial dependence was as follows. First, within participants, we expressed the response error of the current trial as a function of the difference between previous (n-back) and current stimulus orientation (for an illustration please see Supplementary Fig. 1). Positive values of this difference indicate that the previous stimulus was rotated more anti-clockwise than the current stimulus. Consequently, data points in this error plot, whose x and y values have the same sign, indicate that the participant's response error in that trial was in the direction of the previous (n-back) stimulus orientation. For instance, a data point with positive x and y value represents a trial in which the previous stimulus was oriented more anti-clockwise than the current stimulus, and the participant also adjusted the response bar on the current trial more anti-clockwise than the current stimulus orientation. Thus, this data point would represent a trial in which the current percept was biased towards the previous stimulus. Second, we applied a circular moving average to these conditioned response errors to increase robustness of subsequent data analyses. To this end, we computed the circular mean response error in a 20° window, sliding over the relative orientation difference between current stimulus orientation and previous (n-back) stimulus orientation (−90 to 90°). This yielded *serially conditioned error moving averages* for every subject. Third, we combined the single subject serially conditioned error moving averages into a grand moving average and quantified the strength of serial dependence on the group level. In order to quantify the strength of serial dependence, (i.e., how much the response on the current trial was pulled towards the previous [n-back] stimulus orientation), we fitted the first derivative of a Gaussian curve (DoG) to the grand moving average. The DoG is given by

$$y = xacwce^{-(wx)^2}$$

where x is the relative orientation of the previous trial, a is the amplitude of the curve peaks, w is the width of the curve and c is the constant sqrt(2)/$e^{-0.5}$. The constant c is chosen such that parameter a numerically matches the height of the curve peak. The amplitude parameter a was taken as the strength of serial dependence, as it indicates how much the response to the current stimulus orientation could be biased towards a previous (n-back) stimulus for the maximally effective orientation difference between trials. In the remainder of the article, where not otherwise stated, the width parameter w of the DoG curve was treated as a free parameter that was fitted to the data, constrained to a wide range of plausible values ($w = 0.02 - 0.2$). Fourth, we quantified the strength of serial dependence effects exerted by the previous stimulus (1-back) and all stimuli up to 20 trials back in the past (20-back), irrespective of the spatial frequency of previous and current trial.

Further, we split the data into subsets of trials for each combination of previous and current stimulus spatial frequency, yielding sets of low-to-low, high-to-high, low-to-high and high-to-low spatial frequency stimulus pairs, hereafter referred to as spatial frequency conditions. Similarly to the procedure described above, we computed serially conditioned error moving averages for every subject in each spatial frequency condition. How we made use of the serially conditioned error moving averages of each spatial frequency condition to test for object selectivity and how we tested for differences of serial dependence between spatial frequency conditions is described below.

Statistical testing

We used permutation tests for testing all serial dependence effects of interest. In the following, it is first described how we tested for general serial dependence effects, regardless of spatial frequency manipulations. Thereafter, it is described how we tested the object selectivity hypothesis using the spatial frequency conditions, and how we tested for other effects of spatial frequency on serial dependence.

In order to statistically assess general serial dependence effects, regardless of spatial frequency manipulations, we performed permutation tests separately for the serial dependence estimates of all n-back stimuli. Permutation distributions for the strength of serial dependence on the n-back stimulus were computed in the following way: a single permutation was computed by first randomly inverting the sign of each participant's respective n-back serially conditioned error moving average (i.e., randomly flipping the moving averages along the horizontal axis). This is equivalent to randomly shuffling the labels between the empirically observed data and an artificial null distribution of no serial dependence (a flat surrogate serially conditioned error moving average) and subtracting the two conditions from each other per participant. The resulting individual moving averages were combined in a new grand moving average. Subsequently, we fitted a new DoG model to this grand moving average and collected the resulting amplitude parameter *a* in the permutation distribution. We repeated this permutation procedure 10,000 times. As *p*-values we report the percentage of permutations that led to equal or higher absolute values for *a* than the absolute value of the empirically observed *a*. The significance level was set to α = .05 and Bonferroni-corrected for multiple comparisons, resulting in a corrected significance level of α = .0025. The exchangeability requirement for permutation tests is met, because under the null hypothesis of no serial dependence, the labels of the empirically observed data and an artificial null distribution of no serial dependence (flat surrogate serially conditioned error moving average) are exchangeable.

Next, we tested the difference in serial dependence strength for cases in which previous and current stimuli had the same spatial frequency versus cases in which the spatial frequency changed. For this purpose, we used a different permutation testing procedure. With this procedure, we also tested for main effects of current and previous trial's spatial frequency on serial dependence in the current trial, their interaction effect, and the differences in serial dependence between any two spatial frequency conditions.

In the following, the general procedure of the permutation test will be explained by the example of testing differences of serial dependence between trials of the same spatial frequency versus trials of different spatial frequencies. We quantified the difference of serial dependence between these two cases by summing, within participant, the serially dependent error moving averages of low-to-low and high-to-high spatial frequency conditions into a 'same' spatial frequency condition and summing the serially dependent error moving averages of low-to-high and high-to-low spatial frequency conditions into a 'different' spatial frequency condition. We subtracted the 'different' spatial frequency condition from the 'same' spatial frequency condition, again within each participant, and combined the resulting *single-subject contrast moving averages* into a *grand contrast moving average*. After that, we fitted a DoG model to this grand contrast moving average. Here, the width parameter *w* of the DoG model was set to the empirically determined width parameter of the best fitting model to the 1-back data, regardless of the spatial frequency manipulation. This was done to prevent implausible fits to the grand contrast moving average as well as to subsequent permutated versions thereof. The resulting amplitude parameter *a* reflected the difference in serial dependence between the 'same' and the 'different' spatial frequency condition (positive values indicate a stronger serial dependence in the 'same' spatial frequency condition). Next, the permutation distribution was generated as follows: For a single permutation the signs of the single-subject contrast moving averages were randomly inverted (i.e., flipped along the horizontal axis) for each participant. This is equivalent to randomly shuffling the labels of the 'same' and 'different' spatial frequency conditions per participant and subtracting the two conditions from each other. After that, the resulting single-subject contrast moving averages were combined into a grand contrast moving average. Subsequently, we fitted a DoG model with fixed width to the grand contrast moving average and collected the resulting amplitude parameter *a* in the permutation distribution. We repeated this permutation procedure 10,000 times. As *p*-values we report the percentage of permutations that led to equal or higher absolute values for *a* than the absolute value of the empirical *a*. The significance level was set to α = .05. The exchangeability requirement for permutation tests is met, because under the null hypothesis of no difference in serial dependence between 'same' and 'different' spatial frequency conditions, the condition labels are exchangeable.

With a similar procedure we statistically assessed main effects of current and previous trial's spatial frequency, their interaction effect, and the differences in serial dependence between any two spatial frequency conditions. The difference to the above example was how the spatial frequency conditions were combined into the single-subject contrast moving averages. A contrast moving average for the contrast of interest was computed with the formula

$$CMA = w_{ll} * MA_{ll} + w_{lh} * MA_{lh} + w_{hl} * MA_{hl} + w_{hh} * MA_{hh}$$

where *w* are weighting coefficients specific to the contrast, *MA* denotes a serially dependent error moving average of the spatial frequency condition and *CMA* stands for contrast moving average. Subscripts *ll*, *lh*, hl and *hh* denote the spatial frequency of previous and current stimulus (low/high). An overview over the weighting coefficients for the respective contrasts is given in Table 1.

Conditioning on previous response

Additionally to the above analyses, we also investigated the dependence of the current percept on previous response orientations, instead of previous stimulus orientations. The previous

Table 1
Weighting coefficients for the permutation test framework used in Experiment 1

	w_{ll}	w_{lh}	w_{hl}	w_{hh}
Same vs. different	+0.5	−0.5	−0.5	+0.5
Main effect of prev. trial	+0.5	+0.5	−0.5	−0.5
Main effect of current trial	+0.5	−0.5	+0.5	−0.5
Prev. x current trial	+1	−1	−1	+1
Low-Low minus Low-High	+1	−1	0	0
Low-Low minus High-Low	+1	0	−1	0
Low-Low minus High-High	+1	0	0	−1
Low-High minus High-Low	0	+1	−1	0
Low-High minus High-High	0	+1	0	−1
High-Low minus High-High	0	0	+1	−1

response, rather than the physical stimulus itself, should more closely resemble the subjective percept of the previous stimulus orientation and would be a likely candidate for biasing subsequent percepts. However, preempting the results, we found that the estimate of serial dependence on the previous response was severely confounded by a general bias of responses to the oblique orientations. A more thorough description of the problem is given in the results section. Analyses regarding serial dependence on previous responses and its confounds can be found in the Supplementary Information online.

All data analyses were performed with MATLAB (The Mathworks Inc., Natick, MA, USA) and the CircStat MATLAB toolbox for circular statistics (Berens, 2009).

2.2 Results

On average 13.43 of 820 trials were excluded per participant (*SD* = 9.31, low spatial frequency: 7.30 ± 5.6; high spatial frequency: 6.13 ± 4.53). For the remaining data, the mean response error was 9.15° ± 2.47 (*SD*) and it was slightly, but not significantly, higher for low spatial frequency stimuli than for high spatial frequency stimuli (difference 0.16°; $t(22) = 1.23$, $p = .23$, two-sided paired t-test). The mean response time was 1.62 ± 0.48 seconds and there was no significant difference between low and high spatial frequency trials (difference: 4.5 ms; $t(22) = 0.17$, $p = .86$, two-sided paired t-test). Therefore, the onset of the current stimulus occurred on average ~3.87 seconds after the offset of the previous stimulus. Further, when splitting the data into spatial frequency conditions, each condition contained an average of 196 trials per participant.

We found an attractive serial dependence of the current percept on the previous trial's stimulus (amplitude: 1.55°, Fig. 2A; $p = .0005$, two-sided permutation-test), which was considerably smaller than previously reported by Fischer and Whitney (2014). Crucially, we found no significant difference in serial dependence for cases in which previous and current stimuli had the same spatial frequency versus cases in which the spatial frequency changed, although serial dependence was slightly stronger for the 'same' spatial frequency condition (amplitude of difference: 0.2°; $p = .43$, two-sided permutation test). This challenges the object selectivity hypothesis. Further, there were no significant main effects of previous or current spatial frequency, nor was there a significant interaction effect. Interestingly, permutation tests between pairs of spatial frequency conditions revealed a significant difference in serial dependencies between the low-to-high spatial frequency condition, showing the lowest serial dependence, and the high-to-low spatial frequency condition, showing the highest serial dependence (Fig. 2B, amplitude of difference contrast: −1.13°; $p = .02$, two-sided permutation test).

When investigating the dependence of the current percept on the previous response orientation, we found a strong serial dependence effect (Supplementary Fig. 2B, amplitude: 3.96°; $p < .0001$, two-sided permutation test), which was much stronger than the serial dependence effect exerted by the previous stimulus (amplitude of difference contrast: 2.28°; $p < .0001$, two-sided permutation test). However, orientation judgments in our experiment showed a general bias to oblique orientations, which has been reported in previous studies (Tomassini, Morgan, & Solomon, 2010; Wei & Stocker, 2015). This bias, which was solely determined by the current stimulus orientation, acted as a strong confound when assessing serial dependence on the previous response. In particular, the oblique bias had two effects, which in conjunction introduced confounding effects. First, previous responses were more likely to be of oblique orientation than of any other orientation. Therefore, when expressing the orientation difference between previous response and current stimulus orientation,

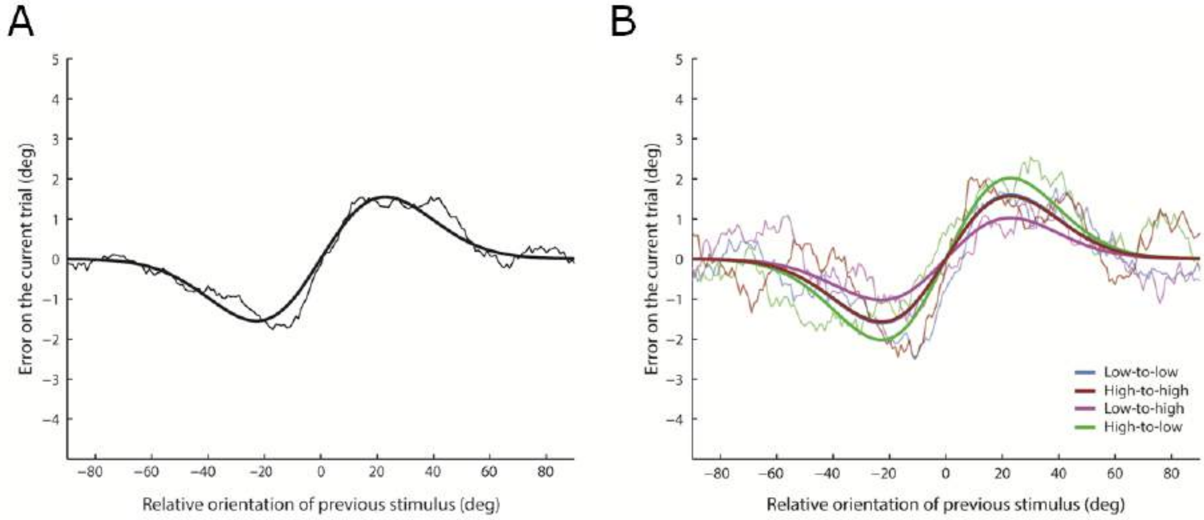

Fig. 2. Serial dependence of orientation perception in Experiment 1: Response errors on the current trial conditioned on the previous stimulus orientation **A.** for all trials irrespective of spatial frequency, **B.** separately for each spatial frequency combination of previous and current stimulus. Positive values on the horizontal axis indicate that the stimulus of the previous trial was oriented more anti-clockwise than the stimulus of the current trial. Positive values on the vertical axis indicate that the response of the current trial was oriented more anti-clockwise than the current stimulus. In **A.** the thin line shows the grand moving average of response errors and the thick line show the best fitting DoG model. In **B.** thin lines show the grand moving averages of response errors for each spatial frequency condition (see figure legend). Thick lines show the best fitting DoG models, respectively. The DoG model of the low-to-low condition (blue) is mostly occluded by the DoG model of the high-to-high condition (red).

the current stimulus orientation was more likely subtracted from one of the oblique orientations than any other orientation. Consequently, the orientation difference between previous response and current stimulus orientation correlated with current stimulus orientation. Second, the response error on the current trial was to a large degree determined by the oblique bias and therefore was strongly dependent on the current stimulus orientation. Together, these two effects had the consequence that response errors, introduced by the oblique bias, were systematically arranged in the serial dependence error plot, and therefore confounded the estimate of true serial dependence. Analyses and simulations demonstrating this confounding effect when estimating serial dependence on previous responses can be found in the Supplementary Information online. Further, we present control analyses, showing that this confound did not substantially affect estimates of serial dependence on previous stimuli.

When investigating the time-course of serial dependence, regardless of spatial frequency, we found that serial dependence decreased for 2- and 3-back stimuli and was not significant (Fig. 3A). Interestingly, we found a repulsive effect on the current percept by stimuli that were presented more than three trials back. This negative serial dependence was maximal for stimuli presented six trials back (amplitude: −1.04°, Fig. 3B; $p < .0001$, two-sided permutation test, Bonferroni corrected significance level). The effect persisted and was significant, with exceptions, for stimuli up to 10 trials back in the past (Fig. 3A).

2.3 Discussion

The aim of Experiment 1 was to test whether serial dependence in orientation perception is object selective. We found no evidence for object selectivity, when manipulating object identity in the secondary feature dimension of spatial frequency. However, spatial frequency of previous and current stimulus, together, had an influence on the strength of serial dependence. When high spatial frequency stimuli were presented after low spatial frequency stimuli, there was a considerably stronger serial dependence than for the opposite order of stimulus presentation. Unexpectedly, we also found an intriguing effect for the time course of serial dependence: while the current percept was biased towards stimuli presented in the previous trial, it was also biased away from stimuli presented four to 10 trials back.

The absence of object selectivity presents potential functional disadvantage for serial dependence in

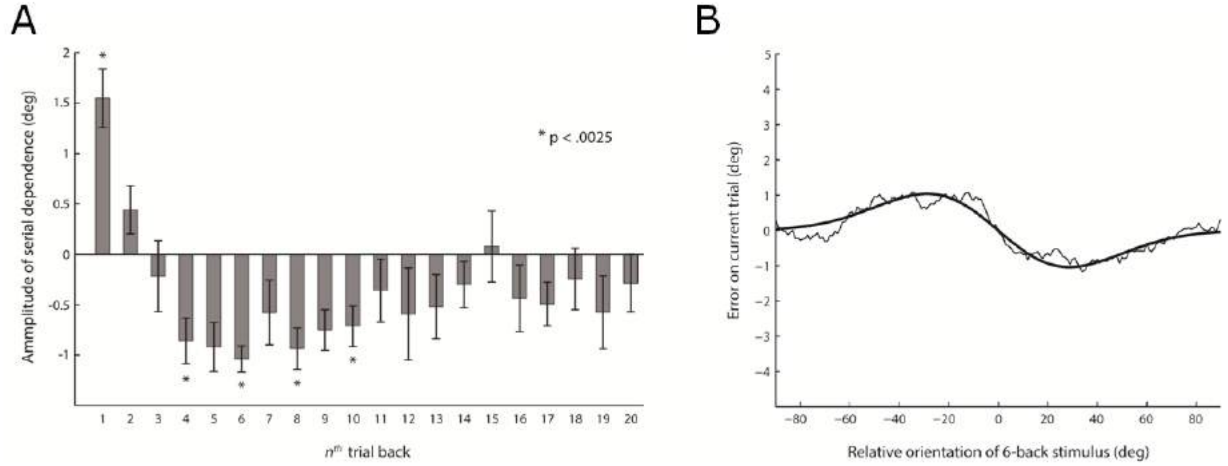

Fig. 3. Temporal dynamic of serial dependence in Experiment 1: **A.** Strength and sign of serial dependence (DoG amplitude parameter) on n-back stimuli; while the current percept is attracted by the previous stimulus, it is repelled by stimuli seen 4 to 10 trials back. Error bars represent 1 s.d. of the bootstrapped amplitude distribution. **B.** Current response error conditioned on the 6-back stimulus orientation. A clear repulsive effect of the 6-back stimulus on the current percept is observable. The thin line shows the grand moving average of response errors and the thick line show the best fitting DoG model.

visual perception. While benefiting from enhanced perceptual stability in a situation where the physical environment remains stable, this perceptual stability appears to come at the cost of decreased perceptual sensitivity to true changes in the environment. In particular, our experiment showed that the strength of serial dependence in the feature dimension of orientation is unaffected by changes in the secondary feature dimension of spatial frequency. It thus appears that visual perception of particular object features is smoothed, even in situations where other object features clearly change. Since abrupt changes in features, such as spatial frequency, rarely occur within the same object over short timescales and thus signal a change in object identity, it would have been functionally more beneficial to make perceptual smoothing in one feature dimension dependent on the continued presence of other features. At this point it must be noted that a recent study reported a breakdown of serial dependence in face perception, when successively presented faces were very different (Liberman et al., 2014). One could argue that this finding already demonstrates a sufficient degree of object selectivity. However, in those cases where serial dependence broke down, all task-relevant features for which perception was probed were quite different from one stimulus presentation to the next. That is, all facial features changed simultaneously. Thus, the previous study leaves open whether the strength of serial dependence is modulated independently in each feature dimension or whether serial dependence for a given feature is influenced by changes in other features. Our current results point to the former alternative. Since changes in object identity not necessarily imply stark changes in all feature dimensions pertaining to an object, a full account of object selectivity is thus challenged by the current findings.

Could it be the case that our manipulation of stimulus spatial frequency was too subtle to be effective? One could argue that our stimuli, with different spatial frequencies, could be interpreted as the same stimulus viewed at different distances. Further, it is possible, and in fact quite likely, that serial dependence also exists for the perception of spatial frequency. Therefore, the perceived difference in spatial frequency could be smaller than the true physical difference. Moreover, the differences in spatial frequency were not task-relevant and thus could be ignored by the observers. However, it must be noted that varying spatial frequency was not ineffective in manipulating the strength of serial dependence per se. That is, we found that spatial frequency of previous and current stimulus, together, had an influence on the strength of serial dependence. We thus assume that our experimental manipulation was in principle sufficiently strong to be effective, however our findings were not consistent with the hypothesis of object selectivity.

What may have caused the differences in serial dependence strength between the low-to-high and high-to-low spatial frequency stimulus pairs? A potential explanation pertains to the difference of orientation information in each stimulus.

Importantly, the number of luminance cycles that are present in the Gabor patch determine the amount of orientation information that can be conveyed by the stimulus. While the processing of spatial frequency information by the visual system is much more complex, and involves differential sensitivity and tuning to different ranges of spatial frequencies (Blakemore & Campbell, 1969; Campbell & Robson, 1968), it is possible that different amounts of orientation information in our stimuli, introduced via the different spatial frequencies, injected a difference in uncertainty about orientation in our observers. In particular, the low spatial frequency stimuli, which contained less orientation information, could have caused higher uncertainty about orientation than the high spatial frequency stimuli, which contained more orientation information. Consequently, the effects of spatial frequency on serial dependence reported in the present experiment could be explained by a dynamic weighting of previous and current input, depending on the amount of information or associated uncertainty about orientation that is present in each stimulus. More precisely, the percept of a low spatial frequency stimulus that contained less orientation information could be more prone to be biased towards previous stimulus presentations, especially if the previous stimulus was of high spatial frequency, and thus contained more orientation information. This way of combining previous and current visual input to smooth perception would be in line with Bayesian models of perception and could present an optimal strategy to ensure smooth, yet sensitive perception under uncertainty.

Additionally, we found a shift in the polarity of serial dependence for stimuli presented further back in the past. In particular, we found a negative serial dependence (i.e., a repulsive effect) of stimuli presented four to ten trials back. This corresponds to a temporal window of ~15.5 to 39 seconds prior to the onset of the current stimulus. However, previous reports of serial dependence in visual perception only investigated the effect of stimuli presented up to three trials back. A possible mechanism that could underlie the negative serial dependence effect is visual adaptation. A well-known adaptation effect in orientation perception is the tilt-aftereffect (Gibson & Radner, 1937). It describes the phenomenon that after a prolonged exposure to an oriented line, which plays the role of the adaptor, a subsequently presented test line is perceived as tilted away from the adaptor's orientation. Thus, positive serial dependence effects and negative adaptation effects are exerting opposite pulls on the current orientation percept. Our current findings are consistent with a scenario in which a given stimulus induces both positive and negative aftereffects, which subsequently decay at different rates. Specifically, the positive aftereffect, pulling subsequent percepts towards the presented orientation, would be initially stronger but would decay much faster than the negative aftereffect. The perceptual bias observed at a given point in time would consequently be a superposition of both aftereffects. While positive serial dependence would outweigh negative adaptation shortly after stimulus offset, this relationship would reverse as more time passes. However, this interpretation must be taken with caution. Although it has been shown that negative aftereffects can persist over a prolonged period of time up to 20 to 30 minutes, very long exposure durations are required for achieving long lasting effects (Magnussen & Johnsen, 1986). Adaptation effects to stimuli presented in the sub-seconds range have been reported (Patterson, Wissig, & Kohn, 2013; Pavan, Marotti, & Campana, 2012), but were of extremely short duration, far from the durations that would be necessary for explaining the current findings. Alternatively, the present effect could be explained by a single serial dependence or adaptation process, which reverses its effects on perception over time. This single process would exert a pull of subsequent percepts towards a presented stimulus for the first couple of seconds after stimulus presentation, which would subsequently change into a repulsive effect as time passes. It remains to be tested, whether the current pattern of effects is caused by one single serial dependence mechanism or by simultaneous activation of positive serial dependence and negative adaptation processes by the same stimuli.

Recently, a study by Chopin and Mamassian (2012) on visual adaptation found a shift from negative aftereffects by stimuli presented in the recent history (up to 3 minutes in the past) to positive aftereffects for stimuli shown in the remote history (~5 to 13 minutes in the past). However, this study capitalised on repeated presentation of the same stimuli in order to induce adaptation effects and thus is difficult to compare to the current study. Importantly, it was criticised that Chopin and Mamassian's findings of positive aftereffects induced by temporally remote stimuli could be explained by simple short-term negative aftereffects and spurious large-range correlations in their stimulus sequences (Maus, Chaney, Liberman, & Whitney, 2013; but see Chopin & Mamassian, 2013). Consequently, one might worry that our negative aftereffects for stimuli presented four to 10 trials back might be similarly caused by a short-term positive aftereffect and random fluctuations of orientations in the stimulus sequence. In order to rule out this possibility, we simulated an artificial observer with serial dependence of various strengths and temporal decays on the stimulus sequences we used in our experiment. We found that the negative aftereffects for more remote stimuli could not be explained by an observer with positive serial dependence alone. Additionally, we simulated artificial observers with combinations of serial dependencies and oblique biases in order to rule out that the negative aftereffects appeared due to a complex interaction of serial dependence, oblique bias and the trial sequences of the experiment. Again, these simulations did not exhibit strong negative aftereffects for more remote stimuli, which were observed in the empirical data (for all simulations see Supplementary Information). Therefore, we conclude that the reported change in polarity of serial dependence is most likely a genuine perceptual effect. However, further experiments will be required to elucidate the underlying mechanisms of this effect. A possible functional role will be discussed in the General Discussion.

3. Experiment 2 – Influence of Sensory Uncertainty on Serial Dependence

Besides the prediction that serial dependence should be object selective, we hypothesised that the degree to which the visual system leverages input of the recent past to smooth perception should depend on the sensory uncertainty associated with previous and current input. This hypothesis was inspired by Bayesian theories of perception that state that our brains combine inherently noisy and ambiguous sensory information with prior knowledge about the world in a probabilistically optimal manner. Some evidence for this was indeed found in Experiment 1. In Experiment 2 we set out to test this hypothesis explicitly. We employed a similar orientation judgment task as in Experiment 1. However, we introduced two levels of stimulus noise (low/high) to manipulate the certainty about stimulus orientation on the previous and current trial. In particular, we used Fourier filtered noise patches for which one can specify distributions of spatial frequencies and orientations to be present in the respective stimuli (Beaudot & Mullen, 2006). According to Bayesian theories, the probabilistically optimal combination of prior knowledge and sensory information predicts that the reliance on prior information should be strongest in cases in which certainty of sensory information is lowest and the certainty of prior information is highest (e.g., see Stocker & Simoncelli, 2006a). If one assumes that recently encountered stimuli serve as prior information, then serial dependence should be strongest in cases where the previous stimulus was of low noise and the current stimulus is of high noise. On the other hand, serial dependence would be expected to be weakest for cases in which the previous stimulus was of high noise and the current stimulus is of low noise. Such a pattern would support the view of serial dependence as a mechanism that takes uncertainty about current stimulus and the perceptual history into account in order to stabilise perception in a probabilistically optimal manner.

3.1 Methods

3.1.1 General procedure

Experiment 2 consisted of three separate sessions. In the beginning of the first session participants were trained on a two-alternative forced choice task (2AFC) that was subsequently used with a staircasing procedure to tailor the noise levels of the stimuli to individual sensitivity. The individually determined noise levels were subsequently used for stimuli in the serial dependence task of the second and third session. Afterwards participants practiced the serial dependence task, which was similar to the task of Experiment 1. Participants were excluded from the experiment after the first session, if their staircasing procedure did not converge to plausible noise parameters (due to chance performance on the 2AFC task). In the second and third session, participants started again with a short practice of the serial dependence task, before performing the main task.

3.1.2 Participants

Twenty participants took part in the experiment. All participants provided written informed consent prior to the start of the experiment. The study was approved by the Radboud University Institutional Review Board. Of the 20 participants, seven were excluded after the first session because of near chance performance during the 2AFC staircasing task, leading to both noise-level stimuli containing virtually no noise. One participant was excluded, because both low- and high-noise stimuli were set to extremely similar and high noise levels by the staircasing procedure, preventing reasonable performance on the subsequent serial dependence task. The remaining 12 participants (11 female, age 21–28 years) reported to have normal or corrected-to-normal visual acuity.

3.1.3 Stimuli

All stimuli of this experiment were generated by filtering white noise in the Fourier domain with an appropriate bandpass filter. The passband of spatial frequencies was defined as a Gaussian with a mean of 1 cycle/° and *SD* of 0.5 cycles/° for all stimuli. The passband for orientations was defined as a von Mises distribution with location parameter μ corresponding to the mean orientation of the respective stimulus and concentration parameter $\varkappa$ reflecting the amount of noise. Effectively, the von Mises distribution determined the distribution of orientations present in the stimulus. A high concentration parameter led to few orientations other than the specified mean orientation to be present in the signal, resulting in low noise stimulus. A low concentration parameter led to a more uniform distribution over orientations, resulting in a noisy stimulus. Concentration parameters $\varkappa$ for low and high noise stimuli were determined for each participant individually with a 2AFC staircasing task described below. After applying the inverse Fourier transform, the root mean square contrast of the filtered noise patches was set to 20 greyscale units, centered on the background colour. The patches were windowed with a Gaussian envelope (1.5° *SD*). Examples of these stimuli are depicted in Figure 4B.

Stimuli were generated with the Psychophysics Toolbox for MATLAB (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) and were displayed on a 24” flat panel display (resolution: 1920 x 1080, refresh rate: 60 Hz). Participants viewed the stimuli from a distance of 51 cm in a dimly lit room, resting their head on a table-mounted chinrest.

3.1.4 2AFC task

In the 2AFC task participants were simultaneously presented with two Fourier filtered noise patches left and right of fixation and had to judge whether the mean orientation of the right patch was tilted more clockwise or anti-clockwise relative to the left patch. The relative orientation difference was either 5 or 20°. During the staircasing phase, concentration parameters $\varkappa$, reflecting the noise in the stimuli, were adjusted such that participants were on average 75% correct for both 5 and 20° relative orientation differences. This resulted in relatively low noise stimuli for the 5° orientation difference and relatively high noise stimuli for the 20° orientation difference.

The sequence of events within each trial is illustrated in Figure 4A. On each trial a central fixation dot (white, 0.6° visual angle diameter) was presented on a mid-grey background. Participants were instructed to maintain fixation. Two Fourier filtered noise patches were simultaneously presented left and right of fixation (6.5° eccentricity) for 500 ms. One of the patches represented the standard with a mean orientation randomly chosen on each trial between 0 and 180°. The other patch represented the probe with a mean orientation tilted either 5 or 20° clockwise or anti-clockwise from the standard. Position of the standard (left or right) and tilt direction of the probe (clockwise or anti-clockwise) was randomly determined in each trial. The concentration parameter $\varkappa$ (noise parameter) was equal for both patches within a trial and varied over trials according to the staircasing algorithm, or was fixed during initial training. 250 ms after the offset of the stimuli, participants could respond whether the mean orientation of the right patch was rotated more clockwise or anti-clockwise relative to the left patch, by pressing the right or left arrow key of a standard keyboard. After the response there was a 2 second inter-trial-interval where only the fixation dot was visible.

Participants initially performed two training blocks with 50 trials each. In this phase the concentration parameters were fixed to $\varkappa = 500$ (very low noise) for trials with a relative orientation difference of 5° and $\varkappa = 6$ (moderately low noise) for trials with a relative orientation difference of 20°. The presentation order of trials with 5 or 20° orientation difference was pseudo-randomised. Further, during training, feedback was given after each trial. The colour of the fixation dot changed to green (red) for 500 ms after correct (incorrect) responses. After that, participants performed the same task, while concentration parameters were adjusted in two interleaved staircase procedures (Quest algorithm, Watson & Pelli, 1983), for 5 and 20° relative orientation difference conditions respectively, such that participants had on average 75% correct responses in both conditions. After 20 initial trials that were discarded, trials of both conditions were presented 200 times each in a pseudo-randomised order, resulting in a total of 400 trials for the staircasing procedure, divided by a self-paced break. No feedback was given during this phase.

The reason for determining the noise levels

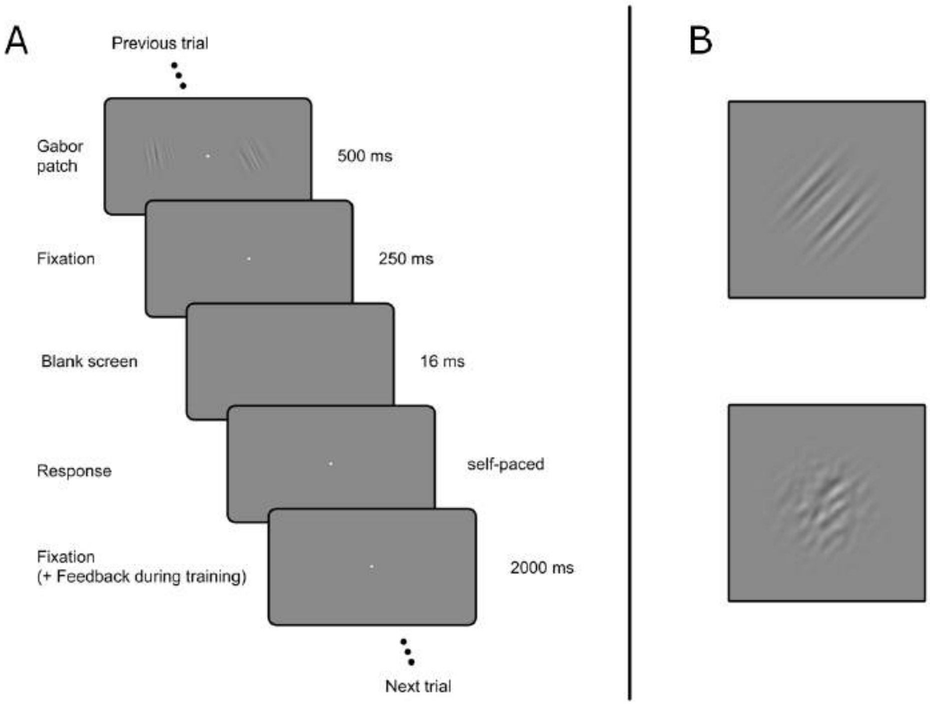

Fig. 4 A. Sequence of events within each trial of the 2AFC task of Experiment 2. Participants simultaneously viewed two Fourier filtered noise patches to the left and right of fixation. The relative orientation difference between the mean orientations in both patches was either 5 or 20° (random sequence). They had to judge whether the mean orientation in the right patch was tilted clockwise or anti-clockwise with respect to the mean orientation in the left patch. The task was used to staircase two individual noise levels for the subsequent serial dependence task. **B.** Group averaged low noise (upper panel) and high noise stimulus (lower panel). In this example both stimuli have the same mean orientation.

individually for each participant, rather than picking two general noise levels, was that individual noise levels promised to reduce the between-subject variability due to different levels of internal noise and differences in sensory encoding of the noisy stimuli.

3.1.5 Serial Dependence Task

The serial dependence task was similar to the task in Experiment 1, with two differences. First, instead of Gabor patches participants were presented with Fourier filtered noise patches. Second, participants were informed about their performance (mean response error) at the end of each block. The sequences of presented stimuli were counterbalanced with respect to the combinations of noise levels on trial n and n−1. That is, stimuli of low/high noise were preceded equally often by stimuli of low/high noise. The stimulus sequences were pseudo-randomised for each participant. In the first session participants practiced on a sequence of 84 trials, divided into two blocks. In the beginning of the second and third session, participants practiced on a sequence of 52 trials, divided into two blocks. In the second and third session, respectively, participants then performed the serial dependence task with 656 trials divided into eight blocks. This resulted in a total of 1,312 trials per participant on the main task that were further analysed.

3.1.6 Data Analysis

Outlier Correction

In the first step of data analysis we excluded those trials in which the response error (shortest angular distance between stimulus orientation and response orientation) was more than three circular standard deviations away from the participant's mean response error. This was done separately for low noise and high noise trials in order to prevent a bias towards removing high noise trails, for which response error distributions had higher variance. The outlier correction was performed in order to exclude trials on which the participant gave random responses due to blinks or attentional lapses during stimulus presentation as well as due to inadvertent responses. After that, response errors were demeaned for each participant to remove general clockwise or anti-clockwise response biases.

Conditioning, Model Fitting and Statistical Testing

As in Experiment 1, we conditioned the response error of the current trial on the relative orientation of the previous trial (1-back conditioning) and on stimuli shown up to 20 trials back (2-, 3-,..., 20-back conditioning). Analogously to Experiment 1 that manipulated spatial frequency of previous and current stimulus, here we obtained four different noise level conditions: low-to-low, high-to-high, low-to-high and high-to-low noise on previous and current trial, respectively. The procedure of computing moving averages and fitting DoG models was done exactly as described in Experiment 1. Statistical testing, in form of permutation tests, was analogous to Experiment 1. That is, we tested serial dependencies on 1- to 20-back stimuli, we conducted a 2 x 2 permutation test to test the influence of noise levels on previous and current trial as well as their interaction, and we performed individual permutation tests between any two noise level conditions.

All data analyses were performed with MATLAB (The Mathworks Inc., Natick, MA, USA) and the CircStat MATLAB toolbox for circular statistics (Berens, 2009).

3.2 Results

3.2.1 Staircasing Procedure

The 12 subjects, that showed successful convergence of the staircasing procedure in the first session, had a mean concentration parameter $\varkappa = 162.49$ for low noise stimuli and $\varkappa = 1.17$ for high noise stimuli. Examples of stimuli with these noise levels are depicted in Figure 4B. Individual concentration parameters can be found in Supplementary Table 1.

3.2.2 Serial Dependence Task

On average 10 of 1,312 trials were excluded per participant (*SD* = 4.34, low noise: 7.08 ± 2.56; high noise: 13.00 ± 3.70). For the remaining data the mean response error was 9.42° ± 1.95 (*SD*). The mean error on high noise trials (12.03° ± 3.27) was significantly larger than the mean error on low noise trials (6.90° ± 1.40; t(11) = 5.51, *p* = .0002, two-sided paired t-test). The mean response time was 2.29 ± 0.70 seconds and there was no significant difference between low and high noise trials (difference: 43 ms; t(11) = 0.16, p = .87, two-sided paired t-test). Thus, the onset of the current stimulus occurred

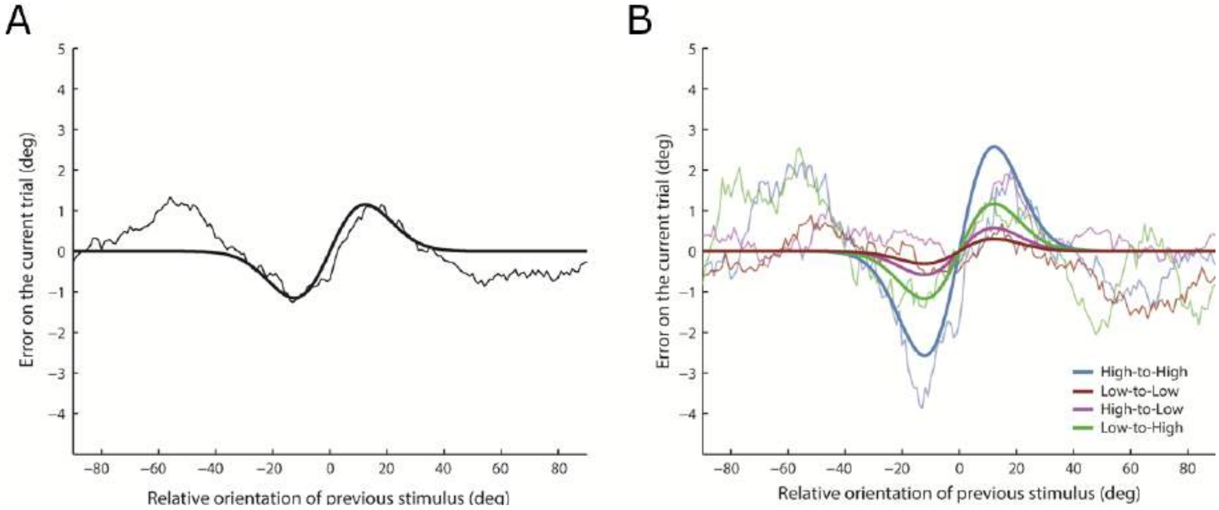

Fig. 5. Serial dependence of orientation perception in Experiment 2: Response errors on the current trial conditioned on the previous stimulus orientation **A.** for all trials irrespective of stimulus noise level, **B.** separately for each combination of stimulus noise level of previous and current stimulus. Positive values on the horizontal axis indicate that the stimulus of the previous trial was oriented more anti-clockwise than the stimulus of the current trial. Positive values on the vertical axis indicate that the response of the current trial was oriented more anti-clockwise than the current stimulus. In **A.** the thin line shows the grand moving average of response errors and the thick line show the best fitting DoG model. In **B.** thin lines show the grand moving averages of response errors for each noise level condition (see figure legend). Thick lines show the best fitting DoG model, respectively.

on average ~4.54 seconds after the offset of the previous stimulus. Further, when splitting the data into noise conditions, each condition contained an average of 315 trials per participant.

We found an attractive serial dependence of the current percept towards the stimulus of the previous trial, which however was not significant after correcting the significance level for multiple comparisons (amplitude: 1.15°, Fig. 5A; $p = .003$, two-sided permutation test, significance level Bonferroni corrected to $\alpha = .0025$). We found a main effect of the current trial's noise level on the strength of serial dependence (amplitude of difference contrast: 1.6°; $p = .005$, two-sided permutation test). However, when processing the data with a moving average window of 30° width, this effect was greatly reduced and not significant (amplitude of difference contrast: 0.67°; $p = .24$, two-sided permutation test). This points to a potentially badly fitting model for this particular contrast moving average, and thus the result has to be interpreted with caution. There was no significant main effect of previous trial's noise level, neither was there an interaction effect. Permutation test between pairs of noise level conditions revealed a significant difference in serial dependencies between the low-to-low noise condition, showing almost no serial dependence, and the high-to-high noise condition, showing the highest serial dependence (amplitude of difference contrast: 2.18°, Fig. 5B; $p = .04$, two-sided permutation test). There was also a significant difference of amplitudes between high-to-high noise condition and high-to-low noise condition (amplitude of difference contrast: 2.84°; $p = .007$, two-sided permutation test). However, this effect was also greatly reduced and non-significant when processing the data with a moving average window of 30° width (amplitude: 1.1°; $p = .34$, two-sided permutation test), again pointing to instabilities in model fitting.

With respect to the time course of serial dependence we found a similar pattern to that of Experiment 1. After apparent positive serial dependencies on previous and second previous trials' stimuli (2-back amplitude: 0.55°; non-significant, $p = .006$, two-sided permutation test, Bonferroni corrected significance level), we found a significant repulsive effect of stimuli, which were presented 6 trials back in the past (amplitude: −0.78°, Fig. 6; $p = 0.0005$, two-sided permutation test, Bonferroni corrected significance level).

3.3 Discussion

In Experiment 2, we investigated the hypothesis that the degree to which the visual system leverages input of the recent past to smooth perception should rely on the sensory uncertainty associated with previous and current input. We found only partial support for this hypothesis. The strength of serial dependence appears to be modulated by the amount of noise in the current visual input: serial dependence was stronger when the current stimulus contained a high amount of noise than

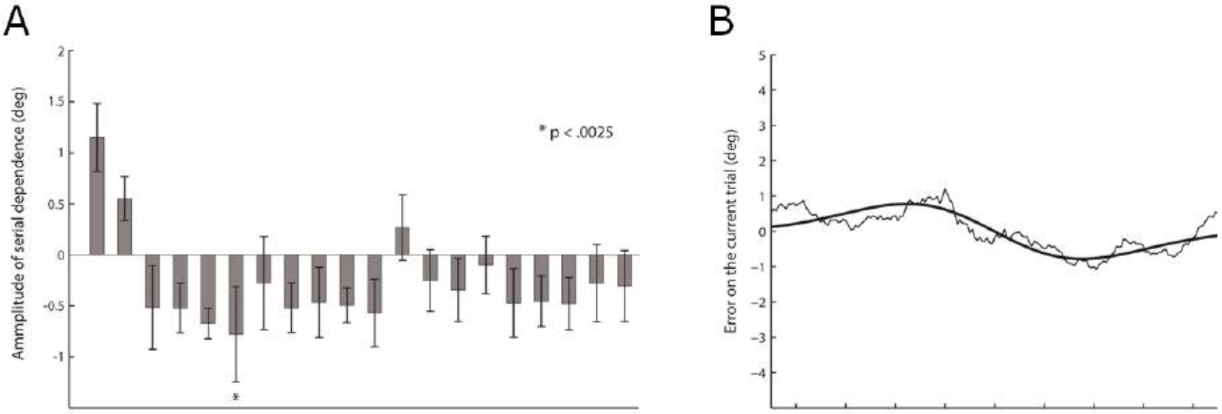

Fig. 6. Temporal dynamic of serial dependence in Experiment 2: **A.** Strength and sign of serial dependence (DoG amplitude parameter) on n-back stimuli; while the current percept is attracted by the previous stimulus (non-significant after Bonferroni correction for multiple comparisons), it is repelled by stimuli seen 6 trials before. Error bars represent 1 s.d. of the bootstrapped amplitude distribution. **B.** Current response error conditioned on the 6-back stimulus orientation. A repulsive effect of the 6-back stimulus on the current percept is observable. The thin line shows the grand moving average of response errors and the thick line show the best fitting DoG model.

when it contained virtually no noise. However, against our prior hypothesis we found that serial dependence was strongest when the previous stimulus also contained a high amount of noise. The latter finding is inconsistent with a model of serial dependence, in which the certainty about the recent perceptual history is maintained and sensibly taken into account when generating the current percept. Further, we confirmed the change in polarity of serial dependence for stimuli presented further back in time that was discovered in Experiment 1. Finally, next to attractive effects of previous stimuli on the current percept, we report repulsive effects of previous stimuli that had a large orientation difference with respect to the current stimulus.

The finding that serial dependence in visual perception is stronger when the current visual input contains a high amount of noise is in line with probabilistic models of perception: when the current visual input is disrupted by noise, the brain relies more strongly on prior information, in the form of previously presented stimuli, when generating the current percept. On the other hand, when the current input is of low noise, the brain relies more on sensory information, and prior information has less impact on the percept of the current stimulus. Interestingly, previous studies on perception of ambiguous or bistable images have similarly reported effects of previous visual input or previous percepts on the current percept and argued in favor of a perceptual stabilisation mechanism (Brascamp, Knapen, Kanai, Van Ee, & Van den Berg, 2007; De Jong, Brascamp, Kemner, Van Ee, & Verstraten, 2014; De Jong, Kourtzi, & Van Ee, 2012). It must be noted, however, that the high noise stimuli in the current experiment were far from ambiguous and participants maintained a relatively good performance on these stimuli. It is an open question whether the history-dependent phenomena in perception of ambiguous stimuli studied in previous experiments and strong serial dependence for noisy stimuli used in the current experiment are based on the same underlying neural mechanisms.

Curiously, we found that serial dependence of the percept of a high noise stimulus was strongest when the previous stimulus also contained a high amount of noise. This finding contradicts our hypothesis that a more uncertain perceptual history should exert a weaker bias on current percept, than a clear and certain perceptual history. Thus, the finding poses a challenge to a Bayesian view on serial dependence. Yet, it remains speculative what caused the reported effect. A recent study by Akaishi, Umeda, Nagase, and Sakai (2014) found that perceptual decisions made on ambiguous perceptual input tended to be repeated on following trials. This effect could not be explained by biases in motor response, sensory processing or attention and was best explained by a learning mechanism of past choices that influenced future choices. It could be possible that in our experiment participants also exhibited such a learning effect that is facilitated by noisy, ambiguous stimuli. However, it must be noted that in Akaishi et al.'s study there were only two response options, whereas in the current study participants had to give a continuous response. Further, in the current experiment the low noise stimuli were not ambiguous, but still contained a sufficient amount of information to give relatively accurate responses. Therefore, it is questionable whether the learning mechanisms proposed in Akaishi's et al. study could explain the current finding.

Another possibility is that increasing stimulus noise selectively reduced negative aftereffects elicited by these stimuli and thus led to a dominance of positive aftereffects. Indeed, previous studies on bistable perception indicate that prior unambiguous stimuli can have positive or negative effects on subsequent bistable perception depending on their 'energy', with low energy stimuli (short duration, low contrast) leading to positive effects and high energy stimuli (long duration, high contrast) leading to negative effects (Brascamp et al., 2007; Pearson & Brascamp, 2008). However, it is not clear whether our noise manipulation targeted similar 'energy' dependent processes.

A third possible explanation for the current result could be a differential engagement of top-down processes for low and high noise stimuli, which in turn could lead to differences in serial dependence. For instance, it is possible that high noise stimuli required more attention than low noise stimuli in order to maintain a satisfactory level of behavioural performance. It has been previously shown that the strength of serial dependence is modulated by spatial attention (Fischer & Whitney, 2014). If indeed, stimuli containing higher amounts of noise would elicit more attention than stimuli containing little noise, similar modulations of serial dependence strengths might be expected. One way of testing this hypothesis could be to make the orientation of the previous stimulus task-irrelevant, thereby presumably abolishing differential amount of attention paid to previous low-noise and previous high-noise stimuli. If a modulation of serial dependence strength would still be observed, an explanation relying on differential amounts of attention due to task demands may be ruled out. In summary, the origin of stronger serial dependence on previous high noise stimuli remains inconclusive.

4. General Discussion

The current study investigated whether serial dependence exhibits properties, which would support its proposed role in stabilizing perception. More specifically, we hypothesised that, if serial dependence would serve to stabilise perception, it would optimally be object selective and should weight visual input of the past and present according to their associated uncertainty when generating the current percept.

Results of Experiment 1 suggest that serial dependence is not object selective. In particular, changes in object identity signaled via changes in the secondary feature dimension of spatial frequency do not affect the strength of serial dependence in orientation perception. Instead, the results suggest a dynamic weighting of previous and current input according to the informational content or uncertainty about task relevant information. In Experiment 2, we manipulated stimulus noise in order to test more directly whether uncertainty inherent to previous and current stimulus is taken into account when generating the current percept. We found that serial dependence was stronger when the noise in the current stimulus was high, especially so when the previous stimulus also contained a high amount of noise. This finding is partially inconsistent with our prior hypothesis and results from Experiment 1. As mentioned in the Discussion section of Experiment 2, possible explanations include a stronger tendency to repeat choices that were based on uncertain sensory evidence, a difference in negative aftereffects induced by low and high noise stimuli and different amounts of attention paid to low and high noise stimuli. Further, the manipulations utilised in Experiments 1 and 2 were quite different in nature. Therefore, it may be possible that our manipulation targeted completely different aspect of serial dependence.

Interestingly, in both experiments we found a change in polarity of serial dependence over time: While the current percept was biased towards stimuli seen just before, it was biased away from stimuli seen in the more remote history (4- to 10-back stimuli with exceptions in Experiment 1 and 6-back stimulus in Experiment 2). Statistical tests and simulations of artificial observers indicate that this effect was neither due to a general bias of perception to the oblique orientations nor due to spurious correlations in our trial sequences. A possible functional role of this effect could be to establish a balance between stabilizing perception and maximizing perceptual sensitivity according to temporal dynamics of the environment. That is, while our environment does not change over very short timescales, changes over longer periods of time can be expected. Therefore, smoothing perception over short periods, but maximizing change sensitivity in the long run, may present the most optimal strategy for visual perception. It will be important to further investigate this dynamic pattern of serial dependence and to explore the underlying mechanisms in future studies.

Finally, one may argue that the effects reported in the current study, which all relate to the response error on the current trial, might not necessarily be perceptual in nature, but could also arise on the response level. We cannot address this concern directly with the current set of experiments. However, it has been convincingly shown by Fischer and Whitney (2014) that serial dependence in orientation perception is of perceptual nature. Thus, we assume that the modulations of serial dependence found in the current study are most likely also occurring on the perceptual level.

In summary, we found only little evidence for properties of serial dependence, which would be desirable for stabilizing perception. The hypothesis of object selectivity could not be confirmed and the prediction of a probabilistic weighting of previous and current input according to sensory uncertainty could only be partially confirmed. However, we found an interesting temporal dynamic of the polarity of serial dependence, which could play a functional role in stabilizing perception. Thus, while serial dependence could in principle serve to stabilise visual perception, it appears to lack properties which would allow to do so in an optimal manner.

5. References

Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous Mechanism of Internal Choice Estimate Underlies Decision Inertia. *Neuron, 81*(1), 195–206.

Barlow, H. B. (1961). Possible Principles Underlying the Transformations of Sensory Messages. In W. Rosenblith (Ed.), *Sensory Communication* (pp. 217–234). Cambridge, Massachusetts, US: MIT Press.

Beaudot, W. H. A., & Mullen, K. T. (2006). Orientation discrimination in human vision: Psychophysics and modeling. *Vision Research, 46*(1-2), 26–46.

Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software, 31*(10), 1–21.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology, 203*(1), 237–260.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*(4), 433–436.

Brascamp, J. W., Knapen, T. H. J., Kanai, R., Van Ee, R., & Van den Berg, A. V. (2007). Flash suppression and flash facilitation in binocular rivalry. *Journal of Vision, 7*(12), 1–12.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197*(3), 551–566.

Chopin, A., & Mamassian, P. (2012). Predictive Properties of Visual Adaptation. *Current Biology, 22*(7), 622–626.

Chopin, A., & Mamassian, P. (2013). Response: Genuine long-term positive aftereffects. *Current Biology, 23*(10), R439–R440.

Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences of the United States of America, 111*(21), 7867–7872.

De Jong, M. C., Brascamp, J. W., Kemner, C., Van Ee, R., & Verstraten, F. A. J. (2014). Implicit Perceptual Memory Modulates Early Visual Processing of Ambiguous Images. *Journal of Neuroscience, 34*(30), 9970–9981.

De Jong, M. C., Kourtzi, Z., & Van Ee, R. (2012). Perceptual experience modulates cortical circuits involved in visual awareness. *European Journal of Neuroscience, 36*(12), 3718–3731.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience, 17*(5), 738–743.

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences, 14*(3), 119–130.

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology, 20*(5), 453–467.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience, 14*(7), 926–932.

Kleiner, M., Brainard, D. H., & Pelli, D. (2007). "What's new in Psychtoolbox-3?" Perception 36 ECVP Abstract Supplement.

Liberman, A., Fischer, J., & Whitney, D. (2014). Serial Dependence in the Perception of Faces. *Current Biology, 24*(21), 2569–2574.

Magnussen, S., & Johnsen, T. (1986). Temporal aspects of spatial adaptation. A study of the tilt aftereffect. *Vision Research, 26*(4), 661–672.

Maus, G. W., Chaney, W., Liberman, A., & Whitney, D. (2013). The challenge of measuring long-term positive aftereffects. *Current Biology, 23*(10), R438–R439.

Patterson, C. A., Wissig, S. C., & Kohn, A. (2013). Distinct Effects of Brief and Prolonged Adaptation on Orientation Tuning in Primary Visual Cortex. *The Journal of Neuroscience, 33*(2), 532–543.

Pavan, A., Marotti, R. B., & Campana, G. (2012). The temporal course of recovery from brief (sub-second) adaptations to spatial contrast. *Vision Research, 62*, 116–124.

Pearson, J., & Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences, 12*(9), 334–341.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*(4), 437–442.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24*(1), 1193–1216.

Stocker, A., & Simoncelli, E. P. (2006a). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience, 9*(4), 578–585.

Stocker, A., & Simoncelli, E. P. (2006b). Sensory adaptation within a bayesian framework for perception. In Y. Weiss, B. Schölkopf, & J. Platt (Eds.), *Advances in neural information processing systems* 18 (pp. 1289–1296). Cambridge, Massachusetts, US: MIT Press.

Tomassini, A., Morgan, M. J., & Solomon, J. A. (2010). Orientation uncertainty reduces perceived obliquity. *Vision Research, 50*(5), 541–547.

Vilares, I., & Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences, 1224*(1), 22–39.

Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120.

Wei, X.-X., & Stocker, A. (2015). A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. *Nature Neuroscience, 18*(10), 1509–1517.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences, 10*(7), 301–308.

Frontal Theta in Control: Overcoming Biased Learning or Biased Choice?

Jessica I. Määttä[1]
Supervisors: Jennifer Swart[1], Hanneke den Ouden[1], Roshan Cools[1,2]

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

Our decisions are shaped by a coupling between value (reward/punishment) and action (approach/avoid). Thus, motivation strongly influences the actions we make, by invigorating behaviour in order to gain rewards, and inhibiting behaviours in order to avoid punishment. This influence can be beneficial, but could also lead to maladaptive choices when motivation interferes with goal-directed behaviour requiring us to act against our automatic bias. It is unclear whether this coupling is the result of a bias in choice or a bias in learning. Resolving motivational conflict when value anticipation interferes with goal-directed behaviour has been proposed to be resolved by midfrontal theta power. The proposed role of midfrontal theta power is to suppress the influence of these biases on our choices to reach desirable outcomes. Here we use a novel Go/NoGo reinforcement learning task, where outcome valence and action requirement are orthogonalised, while recording electroencephalography (EEG) in healthy human participants. We aimed to disentangle whether the influence of motivational valence on action selection is the result of altered choice tendencies, or an effect of biased learning. We further aimed to establish the role of midfrontal theta power in motivational conflict, whether it is involved in suppressing the influence of choice biases and/or biases in learning. Our results indicate that our actions are driven by motivational effects on choice, rather than a learning bias. We also establish that midfrontal theta power is involved in overcoming the influence of these motivational effects on our choices.

Keywords: conflict, cognitive control, motivational biases, reinforcement learning, theta power

---

Corresponding author: Jessica Määttä; **E-mail:** jessica.i.maatta@gmail.com

1. Introduction

In many daily life situations it is optimal to make choices that will maximise reward and minimise, or avoid, punishment. To accomplish this we have automatic behavioural, or 'Pavlovian', biases, which couple outcome valence (i.e., reward/punishment) with action execution (i.e., invigoration/suppression). This coupling allows us to associate appropriate behavioural responses with the promise of reward or the threat of punishment by invigorating behaviours in the anticipation of reward and suppressing behaviours in order to avoid punishment (Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013; Chiu, Cools, & Aron, 2010; Crockett, Clark, & Robbins, 2009; Geurts, Huys, & den Ouden, 2011; Guitart-Masip et al., 2012; Guitart-Masip, Duzel, Dolan, & Dayan, 2014; Hershberger, 1986; Niv, Daw, Joel, & Dayan, 2007; Robinson, Cools, & Sahakian, 2012). This is a computationally efficient approach, presumably shaped throughout evolution (Daw, Niv, & Dayan, 2005; Dayan, Niv, Seymour, & D. Daw, 2006; Guitart-Masip et al., 2014), and is beneficial when the motivational bias steers your actions in the direction of the correct instrumental response.

However, this coupling, between outcome valence and action execution, can also interfere with our goal-directed behaviour in situations where value anticipation and the required action are incongruent with our automatic, Pavlovian, biases (Cavanagh et al., 2013; Geurts et al., 2011; Guitart-Masip et al., 2011, 2012; Talmi, Seymour, Dayan, & Dolan, 2008), resulting in undesired and suboptimal outcomes. In these situations, we need to be able to overcome our automatic response bias in order to make the choice leading to the most desirable outcome (e.g., avoiding to approach a reward in order to obtain the reward). Thus, we need to detect and resolve the motivational conflict occurring when value anticipation interferes with contingent instrumental responses to accomplish this (Cavanagh et al., 2013; Cavanagh, Masters, Bath, & Frank, 2014). The need for cognitive control, for example, when encountering conflict between response tendencies, has been proposed to be signaled by the prefrontal cortex (PFC; Ridderinkhof, Van den Wildenberg, Segalowitz, & Carter, 2004). Specifically, the need for cognitive control appears to be computed and communicated through cortical theta-band oscillations generated in midcingulate cortex (MCC; Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, & Allen, 2012).

Cortical theta-band oscillations are proposed to implement adaptive control by organizing neuronal processes across broad networks in the brain (Cavanagh & Frank, 2014), and the need for control is reflected by a power enhancement in the theta band (Cavanagh & Frank, 2014; Cavanagh, Zambrano-Vazquez, et al., 2012). Theta power has been shown to be predictive of both learning and adjustment of behaviour (Cavanagh, Frank, Klein, & Allen, 2010; Van de Vijver, Ridderinkhof, & Cohen, 2011). Furthermore, midfrontal theta power has been argued to be involved in resolving motivational conflict, specifically by overcoming Pavlovian biases when interfering with a instrumentally optimal action (Cavanagh et al., 2013). The influence of motivational biases on choice differs between participants, and less influence of biases on choice has been associated with greater midfrontal theta power (Cavanagh et al., 2013). Thus, midfrontal theta power does not only seem to reflect the *need* for prefrontal control, but also the *recruitment* of prefrontal control over instrumental processes, allowing us to act in order to avoid punishment and to suppress behavioural responses in order to obtain rewards (Cavanagh et al., 2013).

The effect of motivational valence on action has been proposed to reflect altered choice tendencies (Cavanagh et al., 2013; Guitart-Masip et al., 2012). Thus, the anticipation of reward can influence the vigour with which our actions are performed, making us more likely to approach reward and to withhold or suppress a response in order to avoid punishment. However, the effect of motivational valence on action could also reflect an effect of biased learning, such that assigning credit to specific actions is easier when anticipating reward than when avoiding punishment. Consequently, if the influence of motivational valence on actions would be reflected by altered choice tendencies and/or effects of biased learning, this could potentially refine and improve the understanding of the role of midfrontal theta power in motivational conflict.

In the current study we used a novel design that allows us to assess whether the effect of motivational valence on learning and choice is indeed related to behavioural activation or also relates to the ability to assign credit to specific actions. The aim of our study was to establish whether these well-known influences of motivation on action are due to effects on choices or of biased learning. We then aimed to establish the role of midfrontal theta power in motivational conflict, specifically, whether midfrontal theta power indeed is involved in suppressing choice tendencies and/or is involved in suppressing biases in learning.

2. Methods

2.1 Participants

Thirty-four young adults (age range 18–30, M_{age} = 23.2, *SD* = 3.6; 27 females) participated in the study. All participants reported normal or corrected-to-normal vision. Exclusion criteria included any history of neurological or psychiatric disorders, use of psychotropic drugs, pregnancy, claustrophobia, colour blindness, and left-handedness. Participants were compensated with €20 or course credits for their participation, as well as a monetary bonus of €0–€5 depending on performance (M_{bonus} = 2.12, *SD* = 1.70). Electroencephalography (EEG) data of four participants were excluded, two participants due to excessive movement artifacts, and two participants due to problems during preprocessing. Thus, EEG data of 30 participants were included in the EEG analyses (age range 18–29, M_{age} = 22.87, *SD* = 3.44; 24 females). Consent and procedures were approved by the local ethics committee in accordance with the declaration of Helsinki.

2.2 Experimental procedure

During the experiment, the participant was seated in an electricity-shielded room in front of a computer and was provided with two identical button boxes, one for each hand. The distance between the participant and the computer screen was kept constant across participants, and their chair was adjusted into a fixed position. The experimenter was seated outside the electricity-shielded room and could monitor the experiment via video camera and microphone. During the session, the participant performed two sessions of a Go/NoGo reinforcement learning task and a forced-choice transfer phase, which will be described in detail below.

2.3 Experimental paradigm

2.3.1 Go-NoGo reinforcement learning task

To establish whether the influence of motivational valence on choice is due to a Pavlovian bias in action execution or biases in (learning of) instrumental action-outcome contingencies, we used a modified version of a Go/NoGo reinforcement learning task, where valence and action requirements were orthogonalised (Cavanagh et al., 2013; Guitart-Masip et al., 2011, 2012). Previous studies have shown reliable biases of learning to Go/NoGo in anticipation of reward and when avoiding punishment, respectively (Cavanagh et al., 2013; Guitart-Masip et al., 2011, 2012). We extended this paradigm by providing an additional "Go" option, which enabled us to assess not only the ability to learn whether to act or not, but also specifically which action to make. This allowed us to establish whether the influence of outcome anticipation relates to biases in choice, nonspecific invigoration in the presence of reward (i.e., enhancing all Go responses), and/or reflects a bias in learning (i.e., it is easier to assign credit to Go actions when anticipating reward).

During the Go/NoGo reinforcement learning task, on each trial participants were presented with a gem-shaped cue. Participants could then make one of two active responses: a right button press (Go right) or a left button press (Go left), or withhold from making an active response (i.e., "NoGo"). When participants had made their choice, they were provided with feedback (i.e., reward, punishment, or a neutral outcome). Based on the outcome, participants could learn the required action for each gem. In total, participants were presented with eight different cues to learn. For the purpose of analysis, the eight cues were collapsed into four categories based on valence (win/avoid) and action requirement (Go/NoGo): Go-to-win, Go-to-avoid, NoGo-to-win, and NoGo-to-avoid (Fig. 1).

Participants were instructed that each gem was paired with a particular response that would lead to the most optimal outcome most of the time, and that they were required to learn these action-outcome contingencies by trial and error. They were also informed that they could obtain a reward (or nothing) from half of the gems and get punished (or nothing) from the other half of the gems, but were not informed about which gems belonged to which category. Thus, the possible outcomes for the rewarding gems (win cues) were either neutral or a reward, while the possible outcomes for punishing gems (avoid cues) were either neutral or a punishment. Participants were encouraged to try out all three response options to best learn how to get the most optimal outcome for each gem most of the time. The feedback contingency was 80/20, that is, participants were rewarded or avoided punishment when they made the correct response in 80% of the cases. Each gem was presented 40 times across four blocks resulting in 320 trials, with a short break in between each block. The task was performed twice within the same EEG session, with

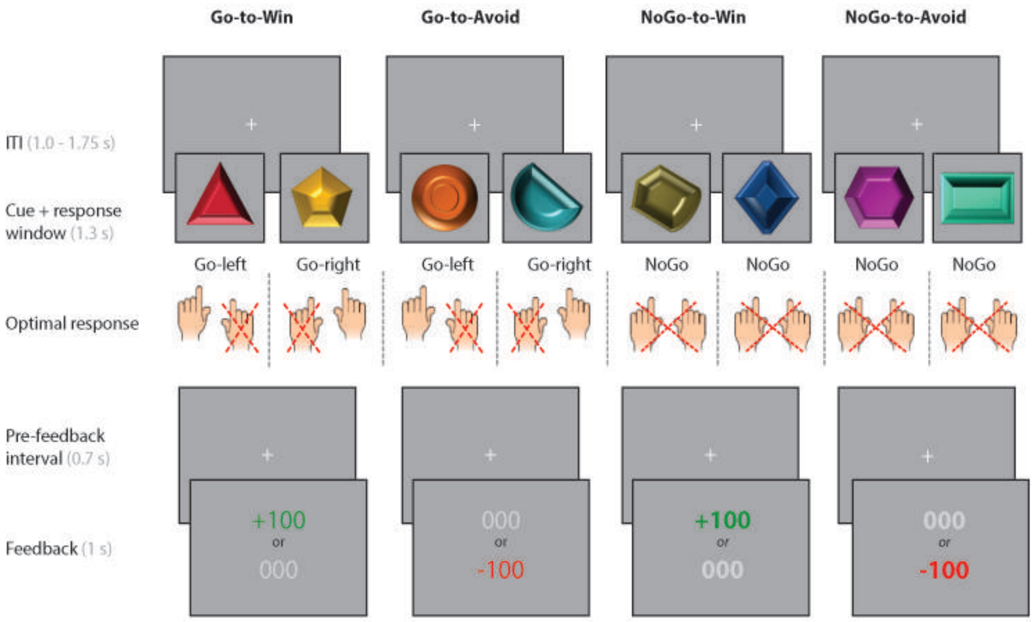

Fig. 1. On each trial participants were presented with a fixation-cross, followed by a gem and feedback. The participants were required to make one of three possible choices while the gem was on the screen (i.e., Go left, Go right, or to not give a response), and were provided with feedback after the gem disappeared. If no response was given during the 1.3 s response window, the response was labeled as a NoGo response.

two different sets of gems, leading to a total of 640 trials per participant. The order of the sets of gems was counterbalanced across subjects, as well as the action-outcome contingencies.

2.3.2 Forced-choice transfer phase

Once the participant had performed the Go/NoGo reinforcement learning task twice they completed a forced-choice transfer phase. They were presented with gems in pairs of two, where each category of gems was paired with each of the other categories, and participants were asked to select the gem they found most rewarding out of the two. No feedback was provided. Only the gems from the most recently presented set were presented in the forced-choice transfer phase. Each paring was presented twice, with in total 12 trials per category pairing, resulting in 48 trials. The forced-choice transfer phase was used in order to indicate the relative preferences between the different gems as a measure of the learned value of the gems (i.e., explicit Pavlovian value).

2.4 Behavioural measures

The collected behavioural data consisted of information about whether participants made a Go response (Go/NoGo), which Go response they made (Go left/Go right), and reaction time (RT). Thus, in our novel design, we were able to establish, not only the subjects' ability to learn "whether to go" in response to a cue (i.e., reflecting behavioural activation), but also "how to go" (i.e., whether they could learn to make the Go response that would lead to the desired outcome 80% of the time). This allowed us to assess the effect of motivational valence on learning and choice related to behavioural activation (i.e., whether or not to go), but also on the ability to assign credit to specific actions (i.e., how to go). Thus, we could disentangle whether motivational bias effects in the Go/NoGo reinforcement learning tasks used in previous studies (Cavanagh et al., 2013; Guitart-Masip et al., 2011, 2012) were a result of altered choice tendencies or an effect of learning/credit assignment. Specifically, whether they were more likely to go for reward and less likely to go to avoid punishment, or whether they were better at learning to make a Go action for a reward than when avoiding punishment.

2.5 EEG data acquisition and preprocessing

EEG data were acquired at a 500 Hz sampling rate from 64 channels using an actiCAP system (Brain Products) placed according to an equidistant arrangement. Two of the electrodes were placed under and above the left eye for vertical electrooculogram (EOG), and two to the left and right side of the left and right eye respectively for horizontal EOG. The electrode at the left mastoid was used for referencing. The ground (GND) was placed on the forehead. The EEG data was preprocessed and analysed with the Fieldtrip software toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) in MATLAB (MathWorks). Offline the EEG data was re-referenced to the weighted average of the left and right mastoid, high-pass filtered at 0.5 Hz, and epoched. Epochs were locked to the cue onset (−1.5 to 4.5 s). Each epoch contained 1.5 seconds before cue onset and after feedback offset to avoid edge artifacts when performing time-frequency convolution. All epochs were linear baseline corrected with a 200 ms pre-cue baseline (−0.2 to 0 s). All epochs were visually inspected for identification of bad channels for interpolation and artifact removal. Interpolation of bad channels was performed before artifact removal and only applied when artifacts in the channel were due to electrode malfunctioning during recording. Interpolation was performed for eight participants and had to be applied only to one channel per participant. Epochs containing artifacts or electromyographic (EMG) activity that was not related to eye blinks were manually rejected, resulting in an average of 98 rejected epochs per participant (*SD* = 41.69), corresponding to an average of 14% trial rejections per participant. On the remaining epochs an independent component analysis (ICA) was performed. Components related to eye blinks or artifacts that were clearly distinguishable from brain activity were removed from the data, resulting in an average of 1.63 rejected components (range = 1–4) per participant. The horizontal EOG signal was left out of the ICA analysis and further analyses. If trials rejected exceeded 30%, the participant was excluded from analysis (see section 2.1 Participants). Lastly, the surface Laplacian of the EEG data was estimated; this method was applied in order to diminish volume conduction trends, by filtering out distant effects while attenuating local effects (Oostendorp & Van Oosterom, 1996; Srinivasan, Winter, Ding, & Nunez, 2007).

2.6 EEG time-frequency decomposition

For condition-specific activation, time-frequency convolution was performed stimulus-locked with a set of Morlet wavelets with frequencies ranging from 1 to 50Hz in 40 logarithmically scaled steps. The width of the wavelet was set to four cycles, in order to have a good tradeoff between temporal and frequency resolution. Wavelet convolution is well suited when interested in localizing frequency information over time (Cohen, 2014), and has previously been used when studying conflict theta power (Cavanagh et al., 2013; Cavanagh et al., 2012; Hajihosseini & Holroyd, 2013; Van Driel, Ridderinkhof, & Cohen, 2012). The wavelet convolution resulted in power values per time-frequency point for each trial. Resulting power values were trial-averaged per condition, and normalised by conversion to a decibel scale ($10*\log10[\text{power}_{tf}/\text{baseline power}_{f}]$) in order to allow direct comparison between frequency bands, electrodes, and participants. The baseline for each frequency consisted of the condition averaged power from −0.25 to −0.05 s before cue onset.

2.7 Statistical data analyses

The goal of our analyses was to establish whether the effect of motivational valence on learning and choice is due to a Pavlovian bias in action execution or a learning/credit assignment bias. We also aimed to establish whether midfrontal theta power is involved in suppressing the influence of Pavlovian biases on choices and/or in suppressing the influence of biased learning. To accomplish these aims we analysed both behaviour (i.e., Go responses, accuracy, and RT) and brain activity (i.e., midfrontal theta power), as well as the relationship between them. Both behaviour and brain activity was analysed by using repeated measures analysis of variance (ANOVA), as well as Spearman's rank correlation to look at the relationship between variables. Post-hoc dependent-sample *t*-tests were performed to explore significant interaction effects. All EEG analyses were performed on trial-averaged data.

2.7.1 Behavioural data analyses

To be able to look at the influence of valence and action requirement on proportion Go actions and accuracy, these were analysed with repeated measures ANOVA with the within-subjects factors Valence (reward vs. punishment) and Action requirement (Go vs. NoGo). We entered beta-coefficients of each individual in the repeated measures ANOVA, acquired from a logistic regression with the 2 x 2 factors Valence (reward vs. punishment) and Action requirement (Go vs. NoGo) performed at the individual level. Beta-coefficients that proved to be poorly fitted to the data were considered outliers and removed from further analysis.

In order to establish whether motivational valence altered choice tendencies and whether participants learned the action requirements of the cues, we looked at the influence of valence and action requirement on proportion Go actions across all conditions. We expected to see an effect of valence on behavioural activation, resulting in more Go actions on appetitive cues and fewer Go actions on aversive cues as a result of reward-based invigoration and punishment-based suppression, respectively. If participants learned the action requirements we expected to see more Go actions on cues requiring Go responses, compared to cues requiring no active response.

Moreover, to investigate whether Pavlovian biases interfere with learning the correct Go action (i.e., learning how to go) we looked at the influence of valence and action requirement on proportion of accurate responses across all conditions. If Pavlovian biases interfere with learning to assign credit to actions incongruent with these biases, we would expect to see an interaction effect of valence by action requirement on learning how to go. This would result in higher accuracy when action requirements are congruent with Pavlovian biases. That is, we would expect better performance at Go-to-Win and NoGo-to-Avoid cues, compared to Go-to-Avoid and NoGo-to-win cues.

The same analysis procedure was used for RT data, but beta-coefficients were acquired using a linear regression. In order to improve normality of the non-normal RT distribution a log transformation was performed. All further analyses were performed on log transformed RT. RTs deviating more than 2.5 standard deviations from the mean were excluded from all RT analysis. We expected motivational valence to influence RTs, such that avoiding punishment would slow responses compared to when anticipating reward (Crockett et al., 2009).

Furthermore, our design allowed us to look at the influence of motivational valence on accurate Go actions on Go cues separately. This enabled us to establish whether the influence of motivational valence was a result of altered choice tendencies or an effect of learning/credit assignment. This was assessed by comparing the proportion of accurate Go actions made on appetitive Go cues [correct-Go/all Go actions on win cues] with the proportion of accurate Go actions on aversive Go cues [correct-Go/all Go actions on avoid cues] with a dependent-sample *t*-test. We would expect altered choice tendencies to result in more Go actions following reward, but not necessarily in accurate Go actions, while punishment would result in fewer Go actions, but likely accurate Go actions when these were made. Thus, the proportion of accurate Go actions would be lower on appetitive Go cues compared to the proportion of accurate Go actions on aversive Go cues. We would expect an effect of learning/credit assignment to result in the reverse pattern, where reward would increase the likelihood of making the accurate Go action when going for reward, but that it would be more difficult to learn the accurate Go action when avoiding punishment. Thus, the proportion of accurate Go actions would be higher on Go-to-win cues, compared to the proportion of accurate Go actions on Go-to-avoid cues.

2.7.2 Pavlovian and Instrumental bias measures

In order to be able to look at individual differences in the extent to which participants relied on Pavlovian biases in their choices and whether they learned better from reward or punishment, two different behavioural measures were created. This was done to quantify participants' dependency on Pavlovian biases in their choices (i.e., a Pavlovian bias score; Cavanagh et al., 2013), and/or whether participants demonstrated a learning/credit assignment bias (i.e., an instrumental bias score). These measures also enabled us to look at the relation between midfrontal theta power and behaviour (see section 2.7.5).

The Pavlovian bias measure reflected to what extent participants were more likely to go for rewards and to what extent they were less likely to go to avoid punishment. This measure was computed as the average of the proportion of Go responses on an appetitive cue (out of the total number of Go actions) and the proportion of NoGo responses on an aversive cue (out of the total number of NoGo actions; Cavanagh et al., 2013). A higher Pavlovian bias score indicated a larger dependency on the bias in one's choices:

$$Pavlovian\ bias = \left(\frac{Win(Go)}{all\ Go} + \frac{Avoid(NoGo)}{all\ NoGo}\right)/2$$

The instrumental bias measure quantified the presence and direction of a learning/credit assignment bias by looking at difference in proportion accurate Go actions between appetitive Go cues and aversive Go cues. Thus, the instrumental bias measure indicated whether it was easier for participants to assign credit to Go actions when anticipating reward, compared to when avoiding punishment, while taking a potential go bias into account (i.e., by accounting for the total Go actions made on the respective cue). A positive value indicated a learning bias from reward, and a negative value indicated a learning bias from punishment. A value of zero would indicate no difference in learning from appetitive and aversive cues:

$$\text{Learning bias} = \frac{Go2win(correct|Go)}{Go2win(all\ Go)} - \frac{Go2avoid(correct|Go)}{Go2avoid(all\ Go)}$$

Previous research has shown that participants exhibit the same forced choice preferences when asked to indicate their relative preference between the different cues, regardless of whether they managed to learn the action-outcome contingencies or not (Cavanagh et al., 2013). In the current study the results of the forced-choice transfer phase were not analysed due to time-constraints. Thus, we could not establish whether we replicate their results.

2.7.3 EEG time-frequency window of interest selection

To establish the role of midfrontal theta power in motivational conflict, we compared midfrontal theta power between conditions (i.e., Go-to-win, Go-to-avoid, NoGo-to-win, NoGo-to-avoid). In order to perform these analyses, we selected a time-frequency window of interest. The selection of the time-frequency window and electrodes of interest was guided by a combination of previous findings as well as the characteristics of our data. The selection was orthogonal to the contrast of interest, and performed on the grand averaged power calculated across all participants and all trials, at subsequent time points between 0 and 1 s, separated by 50 ms. Based on previous research on conflict/control theta power (Cavanagh et al., 2013; Cavanagh, Figueroa, Cohen, & Frank, 2012; Frank et al., 2015), we expected the frequency band of interest to be around 4–8 Hz, with strongest activity at midfrontal electrodes. Topographical plots revealed the strongest activity at electrode 2 and 8, roughly corresponding to electrodes FCz and Fz, at 3–7 Hz. The time-window of interest was defined based on the peak-activity of the time-frequency representation of the selected electrodes, revealing a peak activity between 0.45–0.65 s post-cue-onset (Fig. 2A). Thus, our time-window overlaps with the time of conflict/control theta power observed in a similar reinforcement learning paradigm based on a similar contrast (Frank et al., 2015). The time-frequency window at the time of interest shows the strongest activity at the selected electrodes (Fig. 2B).

2.7.4 EEG data analyses

The selected electrodes and time-frequency window were used to calculate the average theta power per participant and condition, defined by the response made. This enabled us to investigate the influence of midfrontal theta power on motivational conflict in order to establish whether midfrontal theta power was involved in suppressing choice tendencies and/or in suppressing a bias in learning. We looked at the role of midfrontal theta power in suppressing Pavlovian biases with a repeated measures ANOVA with the within-subjects factors Valence (reward vs. punishment) and Action requirement (Go vs. NoGo) when participants learned whether to go, and separately when they learned how to go. If midfrontal theta power is involved in suppressing the influence of Pavlovian biases on our choices, we hypothesise that learning whether to go when the

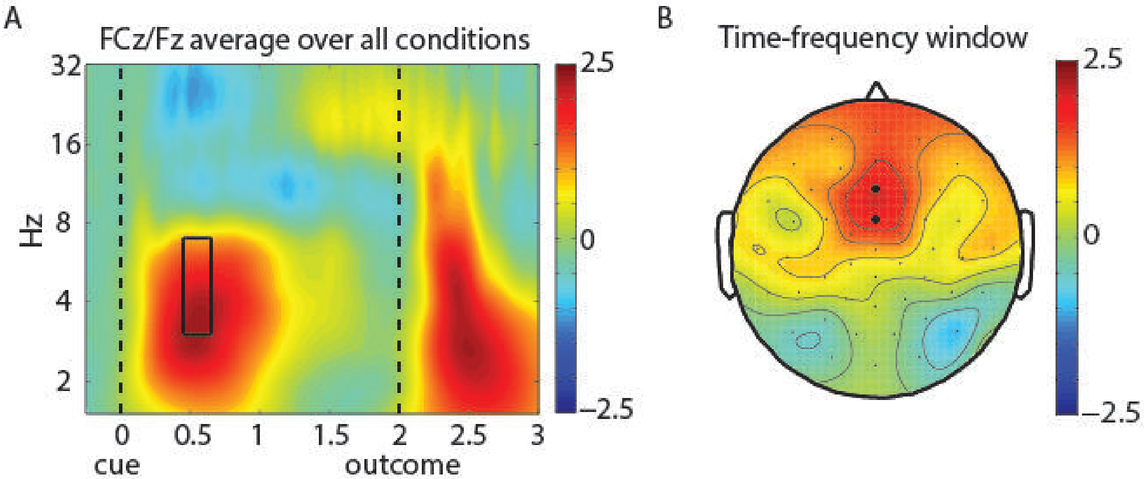

Fig. 2A. The time-frequency representation of the selected electrodes (2 & 8) revealed the peak activity around 0.45–0.65 s post-cue-onset. The square indicates the chosen time-frequency window. **B.** The topographical plot of the chosen time-frequency window show strongest activity at the selected electrodes.

required action was incongruent with our Pavlovian biases would be associated with more midfrontal theta power, compared to when the required action was congruent with Pavlovian biases. Such an interaction effect would provide support for the role of midfrontal theta power in overcoming and suppressing the influence of Pavlovian biases on our choices.

Our design furthermore allowed us to look at the role of midfrontal theta power in learning/credit assignment. To assess the role of midfrontal theta power in learning how to go we analysed the degree of accurate versus inaccurate Go actions on Go cues with repeated measures ANOVA with the within-subjects factors Valence (reward vs. punishment) and Accuracy (accurate vs. error). If midfrontal theta power was involved in suppressing a learning bias, we would expect midfrontal theta power to reduce the difference in biased learning. We would expect this to be implemented by increased midfrontal theta power on accurate Go responses on avoid cues, compared to accurate Go responses on win cues. Thus, more midfrontal theta power would decrease the influence of a learning/credit assignment bias on accuracy on Go trials.

2.7.5 Brain-behaviour correlation

The selected electrodes and time-frequency window were used to calculate a theta effect per participant to look at the relationship between brain activity and behaviour. The theta effect was defined as the difference in midfrontal theta power between responses incongruent, compared to responses congruent, with Pavlovian biases (mean[incongruent] – mean[congruent]). In line with previous studies, the theta effect was first calculated for all responses made on incongruent compared to congruent cues (mean[$\text{incongruent}_{\text{all trials}}$] – mean[$\text{congruent}_{\text{all trials}}$]; Cavanagh et al., 2013), to see whether a similar relationship could be observed. Moreover, our design made it possible to also calculate the theta effect separately for when participants learned whether to go (mean[$\text{incongruent}_{\text{Go/NoGo}}$] – mean[$\text{congruent}_{\text{Go/NoGo}}$]) and when they learned how to go (mean[$\text{incongruent}_{\text{accurate}}$] – mean[$\text{congruent}_{\text{accurate}}$]). This enabled us to further specify the relationship between midfrontal theta power and the influence of Pavlovian biases.

If midfrontal theta power is involved in the ability to overcome Pavlovian biases when conflicting with instrumental requirements, this would be demonstrated by a positive correlation between Pavlovian bias scores and the theta effect when learning whether to go. Thus, as participants rely more on their Pavlovian biases in their choices, the more midfrontal theta power would be required to override these biases in order to perform an action incongruent with Pavlovian biases.

If midfrontal theta power were involved in the ability to overcome a learning bias, this could potentially be demonstrated by a negative correlation between Instrumental bias scores and the theta effect when learning how to go. Thus, we would expect more midfrontal theta power as participants had better performance on Go-to-avoid cues, compared to Go-to-win cues. This would indicate that more midfrontal theta power improves performance on how to go when avoiding punishment.

3. Results

3.1 Behavioural analysis

We hypothesised that anticipation of reward would invigorate responses, while punishment would result in behavioural suppression, which would demonstrate the influence of Pavlovian biases on our behaviour (i.e., biasing approach and inhibition of behaviour). This was confirmed by a significant main effect of valence ($F(1, 32) = 17.11$, $p < .001$) on the proportion of Go responses made in the task. The probability of making a Go response was higher on appetitive cues ($M_\beta = 1.00$, $SEM = 0.17$) compared to on aversive cues ($M_\beta = 0.30$, $SEM = 0.05$; Fig. 3A). If participants indeed learned the action requirements of the cues this would be reflected in the proportion of Go actions to the cues. This was confirmed by a significant main effect of action requirement ($F(1, 32) = 125.85$, $p < .001$). The proportion of Go actions was higher on Go cues ($M_\beta = 1.90$, $SEM = 0.14$) compared to NoGo cues ($M_\beta = -0.55$, $SEM = 0.14$). As expected, we observed no significant interaction effect of valence by action requirement on proportion Go actions ($F(1, 32) = 1.11$, $p = .3$).

If Pavlovian biases interfere with learning action-outcome contingencies incongruent with these biases, we would expect this to be reflected in accuracy. Such interference would result in worse performance on action-outcome contingencies incongruent with Pavlovian biases, compared to action-outcome contingencies congruent with Pavlovian biases. This was confirmed by a significant interaction effect of valence by action requirement on accurate performance ($F(1, 33) = 10.15$, $p = .003$). Planned post-hoc comparison that this congruency effect was driven by a significant difference in accuracy between NoGo-to-win (M_β = 0.17, *SD* = 1.39) and NoGo-to-avoid (M_β = 0.88, *SD* = 0.69; *t*(29) = −2.77, *p* = .009; Fig. 3B), while there was only a trend when comparing Go cues (Go-to-win [M_β = .93, *SD* = 1.32] compared to Go-to-avoid [M_β = 0.55, *SD* = 0.77; *t*(33) = 1.65, *p* = .11]). As expected, we did not observe main effects of valence (*F*(1, 32) = .91, *p* = .35) or action requirement (*F*(1, 32) = 2.20, *p* = .15) on accuracy.

To investigate the influence of Pavlovian biases on accuracy in the different conditions, we looked at the relation between Pavlovian biases and accuracy. Pavlovian bias scores correlated only with accuracy on NoGo-to-win cues (*ρ*(34) = −.83, *p* = .001), but not with the other conditions (*p* > .2). Thus as Pavlovian biases decreased, performance increased on NoGo-to-win cues.

We investigated the influence of motivational valence and action requirement on RTs and expected that the influence of motivational valence would influence RTs, such that avoiding punishment would slow responses, compared to when anticipating reward. This was confirmed by a significant main effect of valence (*F*(1, 33) = 82.13, *p* > .001). Subjects indeed responded faster when playing for a reward (M_β = 0.30, *SEM* = .011) compared to when playing to avoid a punishment (M_β = 0.37, *SEM* = .011). We also observed a significant main effect of action requirement (*F*(1, 33) = 60.59, *p* > .001). Participants responded faster when making Go actions on Go cues (M_β = 0.30, *SEM* = 0.012) compared to NoGo cues (M_β = 0.37, *SEM* = 0.011). No significant interaction effect was observed (*F*(1, 33) = 0.24, *p* = .628).

Furthermore, our novel design allowed us to establish whether the influence of motivational valence was a result of altered choice tendencies or an effect of learning/credit assignment. We would expect this to be demonstrated in the proportion of accurate Go actions on Go cues, where a higher proportion of accurate Go actions on Go-to-win cues compared to Go-to-avoid cues would speak for an effect of learning/credit assignment; the reverse pattern would speak for the influence of altered choice tendencies. Our analysis does not provide support for either of these hypotheses since we observed no difference in the proportion of accurate Go actions between appetitive (*M* = 0.78, *SD* = 0.21) and aversive Go cues (*M* = 0.77, *SD* = 0.14; *t*(33) = 0.25, *p* > .8; Fig. 3C).

3.2 Time-frequency analysis

We hypothesise that the PFC plays a crucial role in resolving motivational conflict through midfrontal theta power. Previous studies have shown that midfrontal theta power is involved in overcoming Pavlovian biases (i.e., whether to go or not), but have not found an effect of midfrontal theta power as a function of congruency of responses (Cavanagh et al., 2013). That is, they have not found more midfrontal theta power on responses incongruent (Go-to-avoid, NoGo-to-win) compared to congruent with Pavlovian biases (Go-to-win, NoGo-to-avoid). Here we perform a similar analysis looking at the

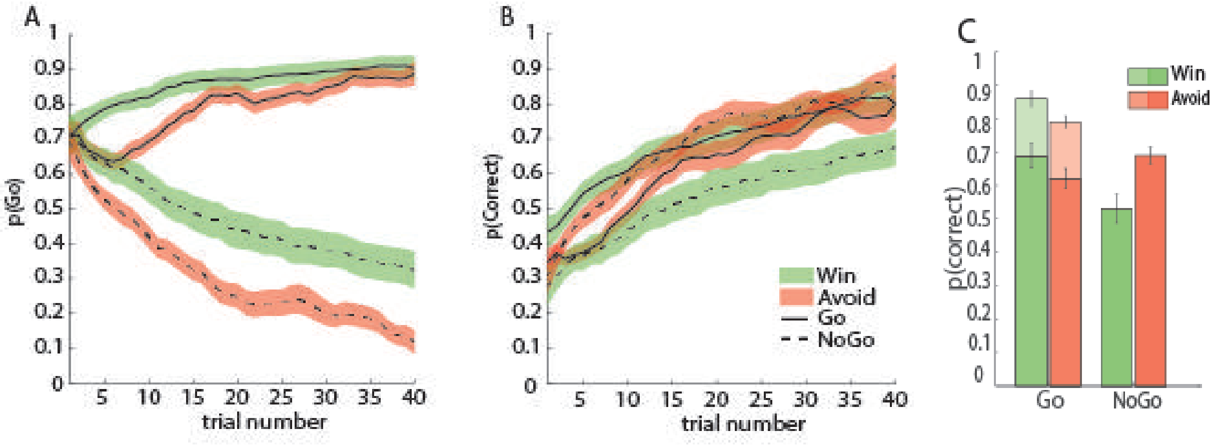

Fig. 3A. Proportion of Go responses over the course of the task for each separate condition, depicted with a moving average. The colored bounds show the standard error of the mean. Cues where participants play for reward are depicted in green and in red when playing to avoid punishment, Go cues are depicted with a solid line and NoGo cues by a dashed line. **B.** Proportion of accurate responses over the course of the experiment in each separate condition, depicted with a moving average. **C.** The proportion of accurate responses (opaque color) on all cues, as well as the proportion inaccurate Go actions (transparent color) on Go cues.

influence of valence and action requirement across the different conditions regardless of response made in order to establish whether we replicate their results in our novel design. In line with their results, we found no significant main effects of valence (*F*(1, 29) = 1.91, *p* = .18) or action requirement (*F*(1, 29) = 0.13, *p* = .7). However, here we did find such a significant interaction effect of valence and action requirement (*F*(1, 29) = 4.47, *p* = .043; Fig. 4A). Moreover, planned post-hoc comparison only showed a marginally significant difference between Go-to-win (*M* = 1.80, *SD* = 1.09) compared to Go-to-avoid (*M* = 2.00, *SD* = 0.99; *t*(29) = −1.66, *p* = .11), but no significant difference between NoGo-to-win (*M* = 2.11, *SD* = 0.90) and NoGo-to-avoid (*M* = 1.96, *SD* = 0.90; *t*(29) = 1.03, *p* = .3).

Unlike previous studies (Cavanagh et al., 2013), our novel design allowed us to assess the role of midfrontal theta power when learning whether to go, but also when learning how to go. If midfrontal theta power is involved in suppressing the influence of Pavlovian biases on choices we would expect this to be reflected by such a congruency effect when participants learn whether to go or not. This hypothesis was confirmed by a significant interaction effect of valence and action requirement (*F*(1, 29) = 6.78, *p* = .014; Fig. 4B). Planned post-hoc comparison of the influence of valence on action requirement showed a marginally significant difference between Go-to-win (*M* = 1.80, *SD* = 1.13) compared to Go-to-avoid (*M* = 2.06, *SD* = 0.99; *t*(29) = −1.77, *p* = .09), and a significant difference between NoGo-to-win (*M* = 2.49, *SD* = 1.01 and NoGo-to-avoid (*M* = 2.26, *SD* = 0.97; *t*(29) = 2.05, *p* = .050). Thus, we observed an effect of midfrontal theta power as a function of congruency of responses when learning whether to go or not, with more midfrontal theta power on incongruent trials compared to congruent trials. Furthermore, we observed no significant main effect of valence (*F*(1, 29) = 0.02, *p* > .8), but a significant main effect of action requirement (*F*(1, 29) = 12.01, *p* = .002), showing more midfrontal theta power on NoGo cues (*M* = 2.37, *SEM* = 0.169) compared to Go cues (*M* = 1.93, *SEM* = 0.18).

Moreover, when investigating the influence of midfrontal theta power when learning how to go, we observed a similar congruence effect, demonstrating the involvement of midfrontal theta power in learning how to go (*F*(1, 29) = 7.91, *p* = .009; Fig. 4C). Planned post-hoc comparison of the influence of valence on action requirement showed a marginally significant difference between Go-to-win (*M* = 2.11, *SD* = 1.00) compared to Go-to-avoid (*M* = 2.38, *SD* = 0.98; *t*(29) = −1.94, *p* = .06), and significant difference between NoGo-to-win (*M* = 2.49, *SD* = 1.01 and NoGo-to-avoid (*M* = 2.26, *SD* = 0.97; *t*(29) = 2.05, *p* = .050). Thus, we observed an effect of midfrontal theta power as a function of congruency of responses when learning how to go, with more midfrontal theta power on incongruent trials compared to congruent trials. This result suggests that midfrontal theta power is important in overcoming Pavlovian biases in order to make an accurate response. We did not observe significant main effects of valence (*F*(1, 29) = 0.05, *p* > .8) or action requirement on accuracy (*F*(1, 29) = 1.31, *p* > .2), indicating that the main effect of action when learning whether to go or not is driven by accuracy (see Discussion).

Our novel design also allowed us to investigate whether midfrontal theta power is involved in learning/credit assignment. To assess the role of midfrontal theta power in learning/credit assignment we analysed the influence of valence and accuracy on accurate versus inaccurate Go actions on Go cues. We observed a significant main effect of accuracy (*F*(1, 29) = 10.65, *p* = .003), with higher midfrontal theta power on accurate Go actions (*M* = 2.24, *SEM* = 0.17), compared to inaccurate Go actions (*M* = 1.59, *SEM* = 0.24). The analysis also showed a marginally significant main effect of valence (*F*(1, 29) = 2.35, *p* = .08), indicating more midfrontal theta power on aversive Go cues (*M* = 2.06, *SEM* = 0.181), compared to appetitive Go cues (*M* = 1.78, *SEM* = 0.209). This suggests that midfrontal theta power is involved in suppressing the influence of Pavlovian biases when making a Go action incongruent with Pavlovian biases. However, there was no significant interaction effect, suggesting that midfrontal theta power is involved in suppressing the influence of motivational valence on choice, rather than learning (*F*(1, 29) = 0.01, *p* = .9; Fig. 5A).

3.3 Brain-behaviour correlation

Previous studies have associated midfrontal theta power during motivational conflict with less influence of Pavlovian biases on choices (Cavanagh et al., 2013). We performed a similar analysis to see whether we would observe a similar effect with our design. However, we observed no such relationship (ϱ(30) = .05, *p* > .7; Fig. 4A). In order to establish whether the relationship between Pavlovian biases and midfrontal theta was instead related to learning whether to go or not, or learning how to go, we looked at Pavlovian biases and midfrontal theta when participants learned whether to go and when

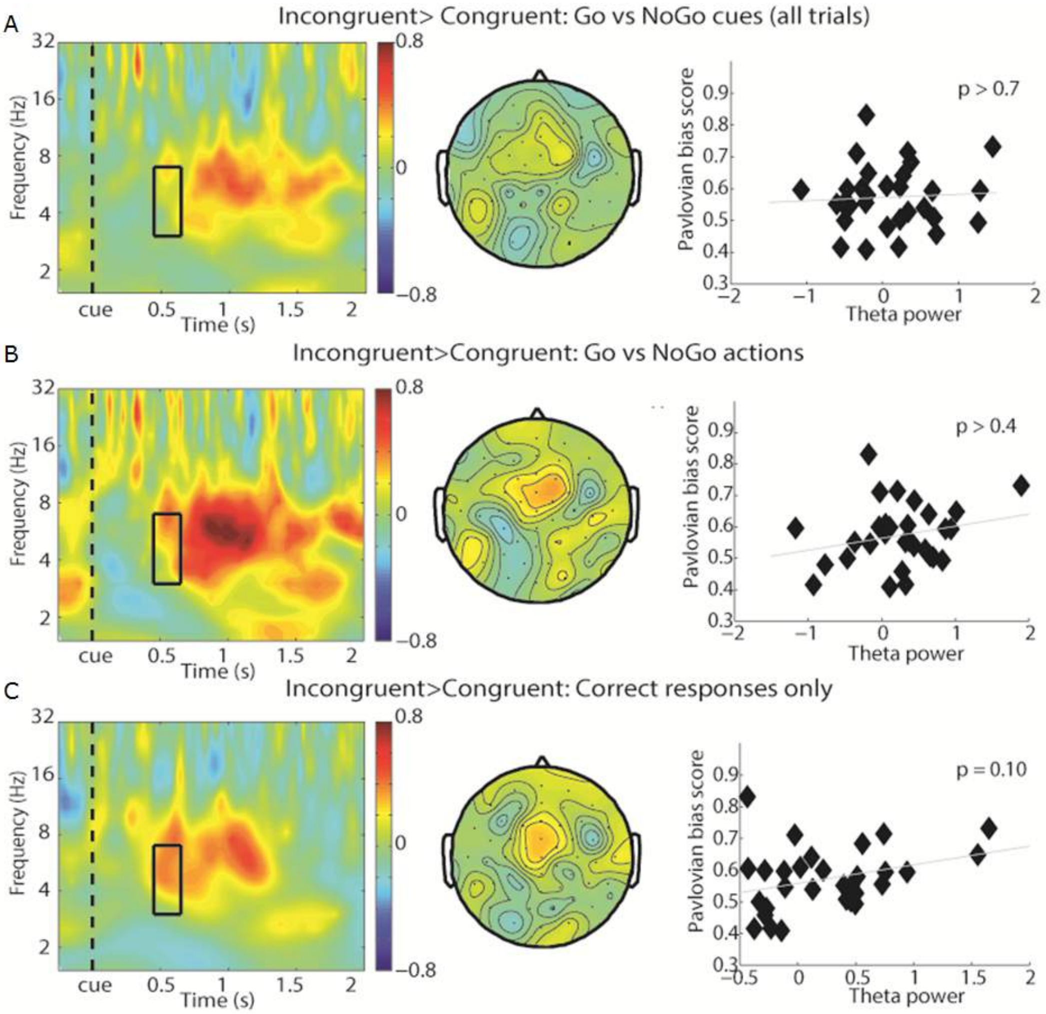

Fig. 4A. A significant interaction effect of valence and action requirement across the different conditions regardless of response made ($p = .04$), but no correlation between Pavlovian bias scores and the theta effect (mean[incongruent[all trials]] − mean[congruent[all trials]]; $p > .7$). **B.** We observe an effect of midfrontal theta power as a function of congruency of responses ($p = .014$), but no correlation between Pavlovian bias scores and the theta effect (mean[incongruent[Go/NoGo]]-mean[congruent[Go/NoGo]]; $p > .4$). **C.** The observed effect of midfrontal theta power as a function of congruency of responses strengthens as accuracy of responses is taken into account ($p = .009$), we further observe a trend between Pavlovian bias scores and the theta effect (mean[incongruent[accurate]] -mean[congruent[accurate]]; $p = .10$).

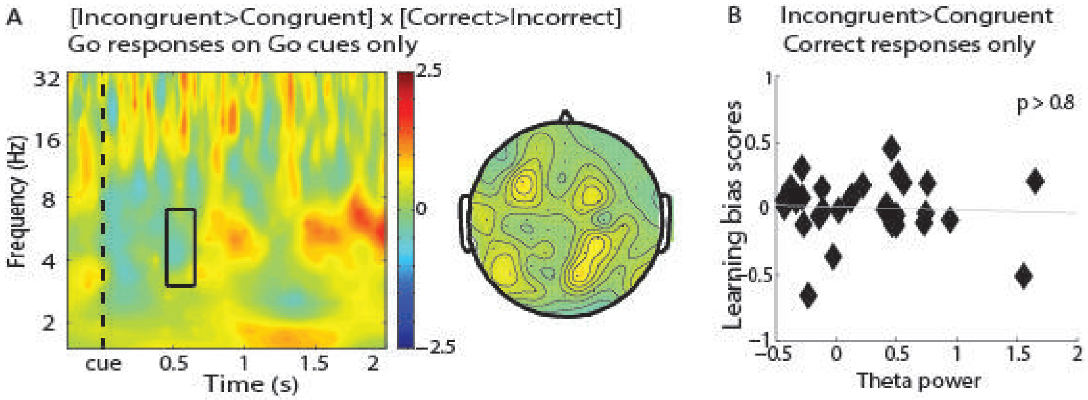

Fig. 5A. No significant interaction effect of accuracy and action on Go-trials ([Go-to-win[accurate-errors]] − [Go-to-avoid[accurate-errors]]; $p > .9$). **B.** No correlation between the Instrumental bias scores and the theta effect (mean[incongruent[accurate]] − mean[congruent[accurate]]; $p > .8$).

they learned how to go. We found no relationship between Pavlovian biases and the theta effect when learning whether to go ($\varrho(30) = .15$, $p > .4$; Fig. 4B).

However, we observed a positive trend between Pavlovian biases and the theta effect when learning how to go ($\varrho(30) = .31$, $p = .10$; Fig. 4C), but no relationship between the instrumental bias scores and any of the theta effects ($p > .8$; Fig. 5B). The observed positive trend suggests that the more participants' actions were driven by their Pavlovian biases, the more midfrontal theta power was associated with suppressing the influence of these biases in order to perform the required, but conflicting, response. Thus, midfrontal theta power does not seem to be involved in merely overcoming our Pavlovian biases to be able to make a response incongruent with our Pavlovian biases, but in reducing the influence of Pavlovian biases on our choices.

4. Discussion

The aim of this study was to establish whether well-known influences of motivation on action were due to effects on choice or learning. Specifically, we asked whether this motivational effect was due to a Pavlovian bias that nonspecifically invigorates responses to reward and inhibit responses to punishment, as previously suggested (Cavanagh et al., 2013), or rather to a bias in learning where valence promotes credit assignment to specific actions. Using a novel task design, we show that midfrontal theta power is involved in the ability to overcome the influence of Pavlovian biases on our choices when these interfere with learning of instrumental action-outcome contingencies. Thus, our results show that midfrontal theta power is not involved in merely overcoming Pavlovian biases on choices, but in suppressing the influence of Pavlovian biases in order to learn how to go; to perform the required action in order to reach desirable outcomes.

4.1 Pavlovian biases on behaviour

Compared to previous studies (Cavanagh et al., 2013; Guitart-Masip et al., 2011, 2012), our design allowed us not only to assess whether the effect of motivational valence on learning and choice related to behavioural activation (whether to go or not to go), but also whether it related to the ability to assign credit to specific actions (how to go). Since our task required learning of specific Go actions, we were able to disentangle whether motivational bias effects previously reported (Cavanagh et al., 2013; Guitart-Masip et al., 2011, 2012) were a result of altered choice tendencies or an effect of learning/credit assignment. Our design required learning the most optimal response out of three possible response options for eight action-outcome contingencies. Our task requirements make the task more difficult than previously used Go/NoGo reinforcement learning tasks (Cavanagh et al., 2013; Guitart-Masip et al., 2011, 2012), where participants are required to learn to either go or not to go, for four cues. Despite the increased task difficulty, participants still managed to learn the action requirements in our task.

Our results show, in line with previous research, that we are more likely to go for a reward and less likely to go to avoid punishment (Cavanagh et al., 2013; Guitart-Masip et al., 2012, 2014), and that we act faster when anticipating a reward (Crockett et al., 2009). In line with previous research, our results also show that we are better at learning action-outcome contingencies congruent with Pavlovian biases (Cavanagh et al., 2013; Guitart-Masip et al., 2012, 2014). Thus, our findings demonstrate the influence of Pavlovian biases on choices when learning action-outcome contingencies incongruent with these biases, interfering with the required instrumentally optimal action. However, we observed no difference in accuracy on Go actions made on win and avoid cues, but only between NoGo responses made on win and avoid cues. The observed congruency effect on accuracy was strongly driven by NoGo-to-win cues, and Pavlovian bias scores correlated only with accuracy on NoGo-to-win cues. Therefore, our results suggest that this impaired performance is due to a bias in choice, since participants made more Go actions at win cues compared to on avoid cues. This does however not exclude the possibility that it is more difficult to associate behavioural suppression with reward, compared to an active behavioural response.

Our novel design allowed us to investigate whether the influence of motivational valence on Go actions specifically was a result of altered choice tendencies or an effect of learning/credit assignment. We observed no difference in proportion of accurate Go actions between win and avoid cues. These results suggest that reward does not increase the likelihood of making an accurate Go action when going for reward compared to when going to avoid punishment (credit assignment), but also not that Pavlovian biases increases the proportion of inaccurate Go actions when going for reward compared to when going to avoid punishment (biased choice). Thus, we observed more Go actions on win cues compared to avoid cues in general, but did not observe a difference in proportion of accurate responses on Go-to-win and Go-to-avoid cues. Taken together, these results indicate that the influence of motivational valence results from altered choice tendencies, rather than biased credit assignment. However, with this analysis we cannot exclude the possibility that the influence of motivational valence was a result of an interaction of altered choice tendencies and biased credit assignment. To establish whether such an interaction exists, a computational modeling approach might be helpful.

4.2 Midfrontal theta power in motivational conflict

We hypothesised that the PFC plays a crucial role in resolving motivational conflict through midfrontal theta power. Here we investigated the influence of midfrontal theta power on motivational conflict to establish whether midfrontal theta power was involved in suppressing choice tendencies and/or learning/credit assignment. Previous studies have shown that midfrontal theta power is involved in overcoming Pavlovian biases (i.e., whether to go or not), but did not find an effect of midfrontal theta power as a function of congruency of responses across participants (Cavanagh et al., 2013). Here we perform a similar analysis, but do observe an effect of midfrontal theta power as a function of congruency of responses. Our novel task furthermore allowed us to look at the influence of midfrontal theta power in suppressing Pavlovian biases when participants learned whether to go, and when learning how to go, enabling us to better understand the observed effect. Our results show that midfrontal theta power was increased when learning whether to go when the action was incongruent with the Pavlovian biases, but also when learning to perform the instrumentally optimal action (how to go). This could not be established by previous studies (Cavanagh et al., 2013). Thus, our results provide further support for a role of midfrontal theta power in suppressing the influence of Pavlovian biases on choices. Furthermore, we observed more midfrontal theta power on NoGo cues when learning whether to go. This effect could be explained by accuracy, since NoGo cues only consisted of accurate responses, compared to Go cues consisting of both accurate and inaccurate Go actions. Thus, our analyses clearly show that the effect of midfrontal theta power as a function of congruency of responses was driven by accuracy; the effect of midfrontal theta power strengthened as the amount of errors reduced in the analyses.

With our novel design, we could furthermore specifically look at learning of required Go actions, enabling us to further understand the observed midfrontal theta effect. Through this analysis we could establish that more midfrontal theta power was needed when performing accurate Go actions compared to inaccurate Go actions on Go cues. We also observed a trend suggesting more midfrontal theta power when making Go actions incongruent with Pavlovian biases, regardless of accuracy. We observed no interaction effect of valence and accuracy, supporting the role of midfrontal theta power in overcoming the influence of Pavlovian biases on our choices. Together with our behavioural results, these results suggest that our actions are driven by motivational effects on choice, rather than learning/credit assignment. Furthermore, our results further suggest that midfrontal theta power is involved in overcoming the influence of these biases on our choices.

Thus, in order to realise goal-directed control over action selection to be able to reach the most optimal outcome, more midfrontal theta power is needed when successfully performing active actions when avoiding punishment, as well as when required to withhold active responses in order to get a reward. Being able to successfully inhibit active responses in anticipation of reward has been associated with successful recruitment of control processes needed for behavioural suppression (Guitart-Masip et al., 2012). The low accuracy on NoGo-to-win cues could be explained by unsuccessful recruitment of top-down control, potentially resulting in unsuccessful inhibition of behavioural responses when Pavlovian biases interfere with instrumental learning (Guitart-Masip et al., 2012). Here we show that midfrontal theta power seems to be important in resolving motivational conflict by suppressing the influence of motivational valence on our choices and might play a role in signaling the need of top-down control to other relevant control processes.

4.3 Midfrontal theta and behaviour across participants

Previous studies have associated less influence of Pavlovian biases on choices with midfrontal theta power at motivational conflict (Cavanagh et al., 2013). We failed to replicate their result, and found no relationship between midfrontal theta power and Pavlovian biases when comparing cues requiring incongruent and congruent responses. However, with our novel design we were able to further look into the relationship between midfrontal theta effects at motivational conflict when participants learned whether to go, but also when they learned how to go. When taking accuracy into account, we observe a trend between Pavlovian biases and the theta effect when participants learned whether to go. This trend suggests that, as participants' actions were more driven by their Pavlovian biases, more midfrontal theta power was associated with successful suppression of the influence of these biases, in order to perform the required, but conflicting, response.

That we fail to replicate the previously observed effect, but do see a trend when taking accuracy into account, might relate to the requirements of our task and how they differ from previously used tasks. That is, it is reasonable to assume that more time was required to learn the action-outcome contingencies in our task, resulting in a bigger proportion of trials per condition where conflict is yet not experienced than in the previous study (Cavanagh et al., 2013). Thus, by taking accuracy into account, we look at trials where the action-outcome requirements have been learned, and when participants are familiar with which cues are associated with reward or punishment. Our results show that there is a relationship between midfrontal theta power and accurately overcoming the influence of Pavlovian biases on our choices.

Moreover, this conclusion is further supported by the absence of a relationship between instrumental learning bias scores and midfrontal theta power when participants learned how to go. However, the absence of a relationship between midfrontal theta power and instrumental bias scores does not exclude the possibility of a bias in learning. The way the instrumental bias scores were quantified here might be an insufficient way of quantifying an instrumental learning bias, and might require a different approach. A more appropriate approach would be to use a model-based analysis to look at trial-by-trial midfrontal theta power to see whether midfrontal theta power influences the degree of biased learning.

4.4 Future directions

Our results show that the influence of motivational valence on our actions is due to altered choice tendencies, and that midfrontal theta power is playing a role in suppressing the influence of motivational biases on our choices, in order to make instrumentally optimal choices. With our analyses we cannot exclude a potential interaction between Pavlovian biases and biased learning. Previous research has used computational models to look at the influence of trial-by-trial theta power on action selection, to investigate the role of midfrontal theta power in motivational conflict (Cavanagh et al., 2013). By using such an approach, we would be able to better understand the dynamic of a potential interaction between an influence of Pavlovian biases and of biased learning on action selection. This would also enable us to establish whether midfrontal theta power is also involved in shifting from Pavlovian to a instrumental control at motivational conflict, to be able to perform the required action.

Future analyses on equivalent tasks should take learning effects into account. In the current study we did not split trials based on whether they appeared early or late during the task. We are likely to have more error trials early on in our task, since the learning requirements were rather difficult. This might introduce unwanted variance when analyzing condition-averaged data, influencing effects of interest. Our analysis on midfrontal theta power in credit assignment is especially sensitive to this, since inaccurate Go actions might have been made early on during the task, when the valence of the cue might not have been learned yet. The proposed modeling approach would resolve this issue.

Another approach to establish whether midfrontal theta power is involved in shifting from Pavlovian to instrumental control at motivational conflict would be to look at whether midfrontal theta power at motivational conflict could predict action selection. This could be done by looking at the correlation, or at the cross-frequency interactions between midfrontal theta power (at motivational conflict) and beta power (i.e., motor activation) within each trial (Aru et al., 2015). Such cross-frequency interactions might play a role in implementing top-down control on action selection, by modulating motor activation following motivational conflict. If we were able to classify which response would be made based on beta power, we could establish whether beta power is dependent on midfrontal theta power at motivational conflict. Thus, to disentangle if cue-locked midfrontal theta power at motivational conflict is predictive of general (i.e., suppression of Pavlovian biases), or specific motor activation (i.e., promoting instrumental responses). This could provide insight about the role of midfrontal theta power during action selection, and how such control is communicated and implemented in the brain.

5. Conclusion

With our novel design we were able to establish that midfrontal theta power is involved in overcoming Pavlovian biases when encountering conflict between Pavlovian biases and required instrumental responses. Participants with bigger Pavlovian bias were characterised by greater frontal theta power at motivational conflict when accurately performing the instrumentally required response. Our findings indicate that the influence of motivational valence on our choices is driven by Pavlovian biases rather than learning/credit assignment. However, with our analyses we were not able to determine whether the influence of motivational valence on choice and learning was affected by an interaction of Pavlovian biases and credit assignment, or whether midfrontal theta power is involved in shifting from Pavlovian to instrumental control at motivational conflict. To better understand the role of midfrontal theta power in motivational conflict, other analysis approaches would be suitable, such as computational modeling.

6. References

Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., … Vicente, R. (2015). Untangling cross-frequency coupling in neuroscience. *Current Opinion in Neurobiology*, *31*, 51–61.

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides pavlovian learning biases. *The Journal of Neuroscience*, *33*(19), 8541–8548.

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, *22*(11), 2575–2586.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421.

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209.

Cavanagh, J. F., Masters, S. E., Bath, K., & Frank, M. J. (2014). Conflict acts as an implicit cost in reinforcement learning. *Nature Communications*, *5*, 5394.

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238.

Chiu, Y., Cools, R., & Aron, A. R. (2010). Opposing Effects of Appetitive and Aversive Cues on Go/No-go Behavior and Motor Excitability. *Journal of Cognitive Neuroscience*, *26*(8), 1851–1860.

Cohen, M. X. (2014). *Issues in Clinical and Cognitive Neuropsychology: Analyzing Neural Time Series Data: Theory and Practice*. Cambridge, US: The MIT Press.

Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *The Journal of Neuroscience*, *29*(38), 11993–11999.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.

Dayan, P., Niv, Y., Seymour, B., & D. Daw, N. (2006). The misbehavior of value and the discipline of the will. *Neural Networks*, *19*(8), 1153–1160.

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. *Journal of Neuroscience*, *35*(2), 485–494.

Geurts, D. E. M., Huys, Q. J. M., & den Ouden, H. E. M. (2011). Aversive Pavlovian Control of Instrumental Behavior in Humans. Journal of cognitive neuroscience, *25*(9) 1428–1441.

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, 1–9.

Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J. M., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *The Journal of Neuroscience*, *31*(21), 7867–7875.

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: interactions between affect and effect. *NeuroImage*, *62*(1), 154–166.

Hajihosseini, A., & Holroyd, C. B. (2013). Frontal midline theta and N200 amplitude reflect complementary information about expectancy and outcome evaluation. *Psychophysiology*, *50*(6), 550–562.

Hershberger, W. a. (1986). An approach through the looking-glass. *Animal Learning & Behavior*, *14*(4), 443–451.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520.

Oostendorp, T. F., & Van Oosterom, A. (1996). The Surface Laplacian of the Potential: Theory and Application. *IEEE Transactions on Biomedical Engineering*, *43*(8), 405–866.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9.

Ridderinkhof, K. R., Van den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140.

Robinson, O. J., Cools, R., & Sahakian, B. J. (2012). Tryptophan depletion disinhibits punishment but not reward prediction: Implications for resilience. *Psychopharmacology*, *219*(2), 599–605.

Srinivasan, R., Winter, W. R., Ding, J., & Nunez, P. L. (2007). EEG and MEG coherence: Measures of functional connectivity at distinct spatial scales of neocortical dynamics. *Journal of Neuroscience Methods*, *166*(1), 41–52.

Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human Pavlovian Instrumental Transfer. *Journal of Neuroscience*, *28*(2), 360–368.

Van de Vijver, I., Ridderinkhof, K. R., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of Cognitive Neuroscience*, *23*(12), 4106–4121.

Van Driel, J., Ridderinkhof, K. R., & Cohen, M. X. (2012). Not all errors are alike: theta and alpha EEG dynamics relate to differences in error-processing dynamics. *The Journal of Neuroscience*, *32*(47), 16795–16806.

The Influence of the First Language on Syntactic Processing in a Second Language: An EEG Study with German Learners of Dutch

Anne Mickan[1]
Supervisor: Kristin Lemhöfer[2]

[1]*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*
[2]*Donders Institute for Brain Cognition and Behaviour, Centre for Cognition, Nijmegen, The Netherlands*

Despite a growing interest in second language (L2) syntactic processing, little is known about how L2 speakers overcome direct conflicts between first language (L1) and L2 syntactic rules. Sentence-final double infinitives in German and Dutch pose such a direct conflict: the two infinitives in the Dutch sentence 'Ik heb het huiswerk niet willen maken' (I did not want to do the homework) would be reversed in a German translation of the sentence (…machen wollen). We used Event Related Potentials (ERPs) to explore how far German learners of Dutch have already overcome L1 rules and show native-like processing signatures to the word order that is incorrect in Dutch, but correct in German. While grammaticality judgments showed comparable error rates in L2 speakers and a Dutch native speaker control group, ERP signatures differed between the two groups. Native speakers showed a clear P600 effect to order violations. For L2 speakers, in contrast, the P600 effect was delayed, and was furthermore preceded by an N400-like effect and an early left negativity. Interestingly, P600 and N400 effect magnitudes were negatively correlated in L2 speakers, suggesting that participants tended to show either a P600 or an N400, rather than both. Differences in N400–P600 dominance could, however, not be explained by individual differences in L2 proficiency or working memory capacity. We conclude that despite native-like behavioral performance, structures with conflicting syntactic rules in L1 and L2 lead to non-native-like online processing signatures. The extent of the impact of the rule conflict, however, differs for individual L2 speakers.

Keywords: second language acquisition, syntactic processing, L1–L2 similarity, syntactic transfer effects, P600, N400, individual differences

---

Corresponding author: Anne Mickan; **E-mail**: amickan1990@gmail.com

1. Introduction

In an increasingly globalized world, a growing number of young adults find themselves facing the challenge of learning a foreign language. How second languages (L2) are learned and processed has become a vivid area of research. Despite the wealth of studies on the topic, a lot of questions regarding the L2 acquisition process remain unanswered. One question that has inspired a great deal of research over the past decades is whether or not native-like processing of foreign syntax is possible at all in late L2 learners.

A number of factors have been proposed to influence the success of ultimate L2 syntactic attainment, most prominent among them are the degree of similarity between the first (L1) and second language (Chen, Shu, Liu, Zhao, & Li, 2007; Clahsen & Felser, 2006; McLaughlin et al., 2010; Sabourin & Stowe, 2008; Tokowicz & MacWhinney, 2005; White, Genesee, & Steinhauer, 2012). L1 and L2 are believed to constantly interact during the L2 acquisition process and learners are assumed to have the tendency to transfer their L1 knowledge over to their L2 whenever possible (MacWhinney, 2005). For structures that are similar between languages, this can result in facilitation of the learning process (positive transfer), but for structures that differ between languages, this may lead to significant interference from the L1 and so-called negative transfer (Caffarra, Molinaro, Davidson, & Carreiras, 2015; Kotz, 2009; Odlin, 1989).

The present study investigates one particularly interesting case of negative transfer, namely structures with opposite syntactic rules in L1 and L2. Despite the growing interest in the influence of the L1 on L2 syntactic processing, it remains largely unclear how learners deal with such directly conflicting syntactic structures. Before delving into detail regarding the design of the present study, we will briefly review some of the existing behavioural and electrophysiological literature on L1–L2 transfer effects.

1.1 Evidence from behavioural studies

The first studies to investigate syntactic transfer phenomena were of a predominantly observational nature and focused on production errors and offline grammaticality judgments (Hulk, 1991; Jansen, Lalleman, & Muysken, 1981). Other, more sophisticated behavioural measurements have included reaction times and eye-tracking (Tuninetti, Warren, & Tokowicz, 2015). Overall, these behavioural studies seem to suggest that the L1 has its greatest influence in the early stages of the learning process and plays much less of a role with increasing L2 proficiency.

In a study on word order differences between languages, for instance, Jansen, Lalleman, and Muysken (1981) had Moroccan Arabic and Turkish learners of Dutch engage in conversations in Dutch and subsequently coded their speech for word order production errors. They were specifically interested in the placement of the verb in main and subordinate clauses, for which rules differ in Moroccan Arabic, Turkish and Dutch. Jansen et al. (1981) found that both learner groups produced anomalous Dutch sentences in line with L1 verb placement preferences. However, importantly, this pattern was mostly observed in low-proficient participants. In contrast, participants with higher proficiency levels approached the Dutch norm and were hardly influenced by L1 verb placement patterns.

In an eye-tracking study with advanced Chinese and Arabic learners of English, Tuninetti, Warren, and Tokowicz (2015) confirmed that the L1 has little to no influence at higher levels of proficiency. They chose violations of noun-adjective and noun-article placement in English, which are realized differently in Chinese and Arabic. Both learner groups were highly sensitive to both types of syntactic violations in the L2 and were not significantly influenced by their L1 as evidenced by few grammaticality judgment errors and overall similar eye-movement trajectories for the two groups in all conditions.

Other behavioural studies on word order violations have reached similar conclusions (e.g., Erdocia, Zawiszewski, & Laka, 2014; Pozzan & Quirk, 2014), thus suggesting that the L1 only has an influence during the early stages of L2 acquisition, but quickly ceases to show negative signs of interaction with the L2.

1.2 Evidence from event-related potentials (ERPs)

More recent studies using ERPs to investigate the impact of the L1 on L2 syntactic processing provide a more nuanced picture. ERP measurements can be acquired in the absence of behavioural responses and have been shown to be sensitive to effects not observable with traditional behavioural methods (Chen et al., 2007; McLaughlin, Osterhout, & Kim, 2004). They thus allow for a more fine-grained investigation of L1–L2 similarity effects on L2 syntactic processing.

1.2.1 *The violation paradigm and common syntactic ERP components*

Studies using these online methods usually have both native speakers and L2 learners read sentences containing syntactic violations and compare the two groups' ERP signatures. The question then is whether the L2 group is sensitive to the violations at all and if so, to what extent the learners' ERP signatures resemble those of native speakers.

In native speakers, syntactic violations reliably elicit a late positive-going voltage deflection starting at around 600 ms post stimulus onset. This effect is largest over centro-parietal scalp sites and is commonly referred to as P600 (Osterhout & Holcomb, 1993). Due to its pervasive appearance in studies on syntactic processing in native speakers, it is believed to reflect syntactic repair (Friederici, 1995) and structural reanalysis processes (Hagoort, Brown, & Groothusen, 1993). Some studies additionally report a left anterior negativity, which, depending on its latency, has been called either an early left anterior negativity (ELAN, 100–300 ms) or simply a left anterior negativity (LAN, 300–500 ms) (Friederici, Pfeifer, & Hahne, 1993). Both components are usually found over left or left anterior electrode sites. Their functional significance is still largely unclear, but they have mostly been interpreted as evidence for rule-goverened automatic structure-building (ELAN), and as an indicator for the early stages of morphosyntactic processing (LAN) (for a review, see Steinhauer & Drury, 2012).

In L2 speakers, syntactic violations can elicit a wide range of ERP signatures (for a comprehensive summary see Bowden, Steinhauer, Sanz, & Ullman, 2013). I will henceforth only focus on results from EEG studies that have specifically investigated the effect of the L1 on L2 syntactic processing.

1.2.2 *Online processing of structures unique to the L2 and structures compatible in L1 and L2*

Previous EEG studies on L1–L2 similarity effects focused on how learners process structures that are either *unique* to the L2 (i.e., absent in the L1), or structures that are the same in L1 and L2. Ojima, Nakata, and Kakigi (2005), for instance, tested low and high proficient Japanese learners of English on subject-verb agreement violations. Since subject-verb agreement does not exist in Japanese, Japanese learners cannot rely on their L1 in processing violations of this kind in English. The ERP results reflect this difference between languages: while English native speakers showed a biphasic LAN–P600 pattern to syntactic violations, the learner group did not. In fact, low-proficient Japanese participants did not show any ERP effects at all, and high proficient learners showed only the LAN, not the P600. The latter finding is all the more striking considering that the high proficient learner group performed on a native-like level on grammaticality judgments.

Chen, Shu, Liu, Zhao, and Li (2007) report very similar results in their study on subject-verb agreement with highly proficient Chinese learners of English. Like Japanese, Chinese does not mark person or number on the verb. While the learners again performed as well as the native speakers on grammaticality judgments, they failed to show both the LAN and the P600 elicited in native speakers. The learner group instead showed a late negativity to violations of subject-verb agreement. This suggests that even though they were sensitive to the violations, they employed different neural substrates in processing them.

Authors of both studies conclude that these differences between native and non-native speakers are essentially driven by the structural differences between the languages, which even at high levels of proficiency and when mastered behaviourally can lead to non-native-like online processing. Structures that are implemented in very similar ways across languages are, in contrast, reported to elicit near native-like ERP signatures (Foucart & Frenck-Mestre, 2012; McLaughlin et al., 2010; Sabourin & Stowe, 2008). Foucart and Frenck-Mestre (2012), for example, tested advanced German learners of French on grammatical gender agreement violations. For rules implemented similarly in German and French (determiner–noun agreement), they report a robust, native-like P600 effect in the learners. For violations of rules that differ or do not exist in the L1 (noun–adjective gender agreement), however, German learners failed to show ERP effects. The authors take these results as strong evidence for both positive and negative transfer.

A similar processing advantage for similar structures over structures unique to the L2 was described by McLaughlin, Osterhout, and Kim (2010). In a longitudinal study, they followed English native speakers enrolled in their first year of French instruction. They tested their participants three times over the course of a year and measured their ERPs to a syntactic rule that is similar in both languages (subject–verb agreement) and a rule that is unique to the L2 (determiner–number agreement). For the similar rule, learners made rapid progress over the year: while showing an N400 to agreement violations at the first session (four weeks into a French course), they showed a native-like P600 at session three (after 26 weeks of French). For the rule unique to the L2, instead, no such changes occurred: learners remained completely insensitive to rule violations over all three sessions.

Overall, these EEG studies suggest that L1–L2 similarity does play a crucial role in determining the rate of acquisition of L2 syntactic structures and the ultimate success of online L2 syntactic processing, both in beginning and advanced learners.[1]

1.2.3 Processing of direct conflicts between L1 and L2 syntactic rules

Strikingly, very little systematic research has gone into the examination of syntactic structures that are implemented in opposite ways in L1 and L2, a scenario where arguably the strongest transfer is to be expected. Grammatical gender, for example, often differs for translation equivalents in different languages. In Dutch and German word gender is often exactly opposite: the word for 'car', for instance, is common (i.e., masculine/feminine) gender in Dutch ('de auto'), while being neuter in German ('das Auto'). Cases like the latter are especially difficult since the two words are identical in form, making it very difficult for the learner to resist transferring syntactic gender knowledge from L1 to L2.

Lemhöfer et al. (2014) investigated such direct gender assignment conflicts for form-similar words in German and Dutch. A behavioural study had shown that these gender conflicts tend to result in wrong subjective gender representations in German learners of Dutch (Lemhöfer, Schriefers, & Hanique, 2010). In the EEG study, Lemhöfer and colleagues (2014) found that learners use these subjective gender representations also during online processing and show a P600 to subjectively rather than objectively incorrect gender assignments. This suggests that knowledge from the first language can lead to the reversal of ERP signatures in situations when learners apply their L1 rather than their L2 grammar during processing.

EEG studies looking at more purely syntactic features, such as word order, are virtually inexistent. Much like for grammatical gender though, the L1 may exert a very strong negative influence fostering similarly wrong subjective representations for L2 word order. The present study aims at closing this gap in the research literature and sets out to test how German learners of Dutch deal with directly conflicting word order rules in their two languages.

The present research is partially inspired by a study from Davidson and Indefrey (2009), who investigated a structure very similar to the one we chose. They looked at verb clusters in subordinate clauses consisting of a finite auxiliary verb ('laten' – let) and a main verb in infinitive form (e.g.: 'raken' – touch). According to German grammar, the infinitive has to precede the finite verb (see Dutch sentence in (2), 'raken laten'), whereas in Dutch there is a preference for the opposite order (see (1), 'laten raken'), though both orders are grammatically possible.

(1) Je zal zien dat wij het rode kruis de blauwe driehoek **laten raken**.
(2) (*) Je zal zien dat wij het rode kruis de blauwe driehoek **raken laten**.
[You will see that we let the red cross touch the blue triangle.]

Davidson and Indefrey (2009) used magnetoencephalography (MEG) to investigate how German learners' processing signatures for these sentences change over time. At three time points over a three-month period, participants read sentences like (1) and (2) and were asked to judge whether they matched a subsequently presented picture. Behaviourally, participants became faster and more accurate in making their judgments over time. MEG signatures, (event-related fields [ERF]), also showed improvement over time: measured on the first verb, Davidson and Indefrey (2009) report a frontal LAN-like effect to the German verb order as in (2) as compared to the preferred Dutch verb order as in (1) after only two weeks of instruction. They interpret the emergence of a differential neural response to the two verb orders to indicate a change in online syntactic parsing of the sentences, due to the application of newly learned Dutch grammatical rules.

Their results generally seem to be in line with studies showing that the learners' online sensitivity to syntactic violations in the L2 changes over time (McLaughlin et al., 2010; Rossi, Gugler, Friederici, & Hahne, 2006; White et al., 2012). However, the fact that Davidson and Indefrey (2009) use MEG instead of EEG makes their results difficult to compare to

---

[1] It should be noted that the absence of a comparable structure in the L1 does not always hinder processing. There are a few studies that have shown near-native-like signatures to such structures (Gillon Dowens, Guo, Guo, Barber, & Carreiras, 2011; Tokowicz & MacWhinney, 2005; White et al., 2012). However, ERP signatures found in these studies are not canonical in terms of topography or latency, thus calling for caution in interpreting them.

earlier studies on L1–L2 similarity effects. Moreover, the fact that they did not have a native speaker control group leaves open whether and to what extent the ERF signatures shown by German speakers in their study were actually "improving" and approaching native-like levels towards the third session.

Furthermore, given the task during the MEG, ERF signatures may not have captured online syntactic sensitivity in the way earlier EEG studies do. German speakers had to match sentences and visual scenes, a task for which in-depth syntactic processing is not necessary. Changes in ERF signatures over time may have been of a different nature had participants been instructed to pay attention to grammatical correctness online (as is common practice in EEG studies on L2 syntactic processing). Finally, the rapid improvement in ERF signatures over time may have been facilitated by the repeated use of the same target words ("laten" with either "raken" or "wegstoten") and objects. These methodological aspects call for further investigation of word order conflicts in German and Dutch.

1.3 The present study

We chose to investigate sentence final double infinitives in main clauses. Double infinitives consist of an auxiliary verb (e.g. "willen" – want) and a main verb (e.g., "maken" – make), both in infinitive form, and Dutch and German differ in how they are ordered. Similar to the structure used by Davidson and Indefrey (2009), Dutch requires the auxiliary to precede the main verb (3), while German requires the auxiliary to be placed after the main verb (4). Note that in contrast to Davidson and Indefrey's structure, both Dutch and German allow only *one* order. A sentence containing the correct Dutch verb order would thus be incorrect in its literal translation in German, and a correct German sentence corresponds to what would be incorrect in Dutch. Using a structure with a clear "right" and "wrong" was important in order to get a robust native speaker baseline against which to compare the learners' performance.

(3) Ik heb het huiswerk niet **willen** **maken**.
** Ich habe die Hausaufgaben nicht wollen machen.*
(I have the homework not *want make*.)
[I did not want to do the homework.]

(4) * Ik heb het huiswerk niet **maken** **willen**.
Ich habe die Hausaufgaben nicht machen wollen.
(I have the homework not *make want*.)
[I did not want to do the homework.]

Knowing that Dutch and German are otherwise syntactically very similar, it is highly conceivable that German learners of Dutch will be influenced by German in forming and analyzing Dutch sentences containing these double infinitives. In the present paper, we ask the question whether and to what extent that really is the case, and in doing so, attempt to shed light on how conflicting syntactic rules in L1 and L2 are processed by learners.

We tested German university students that had been living in the Netherlands for less than two years. They were asked to read sentences containing the afore-mentioned structures in both their correct (3) and their incorrect order (4) and were instructed to judge them for grammaticality. ERPs for these sentences were time-locked to the second last verb (i.e., the first infinitive) and were compared to those of a native speaker control group. Furthermore, in order to make sure that our learners were sensitive to syntactic violations in Dutch at all, we also included a control condition. These control sentences contained violations of the inversion rule after subordinating conjunctions, which is implemented identically in Dutch and German (see section 2.3.1). Like the double infinitive sentences, control sentences thus contain verb order violations. In contrast to the critical sentences, they should be easy to process and are hypothesized to result in ERP signatures comparable to those of the Dutch native speakers.

1.3.1 Hypotheses

While we expected native-like ERP signatures for the control sentences, there are three possible outcomes for the sentences containing double infinitives.

(1) Despite the conflict between L1 and L2 syntactic rules, learners show native-like processing signatures in the form of a P600 to the incorrect order of the two infinitives in Dutch.

This would most likely go hand in hand with native-like behavioural performance and would support the few studies that suggest that native-like processing signatures are indeed possible even for structures that are implemented differently in two languages (Tokowicz & MacWhinney, 2005; White et al., 2012).

(2) Because of the rule conflict, learners show non-native ERP signatures to incorrect Dutch sentences. Among the possible non-native ERP signatures are:

— *Delayed and/or attenuated P600 (Rossi et al., 2006)*
— *N400 effect only (McLaughlin et al., 2010)*
— *Biphasic N400-P600 pattern (Tanner, Inoue, & Osterhout, 2014)*

Regardless of the nature of the signature, qualitatively or quantitatively different ERP signatures would reinforce the majority of earlier EEG studies on the influence of the L1 on L2 syntactic processing, and would add to the conclusion that differences in syntactic rules lead to online processing difficulties and ERP signatures different from those of native speakers. Non-native ERP signatures may be found either together with native-like behaviour or with non-native-like behavioural performance.

Our structure allows for a third possibility. As illustrated by Lemhöfer et al. (2014), a direct conflict between L1 and L2 syntactic rules may lead to wrong subjective representations and, consequently, to the reversal of online processing signatures in line with these subjective rather than objective representations.

(3) German learners of Dutch show ERP signatures in line with German grammar instead of Dutch grammar in the form of a P600 to the correct Dutch verb order instead of to the incorrect one.

Such a signature would indicate a very strong influence from the L1. This, however, is only likely to occur if participants make a lot of mistakes behaviourally.

1.3.2 Individual differences in susceptibility to L1 influence

How much a learner is affected by L1 preferences, and thus what kind of signature he/she shows to violations of infinitive verb order, may depend on the level of proficiency reached in Dutch. L2 proficiency has previously been shown to be an important factor in determining the native-likeness of L2 syntactic processing (McLaughlin et al., 2010; Tanner et al., 2014, 2013). More recent studies looking specifically at individual differences have additionally suggested the following factors to be important for the success of syntactic attainment both in L1 and L2: working memory (Nakano, Saron, & Swaab, 2010; Oines, Miyake, & Kim, 2012; Tanner et al., 2014, 2013), motivation to learn L2 (Tanner et al., 2014, 2013), degree of immersion in L2 (Flege & Liu, 2001; Frenck-Mestre, 2002) and frequency of L2 use (Tanner et al., 2014, 2013).

Next to the EEG measurements, we also acquired such individual difference measures in an attempt to better understand the ERP pattern we find. We included measures of L2 proficiency (self-rated and measured via a proficiency test and the standardized vocabulary test 'LexTALE'; Lemhöfer & Broersma, 2012), motivation to learn the L2 and variables describing the learners' degree of immersion in the L2 environment (e.g., frequency of use and amount of experience with Dutch).

2. Materials and methods

2.1 Participants

Twenty-four Dutch native speakers and 40 German-speaking learners of Dutch participated in the study. One Dutch and ten German participants had to be excluded due to excessive EEG artifacts, or because they had made too many grammaticality judgment mistakes (see section 3.2.1). Only the remaining 23 Dutch (age: *M* = 22.86, *SD* = 2.78, range = 19–29; 14 female) and 30 German participants (age: *M* = 20.80, *SD* = 2.83, range = 18–33; 25 female) entered the final analyses. All of these remaining participants were right-handed, had normal or corrected-to-normal vision, were brought up as monolinguals and reported no history of neurological impairment or dyslexia. They provided informed consent and received either course credit or vouchers (10 €/h) for their participation.

The German participants had been learning Dutch for at least half a year, but at the most two years. We recruited participants from different levels of proficiency in order to be able to look at individual differences in performance in a second step. German participants were living either in the Netherlands (*n* = 25) or close to the Dutch border in Germany (*n* = 5). All of them reported to have knowledge of other foreign languages, but listed Dutch as their currently most frequently used foreign language. All other results from a language background questionnaire are summarized in Table 1.

2.2 Behavioural session

Approximately one week (in days: *M* = 8.41, *SD* = 4.26, range = 2–24) prior to the EEG study, participants came in for a behavioural session, consisting of four to six tasks depending on the participants' mother tongue. This session took up to 30 or 75 minutes, respectively, and participants completed the tests in the following order: language background questionnaire, operation span test,

Table 1
Results from a language background questionnaire given to German learners of Dutch

	Mean	SD	Range
Age of first contact with Dutch (in years)	19.48	3.03	12–24
Amount of experience with Dutch (in months)	13.4	6.17	6–24
Immersion	0.33	0.21	0–1
Self-ratings (1 (low/rarely) to 5 (high/very often))			
Reading frequency	3.73	0.91	2–5
Speaking frequency	3.73	1.05	2–5
Listening frequency	4.43	0.62	3–5
Speaking proficiency	3.17	0.75	2–4
Listening proficiency	3.53	0.58	3–5
Writing proficiency	3.13	0.86	2–5
Reading proficiency	3.83	0.59	3–5

motivation questionnaire (Germans only), Dutch proficiency test (Germans only), LexTALE, and reading span test.

2.2.1 *Language background questionnaire*

At the very beginning, participants were prompted to fill in a questionnaire asking for background information concerning knowledge and usage of foreign languages. For the German participants, the questionnaire furthermore contained questions specifically regarding their knowledge of and experience with Dutch. A Dutch immersion score was calculated based on whether or not a participant lived in the Netherlands, whether they had a job there and whether they had a Dutch partner. The score was calculated as the percentage of positive answers to these three questions, thus resulting in a maximum score of 1 (full immersion) and a minimum score of 0.

2.2.2 *Working memory tests*

Two working memory tests were administered: one for non-verbal working memory (operation span test) and one for verbal working memory (reading span test). Tests were conducted in the participants' native language, but were otherwise identical for the two participant groups.

Operation span

The operation span test was a modified version of a test by Klaus, Oppermann, and Lemhöfer (in preparation). Participants had to judge mathematical equations for correctness and were asked to remember a set of numbers while doing so. A trial began with the presentation of an equation (Verdana, font size 24, bold, black), which participants had to judge for correctness within a maximum of 5000 ms. After each equation, a number appeared on screen for 800 ms, which participants were asked to remember. After four to seven trials, depending on the memory set size of the run, three question marks (Verdana, font size 36, bold, green) appeared on the screen prompting the participant to recall and repeat out loud all numbers he/she could remember. After recall, the experimenter initiated the next memory set.

Before the test phase, participants were familiarized with the task and went through three test runs with a memory set size of four numbers each. The actual test phase consisted of twelve runs with memory set sizes between four and seven numbers, three of each length. Equations and numbers were presented in a fixed, but random, order and memory set size alternated.

Reading span

The reading span test was used exactly as described in Klaus et al. (in preparation) and was very similar in design to the operation span test. Instead of solving equations, participants had to read sentences and judge whether they were plausible. Again, they had a maximum of five seconds to answer. After each sentence, a word appeared on the screen for 800 ms, which had to be remembered. After two to six trials, participants were asked to recall all words they remembered. Two training runs with a memory set-size of three words each preceded the test phase. The test phase itself consisted of 15 runs with memory set-sizes between two and six words, again three times each.

Working memory scores for both the operation and the reading span tests were calculated separately as the average of the mean proportions of correctly recalled items per memory set with scores between 0 (no items recalled) and 1 (recall score of 100%) (cf. Miyake et al., 2000).

2.2.3 Motivation questionnaire

German participants were also asked to answer 16 questions regarding their motivation to learn Dutch. Questions concerned the participants' general motivation to learn Dutch, their level of perfectionism, their level of confidence, and their willingness to learn from mistakes. Questions were phrased positively and negatively equally often and answers were given on a 5-point scale from "I fully agree" to "I strongly disagree". Participants received a score from 1 to 5 per answer, resulting in a maximal score of 20 per question type. Percentage scores were then calculated per question type and an average score was calculated per participant based on these four sub scores. Overall, higher scores indicate a higher degree of motivation to learn Dutch.

2.2.4 Proficiency test

German participants furthermore completed a Dutch proficiency test. The test was constructed by the author and her supervisor specifically for this study. It focused on assessing knowledge of false friends and grammatical structures that are different in the two languages, including the target structure of this research. It thus provides a measure of how good participants are at overcoming their L1 preferences. The test consisted of 58 multiple-choice questions concerning vocabulary (26x), grammar (22x) and idioms (10x). Scores were calculated as the percentage of correct responses. The questions were in part newly designed and in part compiled from a Bachelor thesis on idiom comprehension in German learners of Dutch (Hoff, 2012), and from two online proficiency tests by Transparent Language (http://www.transparent.com/learn-dutch/proficiency-test.html) and Klett (http://www2.klett.de).

2.2.5 LexTALE

Finally, both German and Dutch participants completed the Dutch version of the LexTALE, a standardized vocabulary test available in Dutch, English and German. The test consists of a short, simple and un-speeded visual lexical decision task in Dutch. The LexTALE score was calculated as: [(number of words correct/40*100) + (number of nonwords correct/20*100)] / 2. See Lemhöfer and Broersma (2012) for a detailed description of the English version, or see http://www.lextale.com/ for an online version of the Dutch test.

2.3 EEG session

2.3.1 Materials

In order to ensure that German participants would be able to understand the Dutch sentences, sentences only contained words that were either cognates or frequent non-cognates (with an average frequency of 505 occurrences per million in Dutch according to the SUBTLEX-NL database; Keuleers, Brysbaert, & New, 2010). Cognates are defined as translation equivalents with strong etymological overlap in German and Dutch.

Critical condition

Critical sentences contained the double infinitives described in the introduction and followed a fixed structure (see Table 2):

$$NP_{Sub} - V_{fin} \text{ (form of “hebben” [“have”])} - NP_{Obj} - AdvP - \text{(niet)} - V1_{infin} - V2_{infin}{}^{2}$$

This fixed structure licenses the two infinitives in both languages and was necessary to ensure roughly equal sentence complexity. To avoid monotony and to provide a certain degree of natural variation in sentence meaning, each sentence contained a different content verb and one of five modal auxiliaries: "kunnen" (can), "mogen" (may), "moeten" (must), "willen" (want), "laten" (let). Each of those five auxiliaries was used to build twelve sentences, resulting in a total of 60 critical sentences and five subconditions (for randomization see 2.3.2). Sentences also varied on whether or not they contained a negation, and whether subjects were singular or plural entities.

EEG measurements were time-locked to this first infinitive in all critical sentences, the moment in time when a mistake in verb order can first be detected. Sentences were between seven and twelve words long (*M* = 9.23) and target verbs (1st infinitive) were between five and eleven letters long (*M* = 6.78). Participants saw equal amounts of correct (30x) and incorrect (30x) Dutch sentences and sentence order and correctness were counterbalanced across participants.

---

[2] NP = noun phrase; V = verb; AdvP = adverbial phrase

Table 2
Examples for critical and control sentences with German and English translations

Condition	Example Sentence
Critical Condition - correct	Ik heb het huiswerk voor maandag niet kunnen doen. * Ich habe die Hausaufgaben für Montag nicht können tun. (I was not able to do the homework for Monday.) [I have the homework for Monday not can do.]
Critical Condition - incorrect	* Ik heb het huiswerk voor maandag niet doen kunnen. Ich habe die Hausaufgaben für Montag nicht tun können. (I was not able to do the homework for Monday.) [I have the homework for Monday not do can.]
Control Condition - correct	Elisa koopt graag nieuwe kleren maar zij heeft geen geld. Elisa kauft gern neue Kleider aber sie hat kein Geld. (Elisa likes to buy new clothes but she has no money.) [Elisa buys happily new clothes but she has no money.]
Control Condition - incorrect	* Elisa koopt graag nieuwe kleren hoewel zij heeft geen geld. * Elisa kauft gern neue Kleider obwohl sie hat kein Geld. (Elisa likes to buy new clothes even though she has no money.) [Elisa buys happily new clothes even though she has no money.]
Filler Condition – correct	Ik houd van de winter hoewel ik soms de sneeuw vervelend vind. Ich mag Winter obwohl ich manchmal den Schnee unangenehm finde. (I like winter although I sometimes find the snow annoying.) [I like winter although I sometimes the snow annoying find.]
Filler Condition – incorrect	*Ik houd van de winter maar ik soms de sneeuw vervelend vind. * Ich mag Winter aber ich manchmal den Schnee unangenehm finde. (I like winter but sometimes I find the snow annoying.) [I like winter but I sometimes the snow annoying find.]

Note. Target words are underlined. * indicates an incorrect sentence. Dutch sentences are on the first line, followed by German translations. English translations in round brackets; literal translations into English in square brackets.

Control condition

Apart from the sentences that contained a rule conflict in L1 and L2, participants also read sentences with violations of syntactic rules that are compatible in German and Dutch. To keep the control and critical sentences comparable in violation type and complexity, we also chose a verb order violation for control sentences.

We made use of coordinating and subordinating conjunctions, which differ in the verb order they require: subordinating conjunctions require the verb to be placed at the end instead of immediately after the subject of the subordinate clause. Incorrect control sentences contained a violation of this inversion rule and followed main clause verb order instead of the required inverted verb order (Table 2).

For each incorrect control sentence, we built a correct counterpart. These correct control sentences had the same sentence frame as incorrect controls (i.e., also followed main clause verb order), but instead contained a coordinating conjunction, which rendered the main clause verb order correct. Keeping the sentence frame equal for correct and incorrect control sentences was necessary in order to be able to measure on the same word in both types of sentences.

Both German and Dutch have meaning-equivalent conjunction pairs, for which one conjunction is coordinating (used for correct controls) and the other is subordinating (used for incorrect controls). We chose three such conjunction pairs and created an even number of sentences with each of them: want–omdat [because] (20); maar–hoewel [but/although] (20) and dus–waardoor [thus / whereby] (10).

In total, participants saw 50 control sentences, half of which were correct and half of which were incorrect. ERP measurements were always taken in the second half of the sentence on the finite verb, that means after the conjunction, which fell between the sixth and the ninth word of the entire sentence (M = 7.36). Sentences were between seven and twelve words long (M = 10) and, as for the critical sentences, correctness and sentence order were counterbalanced across participants.

Filler sentences

Given the design of the control sentences, participants may establish expectations concerning the correctness of a sentence based solely on its conjunction. To prevent this from happening, sentences with a correctness mapping opposite to that chosen for the control sentences were included as fillers. Accordingly, correct filler sentences contained subordinating conjunctions and correctly followed inverted verb order, whereas incorrect filler sentences contained coordinating conjunctions and incorrectly followed inverted verb order.

Participants saw 40 filler sentences (maar–hoewel (16); want–omdat (16); dus–waardoor (8)) – half of them correct and half of them incorrect. Sentence length was comparable to that of the control sentences and was between eight and twelve words (*M* = 10.17). No EEG measurements were taken for filler sentences.

2.3.2 List construction

We counterbalanced sentence order and sentence correctness across participants. For that purpose, four lists were created, each of which contained 150 sentences (60 criticals, 50 controls, 40 fillers). All four lists were pseudo-randomized such that no more than three correct or incorrect sentences and no more than two sentences from the same subcondition followed in immediate succession. Lists 1 and 2 were reversed in order to get lists 3 and 4.

2.3.3 Procedure

Participants were tested individually in a quiet room and were seated comfortably in a chair at a viewing distance of about 50 cm from a computer screen. After signing the consent and EEG screening forms, participants were instructed to read each sentence carefully and to judge it for grammatical correctness by pressing one of two buttons (left = incorrect, right = correct). Sentences were presented word by word in black letters (36 pt Arial) in the center of a grey screen using the stimulus presentation software 'Presentation' (Version 18.1, Neurobehavioural Systems). As illustrated in Figure 1, each word was presented for 450 ms and was followed by a 250 ms blank screen. Each sentence was preceded by a fixation cross for 500 ms and was followed by a grammaticality judgment question to which participants had to answer by pressing a button within a maximum of 10,000 ms. Each button press was followed by a 1000 ms blank screen before the next trial started.

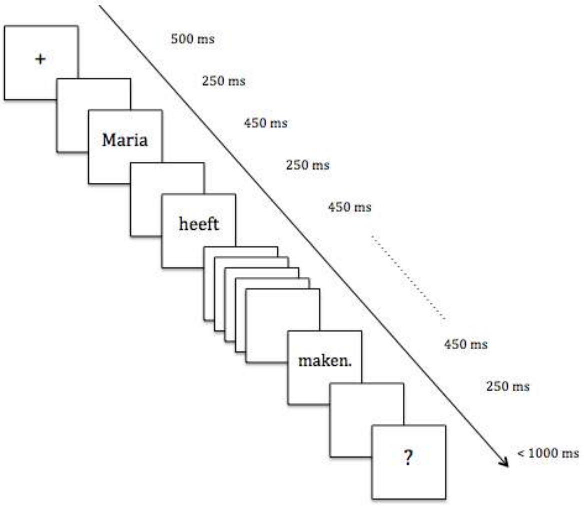

Fig. 1. EEG experimental design.

Experimental trials were preceded by 6 practice trials in order to familiarize participants with the task and the type of errors they would encounter during the experiment. The main experiment consisted of three blocks, each containing 50 trials. After the experiment, German participants were given a list of all sentences on paper and were asked to circle words that were unknown to them. The complete EEG session took 2–2.5 hours for the German and 2 hours for the Dutch participants.

2.3.4 EEG recording

Continuous EEG was recorded from 27 active Ag/AgCl electrodes embedded in an elastic cap (actiCAP, Brain Products) from standard 10–20 locations (Fig. 2) as well as from electrodes placed on the left and right mastoids and the forehead (ground). EEG was referenced on-line to the left mastoid and re-referenced off-line to the averaged activity over both mastoids. Vertical eye movements and blinks were recorded from a bipolar montage consisting of electrodes placed above and below the right eye, and horizontal eye movements were recorded from a bipolar montage consisting of electrodes placed on the left and right temples. The EEG was amplified with a BrainAmp DC / MR plus amplifier, digitized with a 500 Hz sampling rate and filtered online with a high cutoff at 125 Hz and a low cutoff at 0.01 Hz. Impedances for EEG electrodes were kept below 10 kΩ and for EOG electrodes below 15 kΩ.

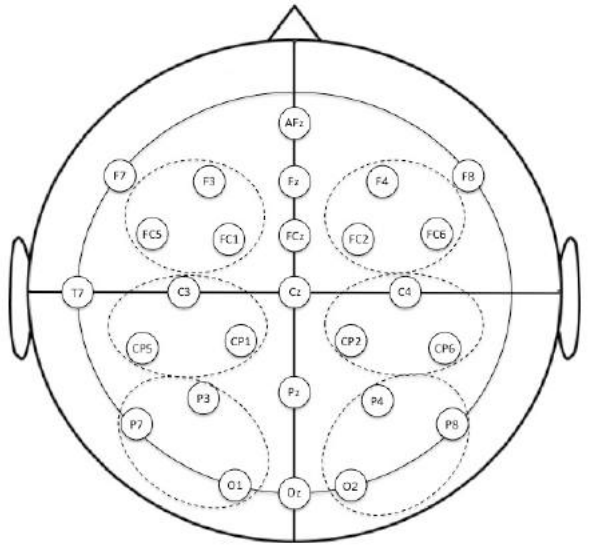

Fig. 2. Electrode positions and regions chosen for statistical analysis.

2.3.5 EEG data analysis

All off-line EEG data processing was done using BrainVision Analyzer 2.0.2.5754 (Brain Products GmbH, 2012). The EEG and EOG signals were segmented into epochs from 200 ms before until 1000 ms after presentation of the target word. The data were low-pass filtered at 30 Hz and high-pass filtered at 0.02 Hz (12 dB/oct roll-off). Filtering was done with Butterworth zero phase filters as implemented in BrainVision Analyzer 2.0, which were applied to the continuous EEG data to avoid edge artifacts. Blink detection and ocular correction were also performed on the continuous EEG data, using the Gratton & Coles algorithm in Brain Vision Analyzer 2.0.2.5757 (Gratton, Coles, & Donchin, 1983). Baseline correction was carried out after segmentation and was based on the average EEG activity in the 200 ms interval before target onset. Finally, trials with amplitudes below −100 μV or above 100 μV, trials with a difference between minimum and maximum voltage of less than 0.5 μV in one or more EEG electrodes, and trials with a voltage change of more than 75 μV between adjacent sample points or a difference in voltage of more than 150 μV within the entire segment, were removed semi-automatically, that is, after inspection (6.64% of critical and control trials).

In order to identify windows for ERP effects in the two groups, we first conducted a time-course analysis via t-tests (p = .05) in consecutive time windows of 50 ms for six regions (left anterior, left middle, left posterior, right anterior, right middle, right posterior; see Fig. 2). Following Lemhöfer et al. (2014), only intervals of at least two consecutive 50 ms windows with significant effects of correctness were chosen for further analysis.

Due to the different time-windows of effects for the German and Dutch participants, the two groups were analysed separately. For both groups, the lateral electrodes were collapsed into 6 regions and analysed using repeated-measures MANOVAs with the factors Hemisphere (right vs. left), Region (anterior, middle, posterior) and Correctness (correct vs. incorrect). Significant interactions including the factor Correctness were followed up by planned simple MANOVAs. For all analyses, only multivariate test statistics and only significant effects, including the experimental factor Correctness, are reported. Other follow-up analyses will be described in the results section.

3. Results

3.1 Separate behavioural session

Table 3 shows the results from the separate behavioural tests for 23 Dutch and 30 German participants. Together with selected variables from a language background questionnaire (Table 1 in section 2.1) and the EEG results (section 3.2.2), correlation analyses were conducted. The correlation coefficients can be found in Table 4. Overall, behavioural scores correlated positively with each other and reinforced common intuitions: higher motivation to learn Dutch and longer overall experience with Dutch, for instance, were associated with better performance on the vocabulary and grammar based proficiency test. Moreover, scores from this proficiency test were highly positively correlated with participants' proficiency self-ratings and marginally positively correlated with scores from a validated vocabulary test (LexTALE), thus reinforcing the validity of our Dutch proficiency test (see Table 4). For the complete correlation matrix refer to Table 4.

3.2 EEG session

3.2.1 Grammaticality judgments during the EEG

Mean grammaticality judgment error rates for the entire group (40 German and 24 Dutch participants) can be inspected in Table 5. The data revealed quite striking individual differences for the German participants, with error rates ranging from 0% to over 60%. Participants with too many mistakes (and

Table 3

Results from the behavioral tests for the 30 German and 23 Dutch participants

	German			Dutch		
	Mean	SD	Range	Mean	SD	Range
Operation span (0–1)	0.77	0.14	0.48–0.97	0.86	0.09	0.61–1
Reading span (0–1)	0.79	0.13	0.46–0.93	0.76	0.12	0.51–0.95
LexTALE (in %)	66.34	9.03	42.5–82.5	93.53	3.63	85–98.75
Motivation (in %)	75.96	8.88	57.5–92.5	-	-	-
Proficiency test (in %)						
Total (in %)	71.73	7.60	55.17–84.48	-	-	-
Grammar (in %)	84.39	9.67	68.18–100	-	-	-
Vocab (in %)	71.66	11.40	50–92.31	-	-	-
Idiom (in %)	44.00	14.29	20–70	-	-	-

Table 4

Correlation matrix for EEG and behavioral results for the 30 German participants

	N400	P600$_{CR}$	RDI	RST	OST	Prof	LexTale	Mot	Imm	Prof$_{SF}$	Freq	NL$_{Exp}$	D'$_{CR}$
N400	-												
P600$_{CR}$	**−.516***	-											
RDI	−.818	.915	-										
RST	−.193	.297	.291	-									
OST	−.195	.254	.262	.553	-								
Prof	−.014	-.106	−.064	.308	.273	-							
LexTale	.069	.027	−.015	.419	.400	.450	-						
Mot	.213	−.336	−.326	.088	.049	**.501***	.376	-					
Imm	−.084	.108	.112	.433	.163	.387	.252	.143	-				
Prof$_{SF}$	−.194	.105	.162	.111	.181	**.548***	.153	.374	.312	-			
Freq	−.079	-.266	−.141	.125	−.131	.297	−.070	.448	.134	.333	-		
NL$_{Exp}$	.069	−.021	−.046	.394	.323	**.527***	.102	.349	**.491***	.273	.342	-	
D'$_{CR}$	.014	.050	.027	.203	.132	.401	.173	.327	.382	.285	.272	**.486***	-

Note. Significance level (Bonferroni corrected): * $p < .004$. Abbreviations: P600$_{CR}$ = P600 to critical sentences; RDI = Response Dominance Index; RST = Reading Span; OST = Operation Span; Prof = Proficiency Test Score; Mot = Motivation to learn Dutch; Imm = Immersion; Prof$_{SF}$ = self-rated proficiency; Freq = Frequency of use of Dutch; NL$_{Exp}$ = Log-Experience with Dutch; D'$_{CR}$ = d' scores for critical sentences.

Table 5

Mean grammaticality judgment error rates for all 40 German and 24 Dutch participants

	German			Dutch		
	Mean	SD	Range	Mean	SD	Range
Critical condition (%)	13.17	15.92	0–61.67	3.4	2.88	0–10
Control condition (%)	11.74	11.39	0–50	4.08	4.11	0–16

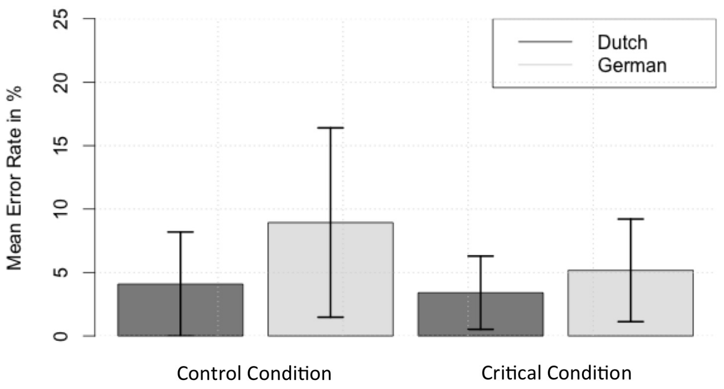

Fig. 3. Mean grammaticality judgment error rates per condition for the remaining 30 German and 23 Dutch participants (in %).

thus too few correct trials per condition) cannot enter the EEG analysis. For that reason, we excluded participants with cell sizes of less than 15 trials per condition and correctness, or participants with more than 15% grammaticality judgment errors for the critical sentences. This resulted in the rejection of 1 Dutch and 10 German participants from the final analyses. For the remaining 30 German and 23 Dutch participants, cell size exceeded the minimum of 15 by far both for the critical and the control sentences (cell size for critical sentences [max. 30] *M* = 28, *SD* = 1.33, range = 21–30; for controls [max. 25]: *M* = 22.38, *SD* = 2.09, range = 15–25).

Mean grammaticality judgment error rates for the remaining participants, both the critical and the control sentences, are shown in Figure 3. The mean percentage of errors for Dutch participants was 3.4% for the critical sentences (*SD* = 2.95; range = 0 – 10%) and 3.57% for the control sentences (*SD* = 3.3; range = 0–14%). The difference between these two error rates was not significant (*t*(22)= 0.23, *p* = .815), illustrating that for native speakers control and critical sentences were equally easy.

German participants had an error rate of 5.16% for the critical sentences (*SD* = 4.04; range = 0–15%) and 8.93% for the control sentences (*SD* = 7.45; range = 0–24%). The higher error rate for the control compared to the critical sentences is somewhat surprising (*t*(29) = 2.24, *p* = .03). Closer inspection revealed that it was driven by only a few participants and by a subset of control sentences, namely those containing the conjunction 'dus' ("thus"). 'Dus' is preferably used with a slightly different word order in Dutch than in German (initial inversion), which lead some participants to consistently judge correct Dutch 'dus'-sentences as incorrect. When 'dus' sentences are excluded, the error rates for critical and control sentences are not significantly different (*t*(29) = −0.76, *p* = .45).

In a second step, between-group comparisons were conducted to evaluate in how far the German participants' performance was already native-like. For the control sentences, the mean error rate for the German group was significantly higher than that for the Dutch native speaker group for reasons explained above (*t*(51) = 3.39, *p* = .001). For the critical sentences, however, between-group comparisons revealed no significant difference between the L2 learners and the native speakers (*t*(51) = 1.75, *p* = .085). This suggests that despite the conflict between L1 and L2 syntactic rules, German learners of Dutch had learned the correct Dutch order of the two infinitives and were able to behaviourally judge correctness for these sentences on a native-like level.

Subsequently, d' scores were calculated per participant and condition [d' = *Z*(hit rate) − *Z*(false alarm rate)] (Table 6). As can be deducted from the correlation matrix (Table 4), higher d' scores for the critical sentences were associated with longer overall experience with Dutch (*r*(28) = 0.49 *p* = .006). D' scores for the control sentences, in turn, were explained by stronger motivation to learn Dutch (*r*(28) = 0.506, *p* = .004) and longer experience with Dutch (*r*(28) = 0.482, *p* = .006) (Suppl. Table 2).

Table 6
d' scores per condition and group for 30 German and 23 Dutch participants

	German			Dutch		
	M	SD	Range	M	SD	Range
Control condition	3.19	0.94	1.41–4.65	3.82	0.61	2.17–4.65
Critical condition	3.57	0.65	2.34–4.65	3.83	0.58	2.61–4.65

Note. A score of 0 reflects chance level performance on the grammaticality judgment task during the EEG, a score of 2.5 or higher indicates very high levels of grammatical sensitivity (i.e., proportion correct over 0.90, Macmillan & Creelman, 2005).

Thus, there seems to be a relation between learners' level of immersion into the Dutch language and their degree of grammatical sensitivity to syntactic violations on a behavioural level.

3.2.2 EEG results

All subsequent analyses are based on correct trials only (i.e., correct grammaticality judgments). Statistical test results for all groups and conditions are summarized in Table 7, which will be referred to at various points in the remainder of the results section.

Control Condition

Control Condition – Dutch native speakers (L1 group)

The grand-averaged waveforms for the native speakers in the control condition are shown in Figure 4. Visual inspection and a time-course analysis revealed an effect of correctness in a time window from 550 ms until 950 ms after target onset. During this time window there was a large, widely distributed positive shift in the EEG signal to incorrect as compared to correct sentences, which was significant bilaterally over central and posterior electrode sites (see Table 7). Both in terms of latency and scalp distribution, this effect corresponds to the classical P600 effect.

Visual inspection additionally seems to suggest the existence of an early left anterior negativity (FC5, ~200–400 ms). However, this effect did not reach significance in the time-course analysis, which is why it is not analysed any further.

Control Condition – German learners of Dutch (L2 group)

The ERP waveforms for the L2 speakers in the control condition can be seen in Figure 5. Similar to the native speakers for this condition, we found a significant positive going shift for incorrect as compared to correct sentences in a window from 500 ms to 1000 ms. As for the native speakers, this positivity was significant over central and posterior electrode sites in both hemispheres and can be classified as a P600 effect (Table 7).

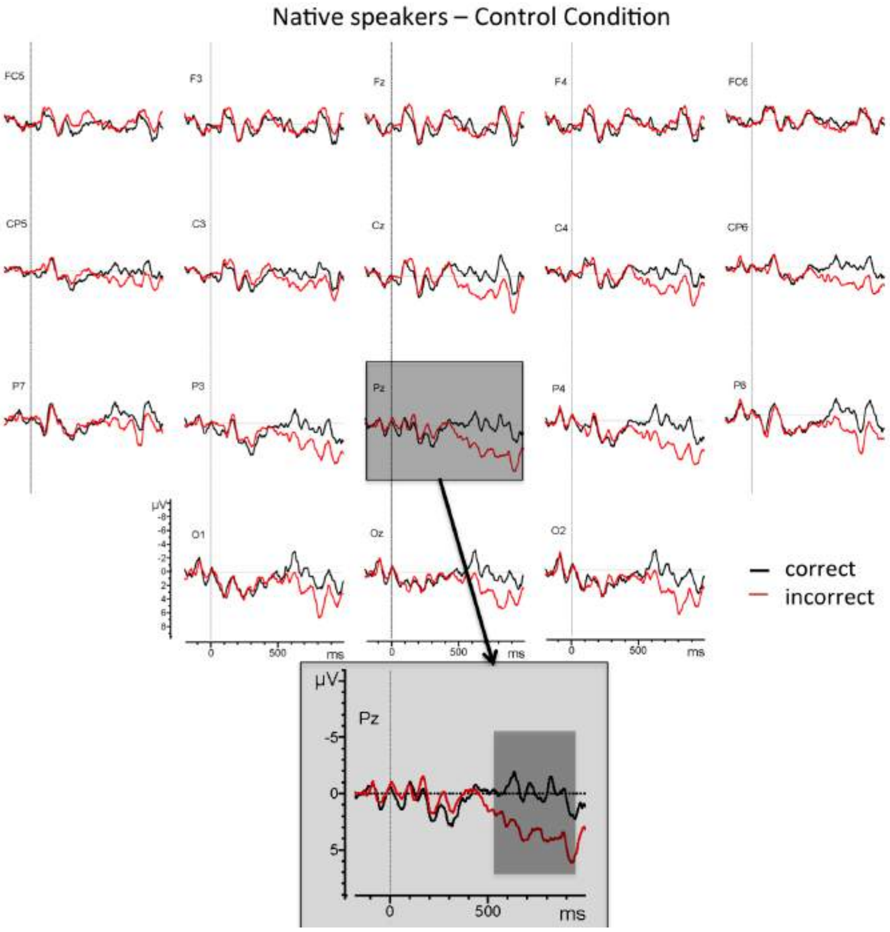
Native speakers – Control Condition

Fig. 4. Grand–averaged ERP waveforms for the critical verb in the control sentences for 23 Dutch native speakers. The graph shows a representative selection of midline and lateral electrodes, and zooms in on Pz, on which the effect is marked in gray in a time-window from 550-950 ms.

Comparison of L1 and L2 group for Control Condition

The effects found for the two groups look very alike. A MANOVA with Group as a between-subject factor in the overlapping time-window from 550–950 ms revealed that although the main effect of Correctness and the interactions between Correctness x Region and Correctness x Hemisphere were found again, there were no significant effects involving the factor Group (all *p*s > .26). This suggests that, disregarding the minimal latency difference, the effects shown by the two groups did not differ in terms of scalp distribution or amplitude. The P600 shown by the learner group can consequently be described as native-like.

Summary of Control Condition

Both the learner and the native speaker group were sensitive to violations of the inversion rule after subordinating conjunctions in Dutch and displayed large, centro-posterior P600 effects in largely overlapping time-windows. The high similarity in ERP signatures for the two groups suggests that the L2 group engaged in native-like processing for structures that are implemented similarly in L1 and L2.

Critical Condition

For the critical sentences the target analysis window only covers the first infinitive, which is the point in time when a violation of verb order first becomes apparent. Longer windows, covering both the first and the second infinitive, can be inspected in the supplementary materials (Suppl. Figs. 1–2), but will not be analysed and discussed in the remainder of this thesis.

Critical Condition – Dutch native speakers (L1 group)

The grand-average for the native speakers for the critical sentences is displayed in Figure 6. As for the control sentences, visual inspection and a time-course analysis revealed a significant positive deflection for incorrect critical sentences as compared to correct ones, this time in a window from 450–950 ms after onset of the target verb. The effect was widely distributed and, in fact, significant over the whole brain (see Table 7). Dutch native speakers

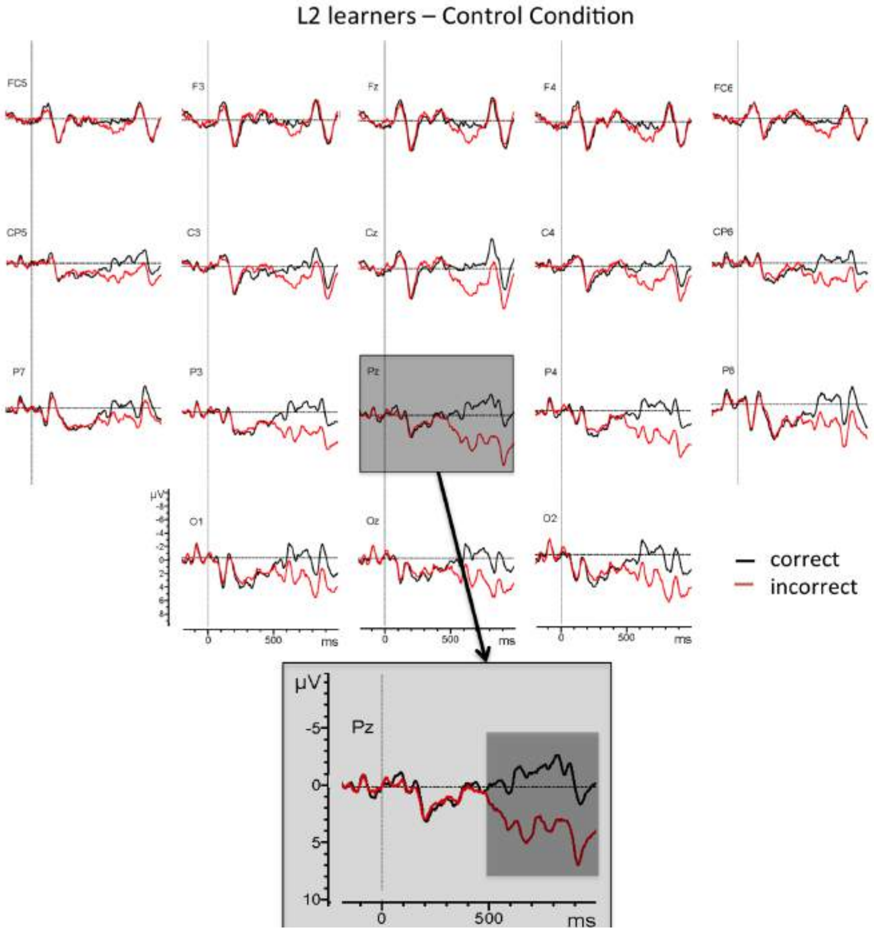

Fig. 5. Grand–averaged ERP waveforms for the critical verb in the control sentences for the 30 L2 learners (German learners of Dutch). The graph shows a representative selection of midline and lateral electrodes, and zooms in on Pz, on which the effect is marked in gray in a time-window from 500-1000 ms.

Table 7

Multivariate test statistics for all groups, conditions, and effect windows for 23 Dutch and 30 German participants

		Dutch group				German group							
		Control Condition		Critical Condition		Control Condition				Critical Condition			
		550–950 ms		450–950 ms		500–1000 ms		100–250 ms		250–600 ms		800–1000 ms	
Midline Analysis	α *level* *	F(df) / t(df)	p	F(df) / t(df)	p	F(df) / t(df)	p	F(df) / t(df)	p	F(df) / t(df)	p	F(df) / t(df)	p
Corr	.05	$F(1,22) = 8.47$	.008	$F(1,22) = 26.98$	.000	$F(1,29) = 23.98$	.000	-	-	$F(1,29) = 4.49$	.043	-	-
Corr x Elec	.05	$F(5,18) = 11.53$	.000	$F(5,18) = 8.5$	.000	$F(5,25) = 9.22$	.000	-	-	-	-	$F(5,25) = 7.65$	.000
AFz	.008	-	-	-	-	-	-	-	-	-	-	-	-
Fz	.008	-	-	-	-	-	-	-	-	-	-	-	-
FCz	.008	-	-	$t(22) = -4.3$	.000	$t(29) = -3.29$	.003	-	-	-	-	-	-
Cz	.008	$t(22) = -3.97$	.001	$t(22) = -5.99$	.000	$t(29) = -4.43$	.000	-	-	-	-	-	-
Pz	.008	$t(22) = -5.69$	.000	$t(22) = -6.51$	.000	$t(29) = -8.26$	.000	-	-	-	-	$t(29) = -4.77$	.000
Oz	.008	$t(22) = -4.33$	.000	$t(22) = -3.52$	.002	$t(29) = -5.23$	.000	-	-	-	-	$t(29) = -5.89$	.000
Cluster Analysis													
Corr	.05	$F(1,22) = 12.28$	.002	$F(1,22) = 35.48$	.000	$F(1,29) = 38.35$	.000	-	-	$F(1,29) = 6.88$	.014	$F(1,29) = 5.19$	.03
Corr x Hemi	.05	$F(1,22) = 4.96$	.037	$F(1,22) = 4.39$	.048	$F(1,29) = 6.15$	.019	$F(1,29) = 8.89$	.006	$F(1,29) = 6.88$	.014	-	-
Left Corr	.025	$F(1,22) = 10.11$	.004	$F(1,22) = 27.35$	.000	$F(1,29) = 28.01$	.000	$F(1,29) = 9.74$	.004	$F(1,29) = 11.87$	.002	-	-
Corr x Region†	.025	$F(2,21) = 12.28$	.000	$F(2,21) = 14.25$	.000	$F(2,28) = 17.79$	.000	-	-	$F(2,28) = 6.43$	.005	-	-
Right Corr	.025	$F(1,22) = 12.47$	.002	$F(1,22) = 39.67$	.000	$F(1,29) = 42.61$	.000	-	-	-	-	-	-
Corr x Region†	.025	$F(2,21) = 21.71$	.000	$F(2,21) = 9.88$	.001	$F(2,28) = 15.01$	.000	-	-	-	-	-	-
Corr x Region	.05	$F(2,21) = 21.41$	.000	$F(2,21) = 15.4$	.000	$F(2,28) = 20.39$	.000	-	-	$F(2,28) = 4.98$	.014	$F(2,28) = 17.49$	.000
Anterior Corr	.016	-	-	$F(1,22) = 14.19$	.001	-	-	-	-	-	-	-	-
Corr x Hemi†	.016	-	-	-	-	-	-	-	-	-	-	-	-
Central Corr	.016	$F(1,22) = 18.14$	.000	$F(1,22) = 37.68$	.000	$F(1,29) = 39.95$	.000	-	-	-	-	$F(1,29) = 7.23$	.012
Corr x Hemi†	.016	-	-	-	-	-	-	-	-	-	-	-	-
Posterior Corr	.016	$F(1,22) = 21.9$	.000	$F(1,22) = 28.89$	.000	$F(1,29) = 46.02$	.000	-	-	$F(1,29) = 11.14$	.002	$F(1,29) = 18.43$	.000
Corr x Hemi†	.016	-	-	$F(1,22) = 7.87$	.01	$F(1,29) = 9.47$	.005	-	-	$F(1,29) = 10.64$	.003	-	-
Corr x Hemi x Region	.05	-	-	-	-	-	-	-	-	$F(2,28) = 3.52$	.043	-	-
Left Anterior Corr	.008	-	-	$F(1,22) = 9.61$	.005	-	-	-	-	-	-	-	-
Left Central Corr	.008	$F(1,22) = 14.6$	.001	$F(1,22) = 30.73$	.000	$F(1,29) = 29.78$	.000	-	-	$F(1,29) = 9.79$	.004	-	-
Left Posterior Corr	.008	$F(1,22) = 20.83$	.000	$F(1,22) = 22.71$	.000	$F(1,29) = 40.77$	.000	-	-	$F(1,29) = 19.97$	.000	-	-
Right Anterior Corr	.008	-	-	$F(1,22) = 17.57$	.000	-	-	-	-	-	-	-	-
Right Central Corr	.008	$F(1,22) = 17.86$	.000	$F(1,22) = 38.7$	.000	$F(1,29) = 43.36$	.000	-	-	-	-	-	-
Right Posterior Corr	.008	$F(1,22) = 19.64$	.000	$F(1,22) = 31.82$	.000	$F(1,29) = 43.88$	.000	-	-	-	-	-	-

Note. * Bonferroni corrected α-levels. Abbreviations: Corr = correctness, Elec = electrode, Hemi = hemisphere. Only significant effects are printed (for complete multivariate statistics see Supplementary Table 1). † Interactions of correctness with region or hemisphere within one hemisphere or region were followed up with separate analyses of clusters, which are reported below the three-way interaction term for reasons of space, even if this term was not significant by itself.

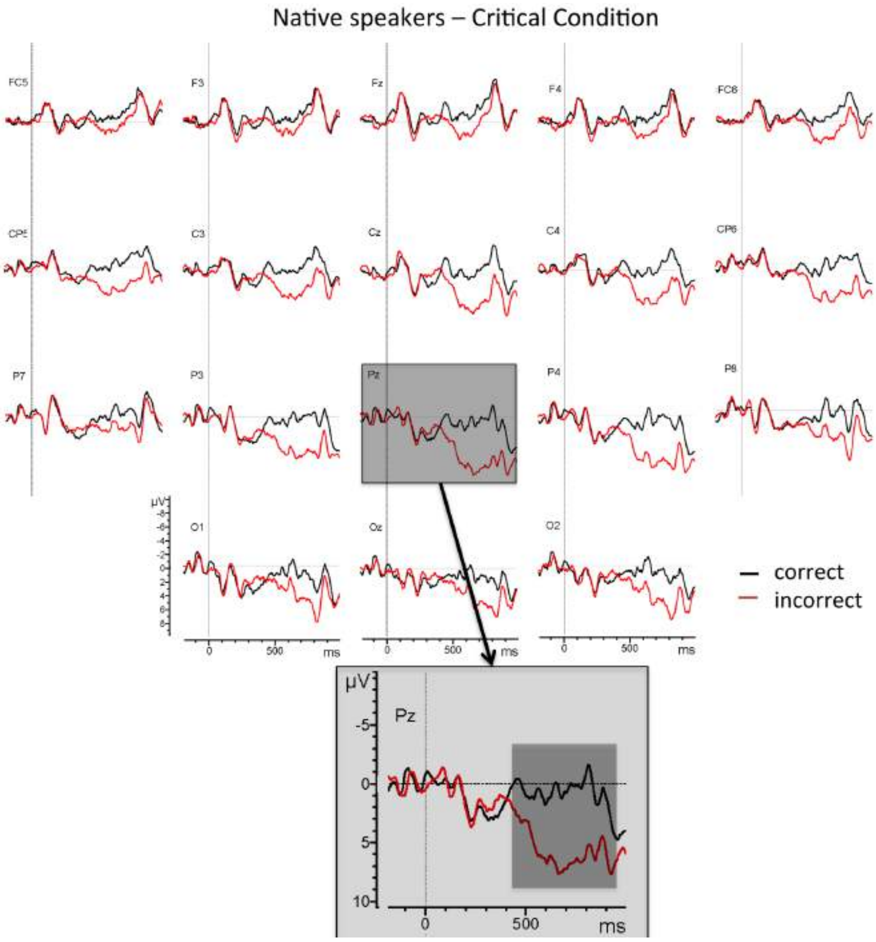

Fig. 6. Grand–averaged ERP waveforms for the critical verb in the critical sentences for the 23 Dutch native speakers. The graph shows a representative selection of midline and lateral electrodes, and zooms in on Pz, on which the effect is marked in gray in a time-window from 450-950 ms.

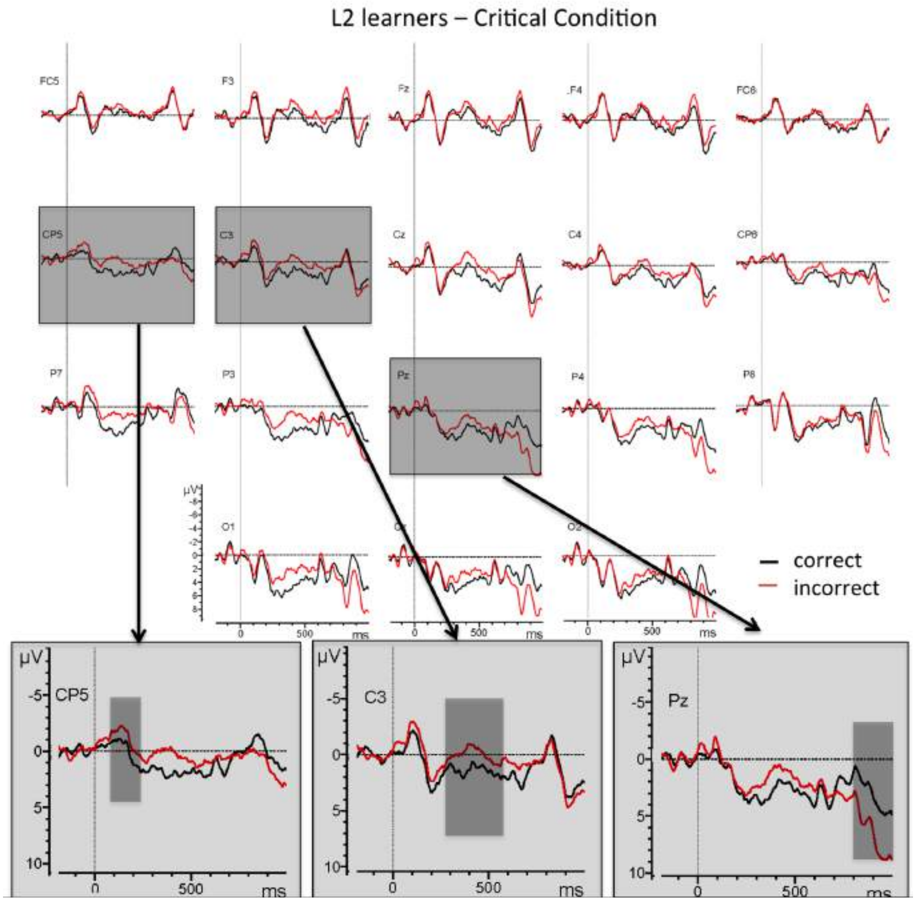

Fig. 7. Grand–averaged ERP waveforms for the critical verb in the critical sentences for the 30 L2 learners (German learners of Dutch). The graph shows a representative selection of midline and lateral electrodes, and zooms in on three electrodes, on which the effects are marked in gray in the following time-windows: 100-250 ms, 250-600 ms and 800-1000 ms.

thus showed a robust P600 effect to violations of infinitive word order.

Again, visual inspection seemed to suggest an additional small N400-like effect prior to the P600 (visible at central and posterior midline sites at 300–400 ms). However, again, this effect did not reach significance in a stepwise (50 ms) time-course analysis and was thus not further investigated.

Critical Condition – German learners of Dutch (L2 group)

ERP waveforms for the critical sentences for the learner group are depicted in Figure 7. A 50 ms time-course analysis revealed two effects: an early and prolonged negativity from 100–600 ms and a late positivity from 800–1000 ms. Based on visual inspection, however, we decided to split the negativity into two time windows from 100–250 ms and from 250–600 ms. This split represents the fact that visually the very early effect seemed to be restricted to the left hemisphere, while the later negativity seemed more widely distributed. Moreover, the very early negativity appeared temporally separated from the later unfolding negativity. All three time-windows will be analysed in turn.

Critical Condition – L2 group – 100–250 ms

Statistical analyses within the first window, from 100–250 ms, revealed a significant negativity over the left hemisphere only (see Table 7). Possible interpretations of this effect will be presented in the discussion section. We will henceforth refer to this effect as an early left negativity.

Critical Condition – L2 group – 250–600 ms

Much like for the first time-window, we found a significant negative deflection in ERP waveforms to incorrect critical sentences from 250–600 ms. This negativity was significant over the midline as well as over left central and left posterior electrode sites (see Table 7). Given its latency and scalp distribution, we suggest to call this effect an N400-like effect. The N400 component is traditionally found in a time-window from roughly 200–600 ms, is largest over centro-parietal electrode sites and is usually reported in response to semantic violations (Kutas & Federmeier, 2011). Why we find such an effect in the context of a syntactic violation remains to be discussed and will be explored in the discussion section.

Critical Condition – L2 group – 800–1000 ms

From 800 to 1000 ms after stimulus onset, we observed a significant positivity in the ERP signal to incorrect as compared to correct sentences. This positivity reached significance over central and posterior electrode sites, but was most pronounced over the posterior portion of the scalp (Table 7). Despite its late onset, but based on its topography and polarity, we interpret this effect to reflect a delayed P600. Visual inspection of a longer time window including the second infinitive reinforced this interpretation as the positivity continues throughout most of the second verb (Suppl. Figs. 1–2).

Comparison of L1 and L2 group for the Critical Condition

Unlike the ERPs in the control condition for which we reported native-like processing signatures in the L2 group, the ERP signatures for the critical sentences clearly differ between the two groups. While native speakers have a robust P600, German participants show an onset-delayed, smaller and less-distributed P600 preceded by an N400 and an early left negativity. The different time-windows make a direct comparison between the groups difficult. To nevertheless get a better grasp of the group differences, we conducted MANOVAs with Group as a between-subject factor for all overlapping (800–950 ms) and non-overlapping (100–250 ms, 250–600 ms, 450–800 ms) component time-windows. An overlapping time-window is defined as a window during which both groups showed an effect of correctness in the same direction.

For the overlapping P600 time-window (800–950 ms), no effects involving Correctness and Group reached significance (all ps $> .084$), suggesting thus that the P600 within this late time-window was comparable for the two groups both in terms of magnitude and topography. For the corresponding non-overlapping P600 time-window (450–800 ms), instead, there was a significant interaction of Group and Correctness ($F(1, 51) = 31.92, p = .000$) as well as of Correctness, Region and Group ($F(2, 50) = 8.45, p = .001$). Subsequent analyses revealed that while Dutch native speakers showed a widely distributed positivity that was significant all over the scalp (main effect: $F(1, 22) = 34.75, p = .000$; interaction Correctness x Region: $F(2, 21) = 12.01, p = .000$; separate analyses for regions: anterior: $F(1, 22) = 17.35, p = .000$; Central: $F(1, 22) = 35.47, p = .000$; posterior: $F(1, 22) = 25.2, p = .000$), the learner group showed no significant effects involving the factor Correctness (all ps $> .096$). This shows that the difference in P600 onset for the two groups was indeed significant.

For the N400 time-window (250–600 ms), there was a significant interaction of Correctness and Group ($F(1, 51) = 12.32$, $p = .001$) as well as of Correctness, Region and Group ($F(2, 50) = 3.47$, $p = .039$). Follow-up analyses for the two language groups revealed that, while the German group had a significant negativity over left central and posterior electrode sites (see above), Dutch native speakers showed a significant positivity, which was most likely driven by the fact that the window analysed here includes the beginning of the P600 (main effect of Correctness: $F(1, 22) = 7.44$, $p = .012$). Together, this shows that the N400 was indeed unique to the learner group.

A similar story can be told for the time-window of the early negativity (100–250 ms). Here, again, there was a significant interaction of Group and Correctness ($F(1, 51) = 4.73$, $p = .034$). Follow-up analyses showed that the German group had a significant negativity over the left hemisphere (see results above), the Dutch native speakers, however, showed no significant effects involving the factor Correctness (all ps > .219). The early negativity was thus also unique to the learner group.

Summary of Critical Condition

Violations of infinitive verb order in Dutch sentences elicited a robust P600 effect in Dutch native speakers, but a biphasic N400–P600 pattern in German learners of Dutch, preceded by an early left negativity. The qualitative differences in ERP patterns between the two groups suggest that the learner group processed sentences containing conflicting syntactic rules in Dutch and German in a different way, despite native-like behavioural performance. This biphasic pattern, furthermore, stands in stark contrast to the robust P600 effect that the learner group showed in response to the violations in the control sentences.

Follow-up analyses – Biphasic N400–P600 pattern

Biphasic N400-P600 patterns to syntactic violations in L2 learner populations have been reported before by Tanner et al. (2014, 2013). For these studies, the biphasic pattern on the grand average level turned out to be a misrepresentation of individual ERP signatures: L2 learners showed either an N400 or a P600, rather than both. In order to see whether our results could replicate this finding, individual N400 and P600 effect magnitudes were calculated and correlated with one another. The N400 effect magnitude was calculated as the mean activity in the correct critical sentences minus the incorrect critical sentences from 250 to 600 ms over a left central and posterior region of interest (C3,

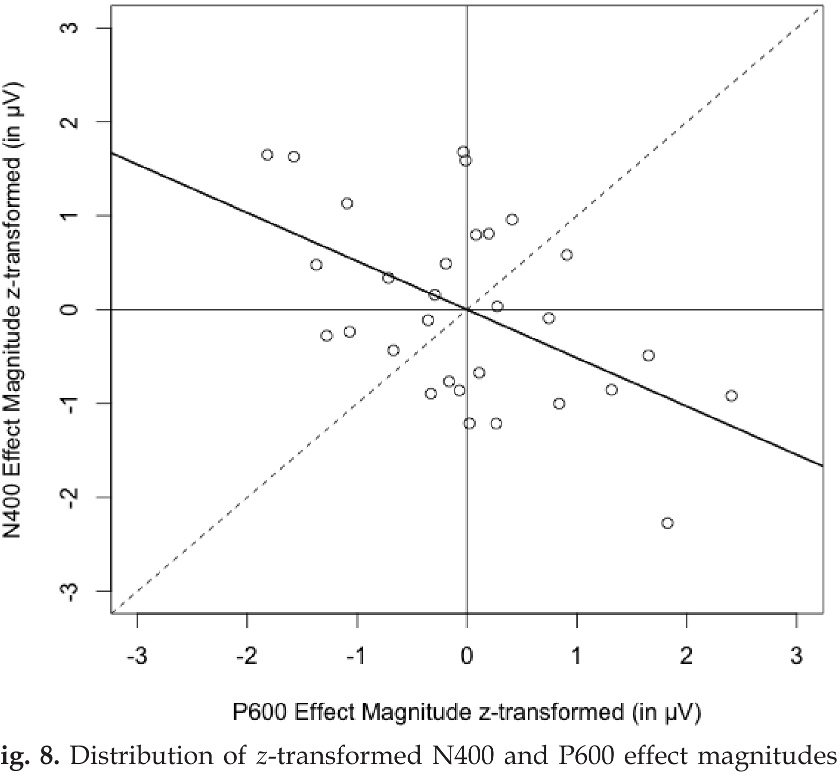

Fig. 8. Distribution of z-transformed N400 and P600 effect magnitudes across participants. Each point represents a single participant. The solid line shows the regression line. The dashed line represents equal N400–P600 effect magnitudes.

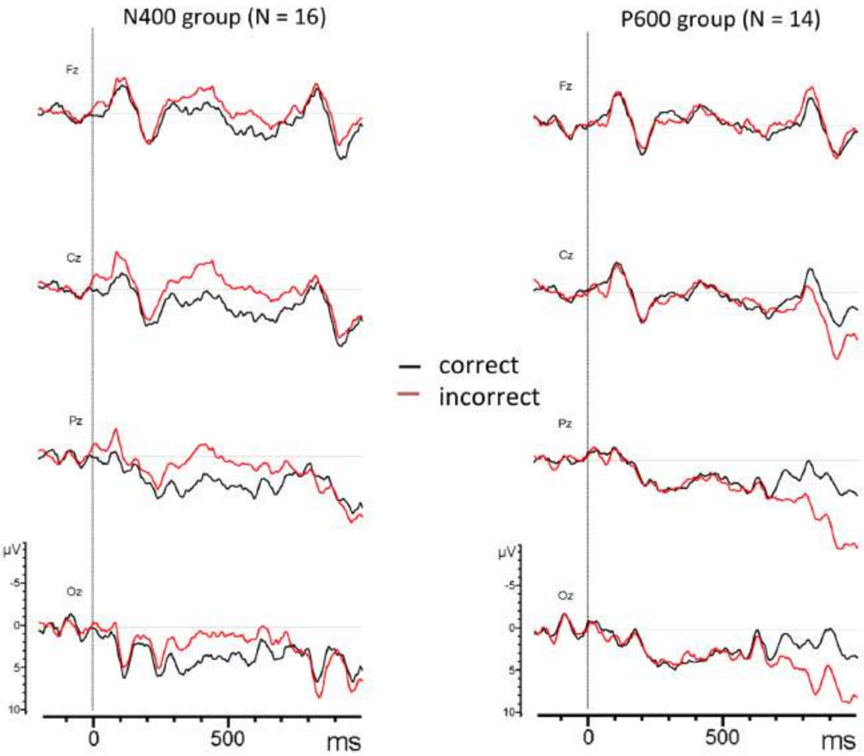

Fig. 9. Grand-averaged ERP waveforms for the N400 and the P600 dominant groups for four representative midline electrodes.

CP5, CP1, P7, P3, O1). The P600 effect magnitude was calculated as the mean amplitude in the incorrect minus the correct critical sentences from 800 to 1000 ms over left and right posterior electrode sites (P3, Pz, P4, O1, Oz, O2). Both magnitude scores were *z*-transformed before entering correlation analyses.

Results showed that the two effects were significantly negatively correlated ($r(28) = -0.516$, $p = .001$, see Fig. 8): participants tended to show either ERP signatures with a negativity-dominance (dots above/left to the dashed line in Fig. 8) or with a positivity-dominance (dots below/right to the dashed line in Fig. 8) such that as one effect increased, the other decreased. It should be noted, however, that the ERP dominance pattern is continuous and that some participants showed a biphasic pattern rather than a clear positivity/negativity.

Following Tanner et al. (2014), we subsequently averaged ERPs for participants with an N400-dominant and participants with a P600-dominant pattern (using the dashed line as a divider). Figure 9 shows the averaged ERP waveforms for these two groups for four representative midline electrodes. While the P600-dominant group showed a clear, albeit delayed P600 effect and no negativity, the N400-dominant group showed virtually no P600, but instead a prolonged negativity spanning almost the entire time window. The clearly distinct ERP waveforms for the two groups reveal stark differences in how individual participants deal with sentences containing rule conflicts in L1 and L2. We will elaborate on possible interpretations of these differences in the discussion section.

Finally, in order to understand what drove these ERP differences, we correlated the individual N400 and P600 effect magnitudes, as well as a response dominance index (RDI) with a set of behavioural scores (sections 3.1 & 3.2.1). The RDI was calculated as the difference between the *z*-transformed P600 and N400 effect magnitudes described above (modification of a formula used by Tanner et al, 2014). As can be seen upon inspection of the correlation matrix in Table 4, none of the behavioural scores could explain the dominance pattern or the ERP effect magnitudes.

4. Discussion

The present study investigated the effect of the first language on online syntactic processing in a second language. More specifically, we examined how German learners of Dutch deal with conflicting syntactic structures in L1 and L2, and to what extent their online processing signatures for such difficult instances of foreign grammar are already native-like. Learners' ERP signatures to sentences containing violations of infinitive verb order were compared to those of Dutch native speakers. Furthermore, to make sure the learner population was at all sensitive to L2 syntax, we tested their online sensitivity to violations of verb placement in subordinate clauses, which is implemented similarly in German and Dutch. In the remainder of this thesis, the results will be discussed with respect to earlier studies and theories on L2 syntactic processing.

4.1 Control Condition – Comparable syntactic structures in L1 and L2

For sentences with compatible syntactic structures in L1 and L2, learners showed a robust P600 effect comparable to that shown by the Dutch native speakers. In the L2 learners, the effect started slightly earlier than in the natives; however, a latency difference of 50 ms is unlikely to be meaningful. Overall, learners' ERP signatures to L1–L2 compatible syntactic structures can thus be described as native-like. This finding confirms our predictions and is, moreover, in line with earlier studies also reporting native-like ERP signatures for structures that exist in similar ways in both the L1 and L2 (Foucart & Frenck-Mestre, 2012; Lemhöfer et al., 2014).

4.2 Critical Condition – Conflicting syntactic structures in L1 and L2

Despite native-like behavioural performance, learners showed a strikingly different pattern compared to native speakers to sentences with conflicting structures in L1 and L2. While Dutch participants showed a robust and widely distributed P600 to violations of infinitive verb order, German learners of Dutch showed a biphasic N400–P600 pattern, preceded by an early left negativity. These results are in line with hypothesis (2) rather than (1). Note also that we did not find evidence for the extreme case of L1 influence, that is a reversed P600 (hypothesis (3)).

Biphasic patterns have been reported before in studies on L2 syntactic processing (Tanner et al., 2014, 2013), but as in earlier studies, our learners actually tended to show either an N400 or a P600, rather than both. A split by N400–P600 dominance made these differences very clear and shows that there were strong differences in how participants dealt with conflicting syntactic structures. Before turning to an interpretation of these differences in ERP signatures within learners, as well as between native and non-native speakers, let us discuss the ERP components observed for the non-native speakers separately.

4.2.1 Early left negativity

The early negativity, observed in L2 speakers over the left hemisphere in a window from 100–250 ms, is somewhat puzzling. To our knowledge, we are the first to report a negativity as early as 100ms post stimulus onset for L2 speakers. Earlier studies had reported slightly later starting (~200 ms) prolonged negativities in L2 speakers and interpreted them as N400 effects. An analysis of the entire time window of the present negativity (100–600 ms) as one single (N400-like) component in our study, however, seemed inappropriate due to the exceptionally early start of the deflection, and the topographic and temporal unfolding of the negativity.

Based only on its latency and scalp distribution, it may correspond to what has been called an ELAN in studies on syntactic processing in monolinguals (Friederici, 1995), which is almost exclusively associated with phrase structure violations (Steinhauer & Drury, 2012). Our sentences did not contain violations of this kind, and neither did they elicit any early left-lateralized effects in native speakers, thus rendering it unlikely that our early negativity is an ELAN.

Steinhauer and Drury (2012) offer a potential alternative explanation. In a critique of earlier studies, they argue that the presence of ELANs in a lot of the studies that report them may actually be artificially induced by baseline problems due to differences in the word preceding the target. Again, however, this explanation seems unlikely in the present context. Words preceding the targets in our study appeared equally often in both correct and incorrect sentences. Any effect of the preceding word should thus cancel out during averaging. Visual inspection of the baseline for the critical sentences further confirms this: the lines appear to be roughly on top of each other, suggesting that we are not dealing with a baseline problem here.

Unfortunately, we can currently not offer any other explanation of this early left-lateralized effect. It remains open for future studies to investigate what it represents and, more importantly, whether it is actually a reliable, meaningful effect.

4.2.2 Biphasic N400 – P600 pattern

As explained above, biphasic patterns are not entirely new in the L2 syntactic processing literature, and neither is the fact that on an individual level, this pattern tends to be monophasic (N400 or P600). Earlier studies had attributed such differences in ERP signature dominance to the participants' level of grammatical sensitivity (Tanner et al., 2013), or to their motivation to speak like a native-speaker and their age of arrival in the L2 environment (Tanner et al., 2014). Together these factors were taken as indicators for overall L2 proficiency: Tanner and colleagues concluded that an N400 occurs as a result of poor L2 learning, whereas a P600 was understood as a manifestation of more advanced levels of L2 proficiency (Tanner et al., 2014, 2013) (see Steinhauer, White, & Drury, 2009 for a similar argument).

Unfortunately, none of our behavioural scores, including measures of motivation to learn Dutch, age of arrival and grammatical sensitivity, could explain the dominance pattern in the present study. We can thus only speculate about the causes behind it. A role of L2, though plausible, is not clearly supported: none of the proficiency-related behavioural scores were significantly correlated with the dominance index or the N400 and P600 effect magnitudes, and participants from the N400 and the P600 dominant groups performed equally well on the grammaticality judgments during the EEG ($t(28) = -0.24$, $p = .814$) and the offline proficiency test ($t(28) = 0.55$, $p = .583$). Both in terms of general L2 proficiency and specific proficiency regarding our target structure the two groups did thus not differ from each other. The same holds for the participants' level of motivation to learn Dutch ($t(28) = 0.81$, $p = .427$) and their age of arrival in the Netherlands ($t(28) = -0.27$, $p = .786$).

Our proficiency test is not a standardized test, and it could be argued that it was simply not sensitive enough. However, what speaks against such an argument are the positive correlations between our proficiency test and a standardized vocabulary test (LexTALE), as well as participants' proficiency self-ratings. Altogether, it thus remains up to future studies to investigate the underlying causes of this ERP dominance pattern, perhaps with a larger set of participants or possibly a different set of behavioural measures (see below for specification). Nevertheless, we will now discuss the possible interpretations of the N400 and P600 components separately.

P600

The P600 in the L2 speakers in the critical condition was significant from 800 ms onwards and present over central and posterior electrode sites. This late positivity does not correspond to the canonical P600 effect in terms of latency. Nevertheless, there were at least two good reasons to interpret it as a (delayed) P600: First, its scalp distribution fits that of canonical P600 effects and was moreover found not to differ significantly from the native speakers' P600 in magnitude and topography in this late window. Second, the effect continued throughout most of the second verb and visually expanded into a typical P600 effect (see Suppl. Fig. 2). Moreover, similarly delayed P600 responses have been reported before (Hahne, 2001; Rossi et al., 2006).

The delay in P600 suggests that it took German participants longer to realize the mistake in word order and to initiate syntactic reanalysis / repair processes. Given that control sentences with *compatible* verb order rules in L1 and L2 elicited a native-like P600 in L2 speakers, it seems most likely that the delay in P600 onset in the critical sentences was caused by the conflict between German and Dutch syntactic rules. The fact that German grammar allows the incorrect Dutch order may have led to a delayed realization that what the participants were reading was incorrect.

A question that arises with respect to the delay of the P600 is whether the effect was driven by the first or the second infinitive. Our experiment was not designed to give a definitive answer to this question. However, the fact that the P600 visually starts at the very moment the second verb first appears on screen (i.e., when it is unlikely to have been processed lexically), speaks against the interpretation that the effect was driven by the second verb only. We would like to argue that it is more likely that the syntactic reanalysis (as indexed by the P600) starts already on the first verb, just later than usual, and is then reinforced by the second verb; however, not initiated by it.

N400-like effect

The N400-like effect was significant in a window from 250–600 ms and was largest over left central and posterior electrode sites. In terms of latency, this effect corresponds to the canonical N400 response; its left lateralization, however, is unusual since N400 effects are normally either distributed bilaterally, or are larger over the right hemisphere (Kutas & Federmeier, 2011). Earlier studies on L2 syntactic processing reported similar, left-lateralized N400 effects though, both in a biphasic pattern with the P600 (Tanner et al., 2014, 2013; Weber & Lavric, 2008) and alone (McLaughlin et al., 2010), thus reinforcing our interpretation of this negativity as an N400.

An N400 to syntactic violations in the L2 can and has been explained in a number of different ways (for a review, see Steinhauer et al., 2009). In fact, it may represent distinct aspects of processing in different studies, depending on the task and the structure under investigation. Given that none of our behavioural scores correlated with either the dominance index or the N400 magnitude itself, we are left with previous studies' interpretations, which I will discuss here briefly one by one.

As noted above, one line of research has interpreted the N400 in L2 syntactic processing as an index of low L2 proficiency. Related to this explanation, McLaughlin et al. (2010) and Steinhauer et al. (2009) claim that learners progress through qualitatively distinct stages in L2 acquisition. According to them, the early stages of L2 learning are characterized by greater involvement of the lexico-semantic system and by a strategy to memorize salient word sequences. They suggest that violations of (morpho-)syntax are not yet recognized as such in low proficient learners: morphologically complex words, for example, are thought to not be decomposed, but rather processed as entire words or word combinations, for which violations result in novel, and thus surprising words or word combinations that prompt an N400 rather than a P600. With increasing levels of proficiency, McLaughlin and Steinhauer argue that learners induce and learn syntactic rules that trigger re-analysis processes when violated and therefore lead to a P600 (McLaughlin et al., 2010; Steinhauer et al., 2009).

While this explanation seems viable in studies with violations of morphosyntax, such as incorrect subject–verb agreement, it cannot account for violations of long distance dependencies as in the present study. In order to recognize the error in placement of the verb 'maken' in a sentence like 'Ik heb het huiswerk niet maken willen', learners have to take into account the entire sentence from 'heb' onwards, that is a sequence spanning *five* words (or more). Shorter sequences from our sentences do not necessarily result in novel word combinations, because a tri-gram like 'huiswerk niet maken' could occur in a sentence like 'Ik wil het huiswerk niet maken' where they are grammatically correct. In order for 'maken' to trigger a surprisal effect, readers would thus need to have internalized representations of exceedingly long word strings. This is highly unlikely. Moreover, if the N400 in some of our participants was really due to a lexically driven processing strategy, and if learners in the early stages of acquisition are not able to generalize and to abstract away from exemplars, as McLaughlin and colleagues propose, learners would need to have heard exactly the combination of words that we are using in our sentences before, in order to recognize the error — an assumption that makes this interpretation of the N400 to our critical sentences all the more unlikely.

Yet another possibility along similar lines is that learners expected sentences like "Ik heb het huiswerk niet" to end in a verb in participle form ("gemaakt"), which would be an alternative grammatical continuation of the sentence in Dutch. The appearance of an infinitive instead of a participle may then have been "surprising". Again, however, in order for learners to expect a participle in that position based purely on a lexical mechanism, they would need to have stored exceedingly long word strings. Moreover, the N400-like effect should then theoretically be equally present in correct and incorrect sentences: both contained infinitives instead of a participle, as well as in native speakers, which was clearly not the case.

What may seem more plausible is that learners store the infinitive sequence ('willen maken') from their daily linguistic input and detect an anomaly when hearing it in the opposite order ('maken willen'). If this were true, however, one would expect the N400 to show only on the second verb, that is after they have seen both infinitives, rather than already on the first verb, which was not the case either (see Suppl. Fig. 2). All these considerations combined thus render an explanation of the N400 in terms of lexico-semantic processing or storage and subsequent surprisal unconvincing for the present findings.

Other accounts of N400 effects to (morpho-) syntactic violations in L2 have included sentence wrap-up and semantic integration difficulty in sentence-final position (Weber & Lavric, 2008). If participants expected the participle "gemaakt" after "Ik heb het huiswerk niet", they may have initiated sentence wrap-up. However, similar to the earlier critique, such sentence wrap-up processes should be equally large for correct and incorrect sentences and would also be expected in native speakers. Moreover, it seems unlikely that after repeated exposure to sentences ending in double infinitives, participants would still expect participles in sentence-final position.

A final interpretation, which to our knowledge has not been brought forth within the context of the "syntactic N400", links the negativity in our study to processes of conflict detection and resolution. Studies using the Stroop or Flanker tasks, which require "the detection of processing conflicts between incompatible competing alternative responses" (Perlstein, Larson, Dotson, & Kelly, 2006, 262), reliably report negative deflections in the ERP signal, which are in some respects comparable to the negativity in the present study (Liotti, Woldorff, Perez, & Mayberg, 2000; West, 2003). Flanker tasks frequently report an N2 and Stroop tasks report an N450 (Larson, Clayson, & Clawson, 2014). Most relevant for our study is the N450, which occurs roughly between 300–500 ms over centro-parietal scalp sites and which has been linked to activity in the anterior cingulate cortex (ACC), a brain region known to be involved in conflict monitoring. Researchers generally agree that the N450 reflects the identification of conflict between two sources of information and subsequent interference resolution (Szűcs & Soltész, 2012; West, 2003).

While the negativity in our N400 group is visually longer and more widely distributed than the typical N450, it could be interpreted as a marker of L1–L2 conflict, that is a conflict between two languages ('two sources of information') regarding the continuation of a sentence (German vs. Dutch verb order). Remember that our critical sentences contain structures that pose direct syntactic rule conflicts in German and Dutch, which need to be overcome and resolved in order to correctly judge a given sentence for grammatical correctness. The set-up of the current study, moreover, encouraged interference from the L1 (instructions prior to the experiment were in German and sentences consisted of a lot of cognates), making this possibility intuitively plausible. Some German participants may have been less successful in inhibiting German and as a consequence of that may have experienced competition between German and Dutch verb order rules, resulting in an N450-like signature. Whether the language conflict occurred specifically in relation to the obligatory grammaticality judgments or more generally in response to a competition between what type of verb to expect remains open.

In order to test the hypothesis of this negativity as a marker for conflict, it would be useful to conduct the present EEG study as a functional magnetic resonance imaging (fMRI) study. In doing so, one could identify the locus of the negativity, which, if it represents conflict monitoring, would be expected to be in the anterior cingulate cortex (and possibly also in other areas associated with cognitive control such as the inferior prefrontal cortex) (Larson et al., 2014; Van Heuven, Schriefers, Dijkstra, & Hagoort, 2008). Future studies should moreover investigate whether the conflict arises specifically in relation to the grammaticality judgments or not. Likewise, the task remains for future studies to identify why only a subset of participants experienced this potential conflict. While it does not seem unlikely that some participants are on some level more advanced and better at inhibiting their L1 than others, behavioural tests sensitive to cognitive control ability, such as the Stroop task, are necessary for future investigations in this matter.

Out of all above-mentioned accounts, explaining the N400-like negativity as a marker of L1–L2 conflict seems most plausible. Note, however, that this interpretation is of a purely speculative and tentative nature, and needs to be backed up empirically.

4.3 Implications and future directions

Regardless of the specific interpretation of the N400 or the delay in P600, our results clearly demonstrate that learners were experiencing difficulties in reading sentences containing conflicting syntactic structures in L1 and L2, in line with hypothesis (2). For structures implemented similarly in German and Dutch, instead, this was not the case; here, learners showed native-like ERP signatures.

Our results thus stress that the L1 has an important impact on the success of L2 syntactic processing. We illustrate this influence both in its positive (control sentences) and its negative form (critical sentences). In doing so, we reinforce earlier studies that show that structures, which are implemented similarly, can elicit native-like ERP signatures, even in late L2 learners (Lemhöfer et al., 2014; McLaughlin et al., 2010; Sabourin & Stowe, 2008) and that differently implemented structures are difficult to process, even at advanced L2 stages, and even if the structure has been mastered at a near native-like level behaviourally (Chen et al., 2007; Ojima et al., 2005).

More importantly though, our study adds insights into the processing of structures for which rules in L1 contradict those of the L2. Our results suggest a complex pattern: while some participants were, interestingly, only marginally affected by the syntactic conflict (delayed P600), others seemed to have experienced quite substantial interference from German (resulting in a negativity). Whether the effects we find here are really due to interference remains to be determined by future studies. For now, an interpretation of the results in terms of L1–L2 dissimilarity and consequently L1 interference, however, seems most plausible.

Apart from the necessary follow-ups described above, it would be interesting to test learners at earlier points in the acquisition process, that means right at the beginning of their stay in the Netherlands, as well as at later stages. Such an extension would not only further our understanding of the developmental nature of the pattern, but would also be likely to increase the chances of finding trends in the data pattern based on individual differences. Moreover, including participants at the very beginning of the acquisition process may also show the hypothesized reversed ERP signature, since learners are more likely to make mistakes at the very beginning of the learning process, and are arguably most likely to be severely hindered by their L1 at this stage. Finally, including more advanced learners will be insightful as to the question of whether or not processing signatures to conflicting structures can become native-like at all, or whether some residual influence of the L1 persists even at very advanced stages of L2 acquisition.

To sum up, the present study adds new insights to the growing body of research on L1–L2 similarity effects and, more specifically, to the processing of conflicting syntactic rules in L1 and L2. We have shown that structures with conflicting syntactic rules in two languages lead to non-native like processing signatures in learners. The extent to which individuals are influenced by this conflict, however, varies; the reasons for this remain to be investigated.

5. References

Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, *51*, 2492–2511.

Caffarra, S., Molinaro, N., Davidson, D., & Carreiras, M. (2015). Second language syntactic processing revealed through event-related potentials: An empirical review. *Neuroscience & Biobehavioural Reviews*, *51*, 31–47.

Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition*, *10*(2), 161–174.

Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*(12), 564–570.

Davidson, D. J., & Indefrey, P. (2009). Plasticity of grammatical recursion in German learners of Dutch. *Language and Cognitive Processes*, *24*(9), 1335–1369.

Erdocia, K., Zawiszewski, A., & Laka, I. (2014). Word order processing in a second language: From VO to OV. *Journal of Psycholinguistic Research*, *43*, 815–837.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*(4), 527–552.

Foucart, A., & Frenck-Mestre, C. (2012). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition*, *15*(1), 379–399.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Herrida & J. Altarriba (Eds.), *Bilingual Sentence Processing* (pp. 217–236). Amsterdam: Elsevier.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*(3), 259–281.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*, 183–192.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*(3), 251–266.

Hoff, V. van 't. (2012). *The influences of L1 on the comprehension and production of idioms in L2*. Radboud University Nijmegen.

Hulk, A. (1991). Parameter setting and the acquisition of word order in L2 French. *Second Language Research*, *7*(1), 1–34.

Jansen, B., Lalleman, J., & Muysken, P. (1981). The alternation hypothesis: Acquisition of Dutch word order by Turkish and Morrocan foreign workers. *Language Learning*, *31*(2), 315–336.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. *Behaviour Research Methods*, *42*(3), 643–650.

Klaus, J., Oppermann, F., & Lemhöfer, K. (n.d.). A parallel version of a reading span test for Dutch, German, and English speakers.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*, 68–74.

Kutas, M., & Federmeier, K. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Larson, M. J., Clayson, P. E., & Clawson, A. (2014). Making sense of all the conflict: A theoretical review and critique of conflict-related ERPs. *International Journal of Psychophysiology*, *93*(3), 283–297.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behaviour Research Methods*, *44*, 325–343.

Lemhöfer, K., Schriefers, H., & Hanique, I. (2010). Native language effects in learning second-language grammatical gender : A training study. *Acta Psychologica*, *135*(2), 150–158.

Lemhöfer, K., Schriefers, H., & Indefrey, P. (2014). Idiosyncratic Grammars: Syntactic Processing in Second Language Comprehension Uses Subjective Feature Representations. *Journal of Cognitive Neuroscience*, *26*(7), 1428–1444.

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, *38*(5), 701–711.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (2nd ed.). New York, NY: Cambridge University Press.

MacWhinney, B. (2005). A Unified Model of Language Acquisition. In J. F. Kroll & A. M. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 49–67). New York: Oxford University Press.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–704.

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain Potentials Reveal Discrete Stages of L2 Grammatical Learning. *Language Learning*, *60*(Suppl. 2), 123–150.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, England: Cambridge University Press.

Oines, L., Miyake, A., & Kim, A. (2012). Individual differences in verbal working memory predict reanalysis vs. integration in syntax-semantics conflicts. In *Architectures and Mechanisms for Language Processing Conference, Riva del Garda, Italy*.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience*, *17*(8), 1212–1228.

Osterhout, L., & Holcomb, P. J. (1993). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Perlstein, W. M., Larson, M. J., Dotson, V. M., & Kelly, K. G. (2006). Temporal dissociation of components of cognitive control dysfunction in severe TBI: ERPs and the cued-Stroop task. *Neuropsychologia*, *44*(2), 260–274.

Pozzan, L., & Quirk, E. (2014). Second language acquisition of English questions: An elicited production study. *Applied Psycholinguistics*, *35*, 1055–1086.

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*(12), 2030–2048.

Sabourin, L., & Stowe, L. (2008). Second language processing: When are first and second languages processed similarly? *Second Language Research*, *24*(3), 397–430.

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162.

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, *25*(1), 13–41.

Szűcs, D., & Soltész, F. (2012). Functional definition of the N450 event-related brain potential marker of conflict processing: A numerical Stroop study. *BMC Neuroscience*, *13*(35), 1–14.

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*(2), 277–293.

Tanner, D., Mclaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*(2), 367–382.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition*, *27*(2), 173–204.

Tuninetti, A., Warren, T., & Tokowicz, N. (2015). Cue strength in second-language processing: An eye-tracking study. *The Quarterly Journal of Experimental Psychology*, *68*(3), 568–584.

Van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex*, *18*(11), 2706–2716.

Weber, K., & Lavric, A. (2008). Syntactic anomaly elicits a lexico-semantic (N400) ERP effect in the second language but not the first. *Psychophysiology*, *45*(6), 920–925.

West, R. (2003). Neural correlates of cognitive control and conflict detection in the Stroop and digit-location tasks. *Neuropsychologia*, *41*(8), 1122–1135.

White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain responses before and after intensive second language learning: Proficiency based changes and first language background effects in adult learners. *PloS One*, *7*(12), e52318.

BOLD Activation in the Oculomotor Network After Perturbation of the Parietal Eye Fields: Implications for Parkinson's Disease

Femke Struik[1]
Supervisors: Ian Cameron[1], Andreea Cretu[2], Ivan Toni[1], Bas Bloem[1,3]

[1]Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2]ETH Zürich, Department of Health Science, Zürich, Switzerland
[3]Radboud University Medical Centre Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands

Parkinson's disease is characterized by bradykinsesia and akinesia, caused by degeneration of the dopaminergic neurons in the substantia nigra, causing problems in voluntary movement. To overcome these problems patients with Parkinson's disease can develop compensational strategies in which they are helped by visual information to initiate movements. However, these compensatory mechanisms are not always beneficial as strong visual cues can trigger freezing. In previous eye movement experiments patients with Parkinson's disease are shown to be more reflexive at pro-saccades, but they have problems in inhibiting these reflexes in anti-saccades. Neuroimaging evidence revealed reduced activity in movement programming regions (frontal eye fields), but increased activity in parietal/occipital visual regions. The aim of the current project is to establish whether this parietal hyperactivity is compensational or pathological. To this end we disrupted activity in the right parietal eye fields (PEF). Two Parkinson's patients and seven healthy controls were tested in a pro- and anti-saccade task in three sessions (baseline, after continuous theta burst stimulation (cTBS) to PEF and after cTBS to a control region, S1). The BOLD activity in left and right parietal eye fields, frontal eye fields (FEF) and dorsolateral prefrontal cortex after cTBS to PEF was compared with cTBS to S1. In healthy controls, we found increased activity in right FEF after cTBS to right PEF, suggesting compensation of right FEF. These results are discussed in relation to previous transcranial magnetic stimulation (TMS) studies, compensation and Parkinson's disease.

Keywords: Parkinson's disease, continuous thetaburst stimulation (cTBS), transcranial magnetic stimulation (TMS), functional magnetic resonance imaging (fMRI), anti-saccades, compensatory mechanisms, eyetracking, parietal eye fields (PEF), frontal eye fields (FEF), dorsolateral prefrontal cortex (DLPFC)

Corresponding author: Femke Struik; **E-mail:** fmstruik@gmail.com

1. Introduction

Parkinson's disease is a movement disorder characterized by bradykinesia (slow movement execution) and akinesia (altered action planning) (Helmich, 2011). These symptoms are thought to be caused by dopamine depletion due to degeneration of the dopaminergic neurons in the substantia nigra pars compacta, part of the basal ganglia. In terms of anatomy, the basal ganglia do not directly project to the spinal cord, but they are interconnected with the cerebral cortex through loops involving the striatum (caudate nucleus, putamen and nucleus accumbens), the pallidum, substantia nigra and specific nuclei in the thalamus. The dopamine depletion in the substantia nigra impacts neural circuitry involving the basal ganglia and frontal cortex, which causes problems with voluntary movement (Albin, Young, & Penny, 1989).

To overcome these problems with voluntary movement, patients with Parkinson's disease can develop compensatory strategies. An example of such a strategy is that they are helped by visual information to initiate movements. For example, horizontal lines on the floor or focusing on an object in the distance helps them to start walking along a hallway. In the absence of this visual information, they can show freezing of gait, demonstrating their deficits in voluntary movement. However, strong visual cues, like crowded places, can also trigger freezing (Bloem, Hausdorff, Visser, & Giladi, 2004). This indicates that a compensatory strategy of using visual cues to overcome voluntary movement problems may not always be beneficial and may even contribute to the problems in voluntary control. It is the aim of this thesis to shed light on these potential compensatory mechanisms.

Apart from purely visually-based compensatory strategies, other compensatory mechanisms in Parkinson's disease were found in laboratory experiments. For example, neuronal compensation in terms of BOLD functional magnetic resonance imaging (fMRI) was found in the somatosensory domain. Helmich, Bloem, and Toni (2012) found that Parkinson's disease patients with resting tremor have increased somatosensory activity (as shown during motor imagery). The authors suggest this may explain why this group of patients has a slower progression of motor symptoms, fewer problems with motor planning and less cognitive dysfunction, demonstrating compensation. This compensatory mechanism was also found in another study by Helmich, de Lange, Bloem, and Toni (2007). Also in a motor imagery task, patients were found to have increased activity in the right extrastriate body area (EBA) and occipito-parietal cortex (OPC) when they rotated their most affected hand as compared with their least affected hand. The authors also found enhanced coupling of the EBA and OPC with the left dorsal premotor cortex (PMd). This could point to a compensatory mechanism between the motor cortex and the somatosensory network with patients making more use of visual areas like EBA and OPC. Similar effects were found in a study by van Nuenen et al. (2012): after perturbation of the right EBA, Parkinson's patients, rather than controls, lost the ability to incorporate the current state of their hand with a mental hand rotation (the "posture congruency effect"). On the contrary, after perturbation of the left PMd, this effect was reduced in controls, rather than in Parkinson's patients. These results indicate the right EBA compensates for problems in PMd in Parkinson's disease and could point to visual compensation for problems in premotor cortex function. While these experiments used motor imagery, they found compensation in similar brain regions that would be active across visual tasks.

Apart from the use of motor imagery to give us insight in motor control and possible compensatory mechanisms, saccades are an interesting example of movements since they involve pre-programmed movements and thus can be a window into the central mechanisms of motor control. There are also well documented deficits in saccade programming in Parkinson's disease that reveal problems in movement initiation and cognitive control.

Saccades are very rapid eye movements used to move the fovea from one target to another. Voluntary saccades are initiated in the frontal eye fields (FEF). In terms of anatomy, FEF projects directly to the areas in the pons that contain the command neurons for the generation of saccades. There are also indirect projections to these pontine areas, namely through a complex circuit with synapses in the caudate nucleus and rostral substantia nigra of the basal ganglia, and superior colliculus. The superior colliculus then projects to the pontine neurons involved in saccade generation (Albin et al., 1989). In terms of pathology, it is thought that in Parkinson's disease, the neurons in the superior colliculus involved in the control of saccades have increased inhibition, due to increased inhibitory output from the substantia nigra pars reticulate. This perhaps over-simplistic model nevertheless accounts for problems Parkinson's disease patients have in saccade generation.

The problems with saccade generation in

Parkinson's disease can give us insight into problems with movement and inhibitory control. To this end, saccades have been studied using a pro- and anti-saccade task (Munoz & Everling, 2004). In this task, participants either make a saccade towards a peripheral stimulus (pro-saccade), which is interpreted as an automatic, reflexive response, or they make a saccade away from the stimulus (anti-saccade), which is a voluntary response. In order to make an anti-saccade, the brain has to incorporate a task set (a rule on how to behave); it has to prepare to execute a voluntary eye movement based on a coloured cue (Cameron et al., 2012), hence, problems of patients in anti-saccade trials may reflect diminished voluntary control. It was shown that Parkinson's patients, as compared with healthy age-matched controls, have shorter reaction times at pro-saccades and are more accurate to go in the correct, cued, direction (Chan, Armstrong, Pari, Riopelle, & Munoz, 2005, Hood et al., 2007). However, in anti-saccades, these patients show longer reaction times and are less accurate to go to the correct direction (they often execute a pro-saccade). Additionally, they display disrupted 'metrics': their saccades land short (hypometria). These impairments were shown by Amador, Hood, Schiess, Izor, and Sereno (2006), who tested Parkinson's patients in an anti-saccade task, delayed anti-saccade task and remembered anti-saccade task. They found these patients had problems in voluntary task execution (their response times were longer than controls) and they also showed problems with the inhibition of reflexive saccades. Evidence from neuroimaging of problems of Parkinson's disease patients in anti-saccades comes from Cameron et al. (2012). In an fMRI experiment using the same anti-saccade task, they found that patients with Parkinson's disease have problems in establishing voluntary task sets to control voluntary motor behaviour, rather than problems in the execution of these saccades. In their experiment, healthy participants showed increased activation for anti-saccade preparation (as compared with pro-saccade preparation), but Parkinson's patients (off medication) did not show this effect. Additionally, Parkinson's patients had a reduced enhancement of an early rise in preparatory activation in FEF on correct anti-saccade trials, as compared with incorrect anti-saccade trials. This may point to problems in generating the task set. Because of the basal ganglia dysfunctions in Parkinson's disease, the authors suggest the basal ganglia may be important in translating executive signals (rule representation) into signals that drive voluntary motor behaviour. Importantly, they also found evidence for increased activity in parietal/occipital visual regions. It is unclear whether this activity reflected compensatory or pathological activity, and a goal of this study is to test this causally with transcranial magnetic stimulation (TMS).

As mentioned above, pro-saccades are considered to be reflexive, whereas anti-saccades involve inhibiting this reflex and voluntarily making a saccade in the opposite direction. This voluntary generation of a saccade involves many cortical structures that are affected by basal ganglia pathology: the superior colliculus (SC), the supplementary eye fields (SEF), the frontal eye fields (FEF), the parietal eye fields (PEF) and the dorsolateral prefrontal cortex (DLPFC). While these regions have different roles in reflexive and voluntary saccades, there is some overlap of functions that would be relevant in terms of potential compensation following neurodegeneration. Munoz and Everling (2004) stress the importance of the SC in the saccade network since it receives input from all cortical and subcortical structures that are involved in saccade control. However, if these incoming signals are incorrect based on the relevant task, erroneous or improperly programmed saccades result. Both SC and FEF contain 'fixation' and 'saccade' neurons that are tonically active, which means they discharge in high-frequency bursts of action potentials during, respectively, fixation and saccades. In monkeys, when a stimulus is shown to instruct an anti-saccade, activity in fixation neurons is enhanced, as compared with a pro-saccade trial. It is important to note that there is a laterality effect: a visual stimulus in the right visual field activates the saccade neurons in FEF and SC in the contralateral hemisphere (left) and inhibits the saccade neurons in the ipsilateral hemisphere (right). Munoz and Everling (2004) argue that in order to make a correct anti-saccade, neurons in the SC and FEF should be inhibited in a top-down manner before the target appears, inhibiting the execution of a pro-saccade. Possibly, the SEF, the DLPFC or the basal ganglia are involved in sending inhibiting signals to SC and FEF. Additionally, in order to make an accurate anti-saccade, the visual target-vector must be inverted into a saccadic vector. In monkeys, area LIP (lateral intraparietal area, analogous to human PEF) and FEF are involved in vector inversion. It is shown in humans that when FEF is perturbed with continuous theta burst stimulation (cTBS, a form of TMS that is thought to have an inhibitory effect), this affects visual processing, rather than saccade generation or executive control (Cameron, Riddle, & D'Esposito, 2015). This demonstrates a critical visual processing role of FEF.

Whereas the FEF may be more involved in disengaging fixation and saccade generation, PEF may have more of a role in triggering saccades that are made reflexively to the sudden appearance of visual targets (Pierrot-Deseilligny, Rivaud, Gaymard, Müri, & Vermersch, 1995). However, this does not mean that PEF is not involved in anti-saccades. According to Medendorp, Goltz, and Villis (2005), PEF remaps its activity to the contralateral visual field when an anti-saccade has to be made, coding the target location for a saccade. Leoné, Toni, and Medendorp (2014) found that both PEF and FEF code direction and amplitude components of a saccade. Therefore, it may be the case that if one region is impaired, some functions could be spared by the similar role of the other region. In this study, we focus on the role of PEF in the programming of saccades. PEF may have a primary role in the saccade amplitude, but it may also influence the hyper-reflexivity observed in Parkinson's disease patients, as it triggers reflexive saccades. This hyper-reflexivity would work against the executive control exerted by DLPFC and FEF. To get a better view on the causal role of PEF, we use inhibitory transcranial magnetic stimulation to perturb activity in this area.

Transcranial magnetic stimulation (TMS) is a non-invasive brain stimulation technique. A large electrical current in a copper coil produces a magnetic field orthogonal to the direction of the wire. When holding this coil over a scalp, the electrical current induces a magnetic field, which in turn induces an electric field in the underlying tissue. The magnitude of this field can be big enough to cause depolarization in the underlying neurons (Rossi, Hallett, Rossini, & Pascual-Leone, 2009). Applying TMS pulses in a specific pattern can increase or decrease cortical excitability. This effect can even last longer than the stimulation itself; so-called offline TMS effects. In cTBS trains of three pulses are applied at 50 Hz intra-train and 5 Hz inter-train frequency continuously. This protocol is thought to cause long-term depression in the underlying tissue (Huang, Edwards, Rounis, Bhatia, & Rothwell, 2005). This long-term depression results in disruptions or perturbations in function of the brain area the cTBS was directed to. By using cTBS we can make conclusions about the causal function of this brain area.

In the current project, cTBS is used to get more insight in the functioning of the oculomotor network in Parkinson's disease patients and healthy participants. More specifically, we aim to establish the neural mechanisms of the parietal hyperactivity and frontal hypo-activity that were found in Parkinson's patients when they made pro- and anti-saccades. In this experiment, we disrupted the activity in PEF using cTBS to get an idea on the causal relationships between these parietal areas and other oculomotor areas. We hypothesize the parietal mechanisms to be compensatory as they seem to help patients in movement. Hence perturbing these parietal mechanisms would cause even greater direction errors in anti-saccades. However, there is also a possibility that these mechanisms are pathological as they can trigger freezing. In this case, saccades would improve after perturbing these parietal areas.

To disrupt these potentially compensational mechanisms, Parkinson's patients (early-mid stage) as well as controls in the same age range received cTBS on parietal/occipital regions; the right parietal eye fields located in the right intraparietal sulcus (this will be referred to as cTBS to PEF) or at a control site (primary somatosensory cortex, leg region, this will be referred to as cTBS to S1). cTBS to S1 was used as a control site because it is considered not to be involved in the oculomotor system, or more specifically in the generation of pro- and anti-saccades. As mentioned above compensatory activity in Parkinson's disease was also found in the somatosensory cortex, however compensatory activity is considered task-dependent leaving the primary sensory cortex a suitable control region for the current experiment.

Participants performed the pro- and anti-saccade task (similar to Cameron et al. 2012) to establish two things. First, if increased parietal/occipital activation found in other studies indicates a compensatory process, this compensation may increase the relative role of the parietal cortex. In this way there could be increased gain in the sensory system which causes improved saccade accuracy in patients. Thus, we hypothesized patients would have more problems with anti-saccade metrics after PEF cTBS (the compensatory process is now impaired); anti-saccades were expected to land even shorter as compared with healthy controls. Second of all, strong visual signals as the result of this compensatory increase in parietal activity could cause patients to produce more errors in anti-saccades. By reducing these signals by cTBS, patients would make fewer pro-saccade errors on anti-saccade trials. However, as mentioned previously, we also considered that parietal/occipital activity is pathological ("maladaptive" as the result of changes in basal ganglia circuits). In this case, the opposite predictions would result after cTBS if pathological activity was reduced.

On a neuronal level, we expected alterations in activity and connectivity in the oculomotor network regions in the right hemisphere (site of cTBS) after cTBS to PEF as compared with the control cTBS condition. We hypothesized that because patients with Parkinson's have a state of compensation, as a result of their disease, increased activity relative to controls would be present in PEF. Therefore, because of this compensatory activity, after cTBS is applied to the PEF, PEF activity in patients would not be reduced as much as in healthy controls. In contrast, patients should have hypoactivity in FEF and DLPFC (as shown previously by Cameron et al., 2012), and due to functional connectivity with PEF, FEF and DLPFC activity might be further reduced after cTBS. Thus, by combining cTBS with fMRI, we can gain insight in the effect of cTBS to PEF on activity in FEF and DLPFC.

However, since the testing of Parkinson's patients takes longer than the span of this thesis, we will mainly focus on the effects of PEF cTBS in healthy control participants. These findings can give us a reference frame on how to interpret the data from patients. Since FEF has a similar role in the vector inversion processes as PEF and vector inversion is necessary for programming saccades to the accurate spatial locations, we expect FEF can conversely compensate for problems in PEF in healthy controls where FEF and DLPFC circuits are not impaired. Thus we expect activity in FEF to increase after cTBS to PEF as compared with cTBS to S1. Since in healthy control participants, hyper-activity of PEF would not work against the inhibition of reflexive saccades by DLPFC, we expect no changes in BOLD activity in DLPFC. In summary, changes in BOLD activity as a result of right PEF cTBS in the cortical oculomotor regions of interest (ROIs) in healthy controls will serve as a reference to the changes in this network in Parkinson's patients.

2. Materials and methods

This study was approved by the local ethics committee (Committee on Research Involving Human Subjects, CMO, region Arnhem-Nijmegen, the Netherlands).

2.1 Participants

Three Parkinson's patients (of which two were female) were recruited from the Neurology department of Radboud University Medical Centre Nijmegen and nine healthy controls (of which two were female) were recruited from the community of Nijmegen. All participants were 50–75 years old (*M* = 64 years). All control participants were right-handed; two of the patients were left-handed. All participants had normal or corrected-to-normal vision, were not colour blind and had no history or current presence of neurological or psychiatric diseases (other than Parkinson's disease for the patients).

The Parkinson's patients were all diagnosed with idiopathic Parkinson's disease according to the UK Brain Bank criteria, not tremor-dominant, with unilateral or bilateral involvement, mild to moderate disease severity (Hoehn and Yahr stage 1−3), with dopaminergic therapy with clear clinical response of non-tremor symptoms (bradykinesia and rigidity) and were tested off dopaminergic medication (last intake of medication > 12 hours prior to testing). They completed the Beck Depression Inventory (BDI), the Mini-Mental State Examination (MMSE) and Frontal Assessment Battery (FAB) to rule out depression and cognitive deficits. Patients that scored lower than 15 on the BDI, higher than 26 out of 30 on the MMSE, and higher than 13 on FAB were eligible for the study.

All participants gave informed consent during the first session and filled in the MRI and TMS screening forms at the beginning of each session. Participants joined voluntarily and received monetary compensation. One male control participant was excluded from the analysis because of technical errors. Additionally, one male control participant was excluded for not being able to participate in all three sessions. One male patient was excluded after the intake session because of TMS exclusion criteria. Hence the final sample consisted of seven healthy controls and two patients.

2.2 Experimental procedure

This study consisted of three sessions. Apart from that, patients also came in for a separate intake session to determine their eligibility. During the intake session, patients were explained the goals and set up of the study. They gave informed consent and completed the BDI, MMSE and FAB. The UPDRS was done as a baseline for their Parkinson's symptoms on medication. Additionally, we tried if they could lie down in the MRI comfortably (by use of a dummy scanner, i.e., an MRI scanner without static magnetic field, MR gradients or RF radiation).

The aim of session 1 was: first, to determine the active motor threshold for cTBS intensity in session 2 and 3 and, second, to make functional scans to determine saccade ROIs for cTBS in session 2 and

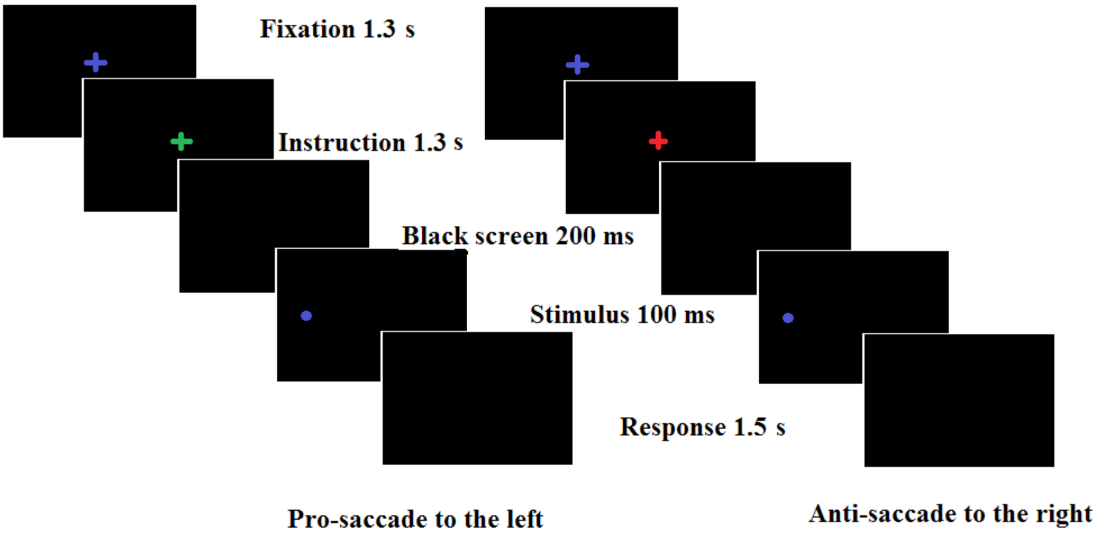

Fig. 1. Pro- and anti-saccade task.

3 and for analysis of all three sessions. In session 1, control participants were explained the goals, set up for the study and gave informed consent, because they did not come in for an intake session. All participants, patients as well as controls, filled in screening forms for the MRI and TMS at the start of all sessions. Participants were explained the pro- and anti-saccade task and they practiced one run of this task. After this, participants were placed in the MRI scanner where they performed the pro- and anti-saccade task. Before or after this, their motor threshold (explained below in 2.5, 'TMS') was determined for the cTBS in session 2 and 3. Due to limited availability of the TMS lab this part was postponed to session 2 for two of the participants.

Session 2 and 3 had the same course. All participants practiced one run of the pro- and anti-saccade task. After this they received 40 seconds of cTBS to S1 or PEF (the order was counterbalanced across participants) on 80% active motor threshold intensity. After this, participants were placed in the MRI scanner and performed the pro- and anti-saccade task. The three sessions were planned at approximately the same time of the day and were one week apart for six of the participants.

2.3 Pro- and anti-saccade task

Participants performed a pro- and anti-saccade task in the scanner (Fig. 1). The task was presented using Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) running in MATLAB version 2010 (MathWorks, Natick, USA) on a PC (Skyra, see below) or Mac (Prisma, see below). In this task, participants fixated on a blue cross projected on a black screen for 1.3 s. In both pro- and anti-trials the cross changed colour for 1.3 s. After this, the screen turned black for 200 ms and then a blue dot appeared at 12.6° (wide) or 8.5° (short) for 100 ms to the right or to the left (this was randomized across trials). Participants were instructed to look towards the dot after the cross had turned green and away from the dot (approximately the same distance to the opposite side) when the cross had turned red. They were instructed to stay at this position until a central stimulus was presented. They had 1.5 s to respond.

Each run consisted of 28 pro-saccade trials and 28 anti-saccade trials. To prevent anticipation and to make sure we would not catch the same part of the BOLD response during scanning, trials were jittered; fixations could take 1.5 s (12 trials), 3 s (eight trials) or 4.5 s (six trials), thus, fMRI data was collected in a rapid event-related design fashion. During random presentation of the trials, the same type of trials could occur after each other. To circumvent this, 500 possible orders were generated. Out of these, the 25 most efficient orders (based on Henson, 2006) were randomly used in the pro- and anti-saccade task. Participants never had the same order in one session. Participants performed three to five runs per session (depending the duration of the cTBS effect in session 2 and 3).

2.4 fMRI

All scanning was done at the Donders Centre for Cognitive Neuroimaging (Donders Institute,

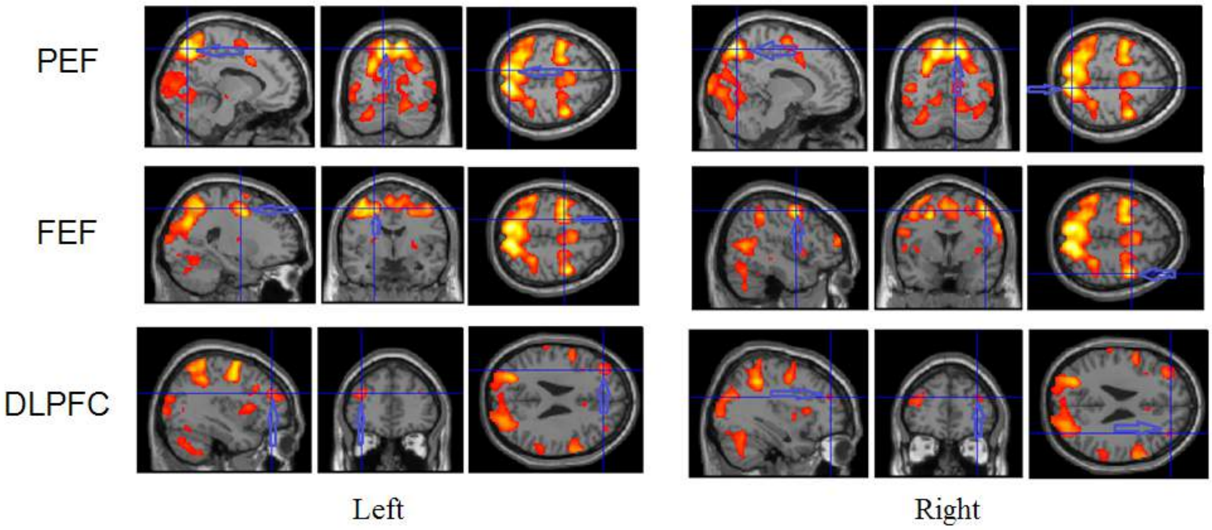

Fig. 2. Example of ROIs in the left and right hemisphere used for analysis.

Nijmegen, the Netherlands). For five participants (of which one patient), a 3T Siemens Skyra MRI system (Siemens Healthcare, Erlangen, Germany) was used to acquire T1 as well as T2* weighted images. Because of technical problems with eye tracking, a 3T Siemens Prisma MRI system (Siemens Healthcare, Erlangen, Germany) was used for the other four participants (of which one was patient). Importantly, all participants completed the three sessions in the same scanner. In both systems, a Siemens Auto-Align scout was used to present the location of imaging volume on a 3-plane localizer collected initially.

On both MRI systems the multiband sequence (Centre for Magnetic Resonance Research, Minnesota, US) was used with a multiband acceleration factor of 3, interleaved multi-slice mode. The TR was 1000 ms, TE 30.00 ms, 33 slices, 341 volumes, ascending slice order, 3.0 mm slice thickness, slice gap 0.51 mm, matrix size 64 x 64, field of view (FOV) 224 x 224 mm, flip angle 60°, fat suppression was used, voxel size 3.5 x 3.5 x 3.0 mm, anterior to posterior phase encoding, 32-channel head coil. At the start of every run, seven Siemens dummy scans (unrecorded) and the two additional scans were discarded to achieve steady-state longitudinal magnetization.

After the functional runs, a T1-weighted magnetization-prepared rapid gradient echo (MP-RAGE) image was acquired for every participant in every session. They had an anterior/posterior phase encoding direction, TR 2.3 s, TE 3.9 ms, voxel size 1 x 1 x 1 mm.

For each participant the functional scans started approximately 20 minutes after the cTBS (due to logistical reason associated with the location of the TMS lab and eye-tracker calibration). The start of the scans was kept the same for the three sessions and ended maximally 50 minutes after cTBS. This period fits earlier experiments using cTBS and fMRI (Cameron et al., 2015).

At the start of each MRI session, participants were given earplugs to protect their hearing; their head was stabilized with foam pads to prevent movement. A tape was placed over their forehead to give them feedback in case they moved. Participants were given a panic button they could press in case of sudden problems inside the scanner. Eye movements were monitored online using an EyeLink 1000 infrared camera (SR Research, Ottawa, Canada) to make sure participants were performing the task. For eight of the participants, their heart rate was recorded in all three sessions with a pulse sensor (Brain Products GmBH, Gilching, Germany) on their left index finger.

For the patients, tremor was measured in the arm of their most affected side using MRI compatible carbon-wired silver/silver-chloride electromyography (EMG) electrodes (Brain Products GmBH, Gilching, Germany) on the muscle bellies of the flexor and extensor and a neutral electrode on the head of the ulna. The data was amplified using a BrainAmp MRI compatible EEG amplifier (Brain Products GmBH, Gilching, Germany) and transmitted to a PC outside the MR room using an optical cable. EMG data was recorded using BrainVision Recorder 1.03.002 (Brain Products GmBH, Gilching, Germany). Patients also had an MRI-compatible 3D acceleration sensor (Brain Products, Gilching, Germany) on the hand of their most affected side.

Table 1
Average (mm) ± SD MNI coordinates of the ROIs of all participants

	X	Y	Z
cTBS sites			
rIPS (rPEF)	17±5	-70±4	55±8
rS1	12±3	-42±4	72±3
Other ROIs			
lPEF	-14±5	-71±4	53±6
rFEF	34±8	-1±5	54±8
lFEF	-27±4	-2±4	52±8
rDLPFC	42±5	41±6	27±6
lDLPFC	-38±3	41±4	27±7

Note. PEF = parietal eye fields, S1 = primary somatosensory cortex, FEF = frontal eye fields, DLPFC = dorsolateral prefrontal cortex, l = left, r = right.

2.5 TMS

2.5.1 Motor threshold

Each participant's motor threshold was determined using a hand-held bi-phasic figure-eight MagVenture butterfly coil (diameter 75 mm), and a MagVenture MagPro X100 stimulation system (Farum, Denmark). Participants were comfortably seated in a chair and were handed earplugs to protect their hearing. They were instructed to sit still during the course of the stimulation. The regions of the EMG electrodes were scrubbed and cleaned with alcohol to lower the impedance. EMG electrodes (Covidien, Minneapolis, US) were put on the first dorsal interosseus (FDI) muscle of their right hand (active and reference) and the extension of the ulna (ground). Single pulses were delivered over the primary motor cortex (M1) by going 5 cm left and 1 cm up from the middle of the head (as measured from left ear to right ear and nasium to inion). Coil position was chosen to induce lateromedial current flow (45° from the mid-sagittal).

The following protocol was used; stimulation started at 20% intensity and increased with steps of 5% until the EMG gave a response higher than 0.050 mV. After this, the intensity went down until it the response exceeded 0.050 mV at least five attempts out of ten. Resting motor threshold was determined with the hand relaxed, active motor threshold was determined with the thumb and index finger loosely touching (in this case a threshold of 0.200 mV was used). Some participants held a pen to make sure they used the right force.

2.5.2 cTBS

Stimulation intensity for cTBS was set at 80% AMT (active motor threshold). The TMS coil was aligned to PEF or S1 using a MRI-based neuro-navigation system (Localite, Sankt-Augustin, Germany). The cTBS paradigm consisted of 50 Hz triplets (three pulses separated by 20 ms) repeated at 5 Hz (every 200 ms) over a period of 40 s (600 pulses in total), sent by Spike2 software (Cambridge Electronic Design, Cambridge, UK). Coil position was chosen to induce lateromedial current flow (45° from the mid-sagittal). The average coordinates of rPEF and rS1 in normalized space are shown in Table 1.

2.6 Data analysis

Because of technical problems with eye tracking in the MRI scanner in this population (ptosis of the eyes, unclear eyes, problems with corneal reflection, partially visible pupil), it was not possible to analyse the eye tracking data. Thus, eye movement data will not be discussed further.

Imaging data was analysed using Statistical Parametric Mapping, SPM 8 (Wellcome Trust Center for Neuroimaging, London, UK) and Statistical Package for the Social Sciences, SPSS 22 (IBM, Armonk, US). The first two volumes were discarded for steady-state magnetization. For each session and each participant, the following preprocessing steps were taken. The data was realigned using a two-pass procedure (all volumes were first realigned to the first volume, then a mean is created, after which the second pass aligns the images to the mean). The data was coregistered using a rigid-body transformation with the structural scan as the reference and the mean realigned scan as the source. The data was normalized to MNI space using the segmented structural scan and then smoothed using an 8 mm FWHM Gaussian kernel.

2.7 ROI analysis

The first session of every participant was analysed using a single-subject GLM and applying a whole-brain contrast of anti-saccade trials (the 1.5 s response period after the visual stimulus) vs. baseline (the fixation epochs, not modelled explicitly in the GLM). We defined the following ROIs as 5

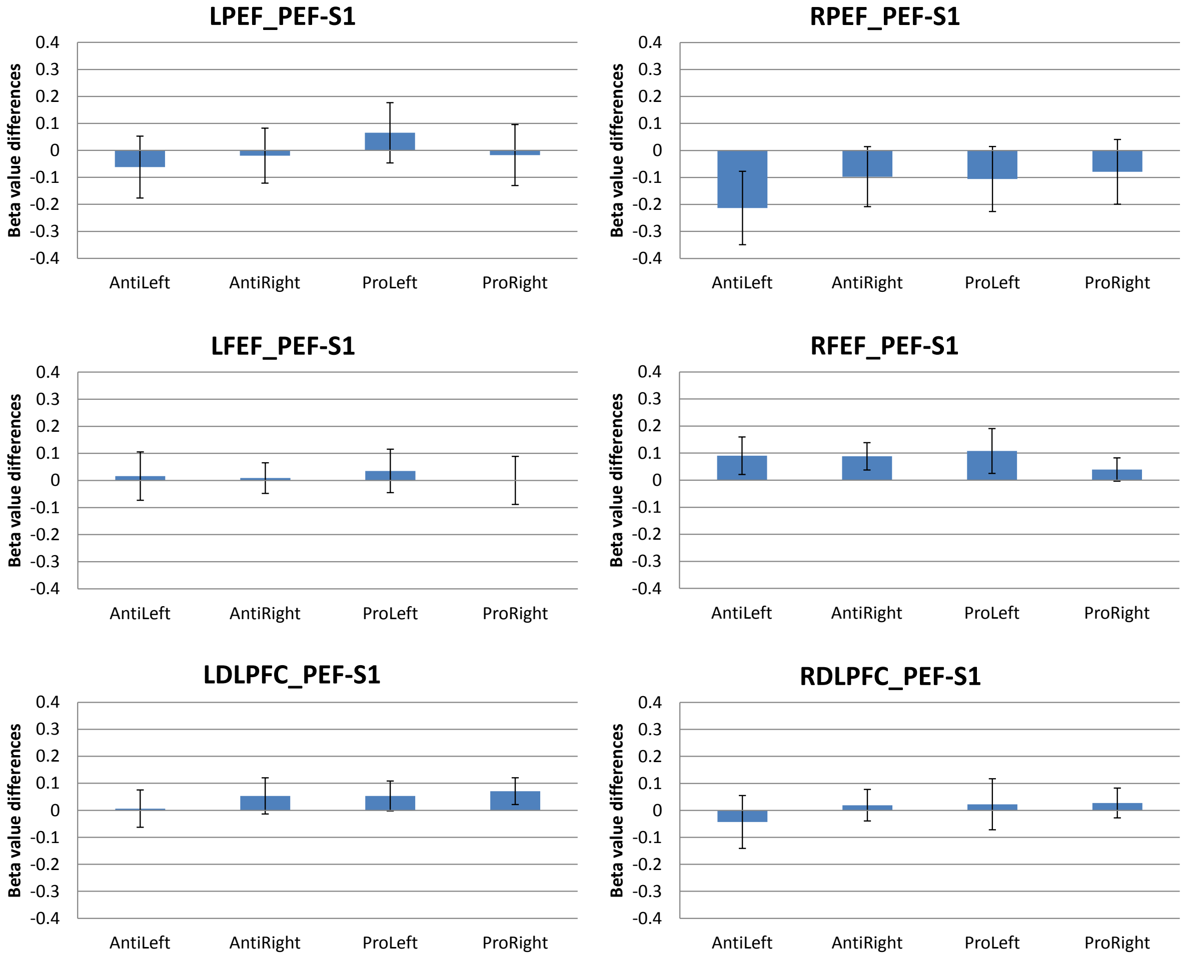

Fig. 3. Beta value differences (cTBS to S1 subtracted from cTBS to PEF) of control participants for each ROI.

mm cubic voxels on a participant's anatomical scan centred on the coordinates for peak activations: right parietal eye fields (rPEF), left PEF, right and left frontal eye fields (FEF) and right and left dorsolateral prefontal cortex (DLPFC). The frontal eye fields were chosen as the most posterior area of the middle frontal gyrus and the precental sulcus. The parietal eye fields were chosen as the most medial peak activation in the intraparietal sulcus. The dorsolateral prefrontal cortex was chosen as the middle frontal gyrus, anterior to the lateral ventricles. An example of these ROIs can be found in Figure 2. As a control site, the right S1 was chosen as the superior extent of primary sensory cortex, approximately 9 mm from the sagittal axis (Cameron et al., 2015), thus corresponding to the somatosensory region of the lower limb.

The right PEF and right S1, as determined after ROI analysis of session 1, were used to direct the TMS coil during the two cTBS sessions, session 2 and 3. The six ROIs (r/l PEF, r/l FEF, r/l DLPFC) were used in an ROI-based GLM comparing the two cTBS sessions. The average MNI coordinates of all ROIs are shown in Table 1. Additionally, Figure 2 depicts an example of the ROIs used for analysis. The betas of these six ROIs were extracted using MarsBaR (Brett, Anton, Valabregue, & Poline, 2002) running in SPM8. The normalized coordinates from session 1 were used to extract betas of the six ROIs from single-subject GLMs in session 2 and 3, modelled in the same fashion.

2.8 Statistical analysis

The betas of left/right PEF, left/right FEF and left/right DLPFC were analysed using SPSS in a 6 x 2 x 2x 2 repeated measures ANOVA with the six ROIs (left/right PEF, left/right FEF, left/right DLPFC), stimulation site (PEF, S1), task (pro-saccade, anti-saccade) and saccade direction (left, right) as within-subject factors. Post-hoc *t*-tests were conducted on the difference between PEF stimulation and

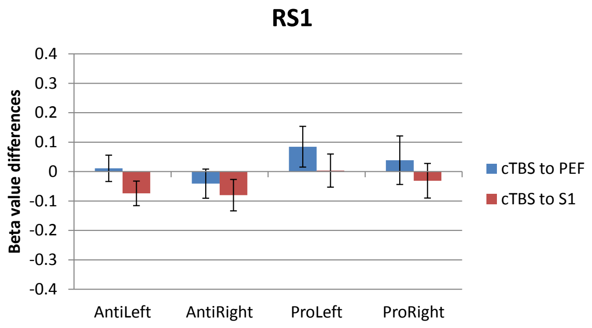

Fig. 4. Beta value differences of right S1 for control participants after cTBS to PEF and S1.

S1 stimulation. When sphericity was violated, a Greenhouse-Geisser correction was applied.

3. Results

A 6 x 2 x 2 x 2 repeated measures ANOVA with the six ROIs (left/right PEF, left/right FEF, left/right DLPFC), two stimulation sites (PEF, S1), two tasks (pro-saccade, anti-saccade) and two saccade directions (left, right) as within-subject factors was conducted on the betas from session 2 and 3. We found a main effect of ROI, $F(2.56, 15.33) = 5.54$, $p < .05$, $\eta^2 = 0.48$. This means that across the two stimulation sites, across the two tasks, and across the two saccade directions the activity was different in the six ROIs; these ROIs of the oculomotor network had different activity in eye movements (Fig. 3 illustrates the beta value differences). We also found a main effect of task, $F(1, 6) = 8.32$, $p < .05$, $\eta^2 = 0.58$. Across the ROIs the BOLD activity was higher for anti-saccades ($M = 0.49$, $SD = 0.33$) than for pro-saccades ($M = 0.41$, $SD = 0.28$). We also found a main effect of saccade direction, $F(1, 6) = 8.36$, $p < .05$, $\eta^2 = 0.58$, with saccades to the left having greater BOLD activation ($M = 0.48$, $SD = 0.32$) than pro-saccades ($M = 0.42$, $SD = 0.30$).

The main effect of stimulation site was not significant; this means across all ROIs, BOLD activation was not generally higher or lower after PEF stimulation as compared with S1 stimulation.

The interaction effects were not significant either, although the ROI x task interaction showed a trend with a large effect size, $F(2.13, 12.8) = 2.93$, $p = .09$, $\eta^2 = 0.33$.

Because PEF and FEF share similar roles in the visuomotor aspects of anti-saccade programming (Medendorp et al., 2005; Moon et al. 2007), we more specifically examined whether cTBS to PEF resulted in changes in BOLD signal in either PEF or FEF. Specifically, we tested if the difference in betas of PEF-S1 cTBS was different from zero for the four task / saccade direction combinations. Running these eight one-sample *t*-tests (2 ROIs, four task / saccade direction combinations), uncorrected, revealed no significance (for all, $p > .1$). This could be due to the large variation in betas across participants. When inspecting the data for individual participants (supplementary Fig. 1), we see that BOLD signal was reduced in the right PEF after cTBS to PEF for five out of seven participants. For the right FEF, BOLD increased for six out of seven participants. This data indicates the effect is quite consistent. For left PEF, left FEF and both left and right DLPFC's, the individual activation patterns are less clear.

To test if right S1 was indeed an appropriate control region for the oculomotor network, four one-sample *t*-tests were conducted across the task/saccade direction combinations on the difference in betas of PEF-S1. None of these tests, uncorrected, were significant (for all, $p > .3$). This indicates that there was no effect in the control region S1. Inspection of Figure 4 illustrates that the beta value differences in right S1 are around zero after cTBS to PEF and cTBS to S1.

Figure 5 captures the beta differences of the two patients. Examination of this data reveals different patterns as compared with the healthy controls (albeit only data from two patients), but we did not conduct statistical analyses on this data. The data of the Parkinson's disease patients will further be discussed in the discussion.

4. Discussion

In summary, we found BOLD activity to be different in the six ROIs (left/right PEF, left/right FEF, left/right DLPFC), for the two tasks (pro-saccade and anti-saccade) and for the two saccade directions (left and right). While the individual *t*-tests showed no significant effects and we cannot draw final conclusions from these results, we will describe the implications of the trends and the implications of the main effect of ROI in the context of previous studies and in the context of compensation. Specifically, we found that BOLD activity was reduced in right PEF and increased in right FEF after cTBS to PEF as compared with S1. Thus these trends do suggest potential compensation by FEF for perturbations to PEF in healthy controls.

4.1 BOLD signal after cTBS

After cTBS to right PEF, BOLD activity was

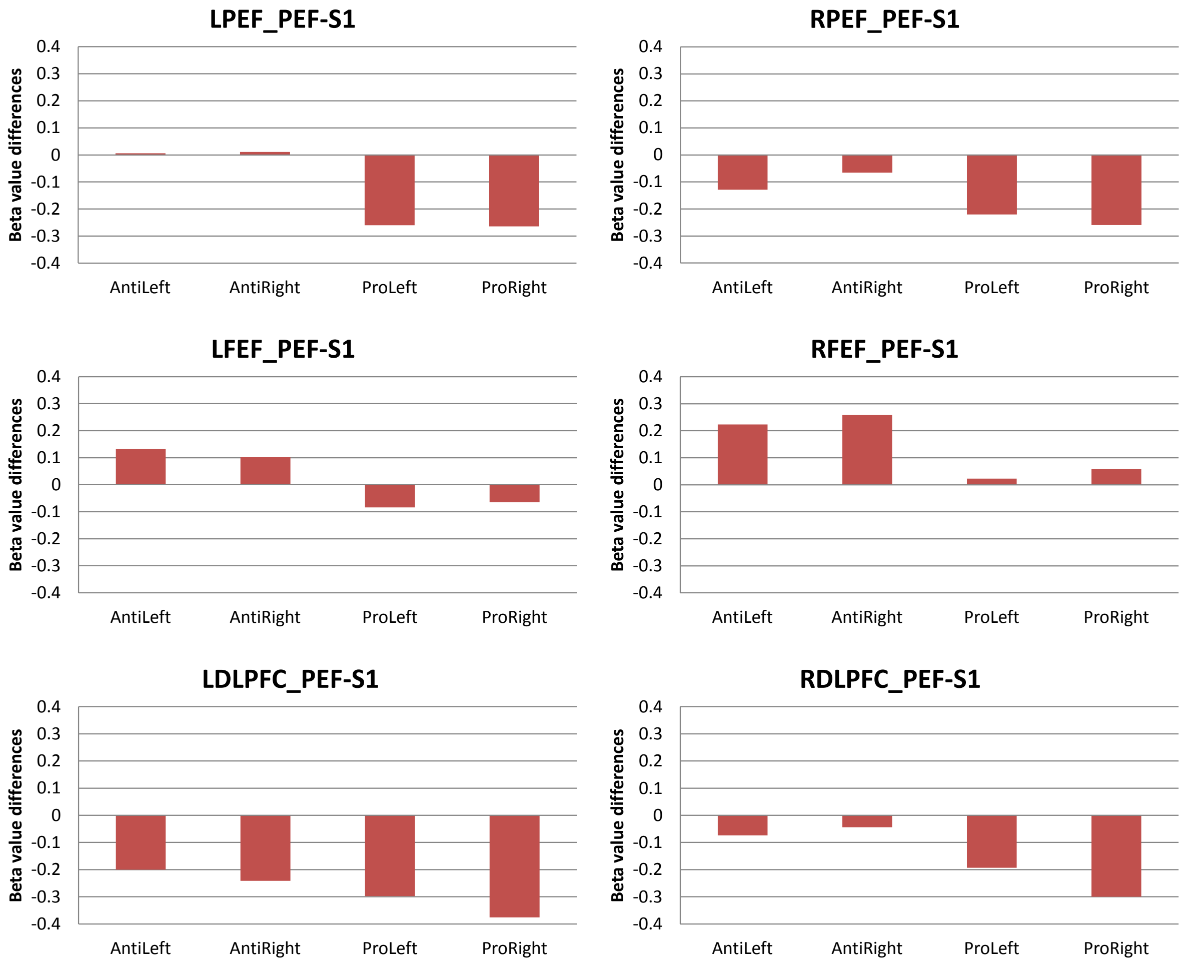

Fig. 5. Beta value differences (cTBS to S1 subtracted from cTBS to PEF) of patients for each ROI for patients.

reduced in this region, as compared with cTBS to S1. This finding suggests cTBS did have an inhibitory effect, if we assume that reduced BOLD signal corresponds to reduced function, an assumption that is common to the majority of neuroimaging studies (e.g., Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Schäfer et al., 2012; Stefanovic, Warnking & Pike, 2004). This inhibitory effect was not different for the different task/saccade direction combinations. In contrast to this decrease in right PEF, we found BOLD activity increased in the right FEF. However, there seemed to be no increase or decrease for left PEF, right/left DLPFC. This indicates compensation for perturbation of right PEF does not occur contralaterally in a region also involved in saccade programming, but rather occurs ipsilaterally in FEF. This could mean that though PEF and FEF are shown to have distinct roles in saccade programming, in a way these regions can take over each other's functions.

4.2 Previous TMS studies on compensation

In order to draw conclusions on whether increased BOLD activity, as shown in previous experiments, relates to compensation or pathophysiological activity, it is crucial to combine neuroimaging methods with a causal technique like TMS (Martinu & Monchi, 2013). Changes in BOLD activity in the visual cortex after TMS to FEF or PEF region have been shown by Ruff et al. (2009). Compensatory activity in distinct areas after TMS was found by O'Shea, Johansen-Berg, Trief, Göbel, and Rushworth (2007). After rTMS to left dorsal premotor cortex, activity in right dorsal premotor cortex increased. Importantly, activity in connected medial premotor areas increased as well, but only for task selection (as compared with task execution). These results show this compensation is functionally specific. This is in accordance with our findings; after perturbing right PEF, we found increased activity in right FEF. However, increased contralateral activity after cTBS has also been found in speech processing (Hartwigsen et al., 2013). This indicates compensation could also occur contralaterally in a functional similar region. Contralateral compensation was not found in the current experiment, and this might be due to the dominance of the right hemisphere in visuo-spatial processing in humans (Pisella et al., 2011).

Comparing the trends found in this experiment to previous TMS studies involving saccades, we could conclude they fit in an internal feedback system where the different oculomotor regions can compensate for each other's functions if one of the regions is impaired. For example, in a single pulse TMS paradigm where the cerebellum, parietal cortex or vertex were perturbed, Xu-Wilson, Tian, Shadmehr, and Zee (2011) found a delay in ongoing saccades. However, the oculomotor system could compensate for the perturbation because participants were still able to execute pro- and anti-saccades. Additionally the authors found no difference when they perturbed the cerebellum, parietal cortex or vertex. According to the authors, the finding that the oculomotor system could compensate for perturbation, points to an internal feedback system. Though this is evidence from a single-pulse paradigm, it is in accordance with our findings, in which the different oculomotor regions we thought to compensate for each other. This could also be concluded from the study of Cameron, Riddle, and D'Esposito (in preparation) that also used a pro-and anti-saccade task. They found after cTBS to FEF, BOLD activity in the oculomotor network was higher (as compared with cTBS to S1), especially when the stimulus was on the left (pro-left or anti-right trials). The authors suggested this may involve a compensatory network that did not succeed in restoring behaviour since anti-saccades were still impaired. Using Dynamic Causal Modelling (DCM), they found an increased role of PEF in anti-saccades, which also points to a compensatory role of PEF in the vector inversion process in anti-saccades. These findings and the results of the current experiment could demonstrate a system in which, in healthy participants, PEF and FEF can compensate for each other's function when one of the regions is perturbed.

4.3 Parkinson's disease

In the current study, only two Parkinson's patients were tested. Because of this small sample size, we did not statistically analyse activity in the six ROIs. However, by inspecting the difference in BOLD activation after cTBS to PEF as compared with cTBS to S1, we observed different patterns for pro-saccade trials and anti-saccade trials. For anti-saccade trials, activity in the right FEF increased. For pro-saccade trials, the results are less clear: activity in the left PEF seems to decrease. If these results are representative of the results of 30 patients, this may indicate that for voluntary saccades, the activation pattern is not different than for healthy control participants, because activity in right FEF increases after perturbation of right PEF. This can indicate that information about task set and vector inversion cannot be compensated for contralaterally by left PEF. In other words, the right (ipsilateral) FEF, although affected by Parkinson's pathology, has to compensate for it. An earlier study that examined saccades in Parkinson's disease (Cameron et al., 2012) found that the problems with voluntary saccades and the reduced activation of FEF in Parkinson's disease are mainly related to problems with task set signals, rather than the execution of the saccades.

As for pro-saccades, the current results suggest cTBS to right PEF also causes inhibition of left PEF. This may indicate in Parkinson's patients increased connectivity between left and right PEF exists. The decrease in PEF seems to be in conflict with theories that assume a dominance of the right hemisphere in visuo-spatial processing (Pisella et al., 2011). If the right hemisphere would be dominant, we would expect increased activity in left PEF after perturbation of right PEF, due to decreased inhibition of right PEF on left PEF. Importantly, in the current experiment we were not able to accurately measure eye movement behaviour. Hence, we do not have information on increased latency or reaction time of pro-saccades. Additionally, the current project focused on activity in PEF, FEF and DLPFC, but other oculomotor regions could have altered activity following cTBS to right PEF. These limitations make interpretation of the trends of pro-saccades in Parkinson's patients difficult.

Possibly, this difference (as compared with healthy controls) in activity in pro-saccades is part of the parietal compensation network that we think exists in Parkinson's disease. DCM could give us more insight into how the connectivity between the oculomotor regions in both hemispheres changes after perturbation of the right PEF in Parkinson's disease.

4.4 Limitations of the present study

To further examine the BOLD activation patterns after cTBS to PEF, it is critical to take eye movement behaviour into account. In the current study, this was not possible due to technical difficulties with eye tracking in older participants. Consequently, we cannot draw conclusions on the effect of cTBS to PEF on saccadic amplitude, reaction time, etc. in Parkinson's patients as well as healthy participants. This also means it was not possible to correct for errors, especially in the anti-saccade trials. Errors could have contaminated the effects of BOLD activation we found. However, for most participants, it was possible to observe their eye movements during testing, even though they were not correctly recorded. By observing these, we noticed the error rate in anti-saccade trials was less than 10%. This observation makes it unlikely that error-related activity drove the effects we found. Even without recording eye movements, it is possible to get more information from saccadic behaviour using the current paradigm. In the pro- and anti-saccade tasks, stimuli of different eccentricities were used. In the current analyses we collapsed across these amplitudes due to the limited effect of cTBS. However, future research could take into account the effect of amplitude without the use of saccadic behaviour.

Another limitation of the current study could be the limited amount of ROIs that were used. In the current analyses we focused on six oculomotor ROIs (left/right FEF, left/right PEF, left/right DLPFC), as determined prior to the start of the study. Because of multiple comparisons and the size of other regions, we did not examine activity in the supplementary eye fields, basal ganglia and superior colliculus. Hence, no conclusions can be drawn with respect to these other regions, both in Parkinson's disease pathology as well as compensational networks after cTBS in healthy participants.

4.5 Future directions

The paradigm described here makes it possible to study the altered activation patterns of the oculomotor network that were observed in Parkinson's patients. Studying eye movements, voluntary eye movements in particular, might tell us more about higher cognitive functions of Parkinson's disease patients. For example, in the study of Amador et al. (2006), deficits in voluntary eye movements were correlated with severity of the disease. Saccades can thus be a powerful and direct measure of cognitive functions in Parkinson's disease. Hence, future research could focus on the relation between the progression of the disease on the one hand, and the change in saccadic behaviour and the functioning of the oculomotor network on the other hand. Apart from gaining more insight in Parkinson's disease pathology and compensatory mechanisms, the current project could also contribute to the development of therapeutic interventions to prevent freezing and difficulties in initiating movements. In general, knowledge about compensatory networks in Parkinson's disease could give us more insight in how the brain can adapt to and compensate for pathology.

5. References

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences, 12*(10), 366–375.

Amador, S. C., Hood, A. J., Schiess, M. C., Izor, R., & Sereno, A. B. (2006). Dissociating cognitive deficits involved in voluntary eye movement dysfunctions in Parkinson's disease patients. *Neuropsychologia*, *44*(8), 1475–1482.

Bloem, B. R., Hausdorff, J. M., Visser, J. E., & Giladi, N. (2004). Falls and freezing of gait in Parkinson's disease: a review of two interconnected, episodic phenomena. *Movement Disorders*, *19*(8), 871–884.

Brainard, D. H. (1997). The Psychophysics toolbox, *Spatial Vision*, *10*(4), 443–446.

Brett, M., Anton, J., Valabregue, R., & Poline, J. 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. Available on CD-ROM in *NeuroImage*, *16*(2).

Cameron, I. G. M., Riddle, J. M., & D'Esposito, M. (2015). Dissociable roles of dorsolateral prefrontal cortex and frontal eye fields during saccadic eye movements. *Frontiers in Human Neuroscience*, *9*, 613.

Cameron, I. G. M., Riddle, J. M., & D'Esposito, M. (2016). *Theta burst TMS to frontal eye fields affects visuomotor processing rather than executive control.* Manuscript in preparation.

Cameron, I. G. M., Pari, G., Alahyane, N., Brien, D. C., Coe, C. C., Stroman, P. W., & Munoz, D. P. (2012). Impaired executive function signals in motor brain regions in Parkinson's disease. *NeuroImage*, *60*(2), 1156–1170.

Chan, F., Armstrong, I. T., Pari, G., Riopelle, R. J., & Munoz, D. P. (2005). Deficits in saccadic eye-movement control in Parkinson's disease. *Neuropsychologia*, *43*(5), 784–796.

Hartwigsen, G., Saur, D., Price, C. J., Ulmer, S., Baumgaertner, A., & Siebner, H. R. (2013). Perturbation of the left inferior frontal gyrus triggers adaptive plasticity in the right homologous area during speech production. *Proceedings of the National Academy of Sciences*, *110*(41), 16402–16407.

Helmich, R. C. G. (2011). *Cerebral reorganization in Parkinson's disease* (Doctoral thesis). Radboud University Nijmegen, Nijmegen, the Netherlands.

Helmich, R. C., Bloem, B. R., & Toni, I. (2012). Motor Imagery Evokes Increased Somatosensory Activity in Parkinson's disease Patients with Tremor. *Human Brain Mapping*, *33,* 1763–1779.

Henson, R. N. (2006). Efficient experimental design for fMRI. In K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny (Eds.), *Statistical Parametric Mapping: The analysis of functional brain images* (pp. 193–210). London: Elsevier.

Hood, A. J., Amador, S. C., Cain, A. E., Briand, K. A., Al-Refai, A. H., Schiess, M. C., & Sereno, A. B. (2007). Levodopa slows prosaccades and improves antisaccades: an eye movement study in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *78*(6), 565–570.

Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron*, *45*(2), 201–206.

Leoné, F. T. M., Toni, I., & Medendorp, P. (2014). Two-dimensional spatial tuning for saccades in human parieto-frontal cortex. *NeuroImage*, *87,* 476–489.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.

Martinu, K., & Monch, O. (2013). Cortico-basal ganglia and cortico-cerebellar circuits in Parkinson's Disease: Pathophysiology or compensation? *Behavioral Neuroscience*, *127*(2), 222–236.

Medendorp, W. P., Goltz, H. C., & Vilis, T. (2005). Remapping the remembered target location for anti-saccades in human Posterior Parietal Cortex. *Journal of Neurophysiology*, *94*(1), 734–740.

Moon, S. Y., Barton, J. J. S., Mikulski, S., Polli, F. E., Cain, M. S., Vangel, M., Hämäläinen, M. S., & Manoach, D. S. (2007). Where left becomes right: A magnetoencephalographic study of sensorimotor transformation for antisaccades. *NeuroImage*, *36*, 1313–1323.

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience*, *5*, 218–228.

Van Nuenen, B. F. L., Helmich, R. C., Buenen, N., van de Warrenburg, B. P. C., Bloem, B. R., & Toni, I. (2012). Compensatory activity in the extrastriate body area of Parkinson's disease patients. *The Journal of Neuroscience*, *32*(28), 9546–9553.

O'Shea, J., Johansen-Berg, H., Trief, D., Göbel, S., & Rushworth, M. F. S. (2007). Functionally specific reorganization in human premotor cortex. *Neuron*, *54*(3), 479–490.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision*, *10*(4), 437–442.

Pierrot-Deseilligny, C., Rivaud, S., Gaymard, B., Müri, R., & Vermersch, A. I. (1995), Cortical control of saccades. *Annals of Neurology*, *37*(5), 557–567.

Pisella, L., Alahyane, A., Blangero, A., Thery, F., Blanc, S., & Pelisson, D. (2011). Right-hemispheric dominance for visual remapping in humans. *Philosophical Transactions of the Royal Society*, *366*, 572–585.

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*(12), 2008–2039.

Ruff, C. C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Weiskopf, N., & Driver, J. (2009). Hemispheric differences in frontal and parietal influences on human occipital cortex: Direct confirmation with concurrent TMS-fMRI. *Journal of Cognitive Neuroscience*, *21*(6), 1146–1161.

Schäfer, K., Blankenburg, F., Kupers, R., Grüner, J. M., Law, I., Lauritzen, M., & Larsson, H. B. W. (2012). Negative BOLD signal changes in ipsilateral primary somatosensory cortex are associated with perfusion decreases and behavioral evidence of functional inhibition. *NeuroImage*, *59*(4)*,* 3119–3127.

Stefanovic, B., Warnking, J. M., & Pike, G. B. (2004). Hemodynamic and metabolic responses to neuronal inhibition. *NeuroImage*, *22*(2), 771–778.

Xu-Wilson, M., Tian, J., Shadmehr, R., & Zee, D. S. (2011). TMS perturbs saccade trajectories and unmasks an internal feedback controller for saccades. *The Journal of Neuroscience*, *31*(32), 11537–11546.

The Effect of Whole Body Translation on Hand Choice

Roel Weijer[1]
Supervisors: Pieter W. Medendorp[1], Luc Selen[1], Romy Bakker[1]

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

We frequently choose which hand we use to interact with objects in our surroundings. Three major factors have been found to influence the decision process: the energetic cost of the movement, the likelihood of accurately reaching the target, and a bias that shifts the decision in favour of the dominant hand. Previous literature has mostly focused on this decision process under body stationary conditions. Here we test the hypothesis that the decision governing hand choice is affected by whole body translation. Participants were seated in a vestibular sled, which accelerated sideways in a sinusoidal motion, while they had to move to a body fixed target. They were free to choose the hand they preferred. We recorded hand choices over a range of targets and fitted a psychometric line to estimate the point of subjective equivalence (PSE) which served as a measurement of the amount of right hand usage. Targets were presented at 4 different phases during the sled motion: maximum right- and leftwards velocity and maximum right- and leftwards acceleration. Results showed that hand usage was only significantly affected when targets were presented at maximum acceleration, not at maximum velocity. At maximum rightwards acceleration participants increased their left hand usage; an opposite effect was found for maximum leftwards acceleration. Results regarding reaction times were inconclusive as to how the brain anticipates inertial forces on the hand. However, we argue that it is possible that the brain can predict the path of the sled motion and hence, predict the inertial forces acting on the arm during the reach.

Keywords: hand choice, decision, reaching task, acceleration, inertia, LATER

Corresponding author: Roel Weijer; **E-mail:** roelweijer@live.nl

1. Introduction

One of the more frequent decisions we make on a daily basis is which hand to use to interact with objects. We pick up an object from the table, or point someone in the right direction without thinking about the hand to use. Even when our whole body is in motion, we do not think about which hand to select for the actions we make. Imagine yourself in a bus, for example, while approaching your stop. To notify the driver to get off, you must press the bell in front of you. Being right handed, you typically use your right hand in this act. However, just before you reach out, the bus takes a sharp turn to the left and you feel your body being pulled to the right. Due to these inertial forces it may be more preferable to use your left hand since it is already being pulled in the direction of the bell. But what drives the hand selection process when the body is in motion? So far, not much is known about this decision process while the body is in motion, but studies performed in body-stationary conditions point to various factors that influence the decision process governing hand choice. In this study we aimed to clarify how body motion affects hand choice.

First, in body-stationary conditions, the hand that is typically selected is located on the same side of the body midline as the object that needs to be acted upon, although there is a bias towards the dominant hand (Bryden, Pryde, & Roy, 2000; Fisk & Goodale, 1985; Gabbard, Tapia, & Helbig, 2003). Thus, people that are right handed tend to use their right hand to reach for targets located slightly left of the midline relative to the body (Oliveira, Diedrichsen, Verstynen, Duque, & Ivry, 2010).

Second, task complexity has been shown to affect hand choice (Bryden et al., 2000; Gabbard et al., 2003). Some tasks require higher precision or more force. For instance, to unlock a door you probably use your dominant hand to guide the key to the keyhole. However, if you only have to manipulate the door handle you could use your non-dominant hand too, since precision constraints are much lower.

A third factor in the decision process relates to the biomechanics of the arm, which affects movement variability and energy expenditure (Cos, Duque, & Cisek, 2014). Cos et al. (2014) showed that people predict biomechanical costs of movements before movement onset and use these predictions to choose the movement with the lowest energetic cost. For people to predict the biomechanical costs of a movement, they need to have an internal representation of both the limb and environmental dynamics. They need to take into account both the current position of the hand, as well as the distance from the hand to the target. While the findings by Cos et al. (2014) concerned target choice, it is also likely, but not yet shown, that people also use predicted biomechanical cost in the decision governing hand choice.

The final factor that influences the decision process is the success of the hand in previous acts. Stoloff, Taylor, Xu, Ridderikhoff, and Ivry (2011) showed that by increasing the likelihood of successfully reaching a target with the non-dominant hand or decreasing the likelihood of successfully reaching it with the dominant hand, people increased their use of the non-dominant hand for those targets (Stoloff et al., 2011).

Recently, Schweighofer et al. (2015) proposed a hand choice model that incorporated all these factors in the forms of effort, handedness bias and success rate (Schweighofer et al., 2015). Their model was based on findings in unimanual tasks, both forced and free hand choice trials. They showed that success rate of the left hand decreased more with target width than success rate of the right hand. Furthermore during the forced choice trials they showed that effort was larger for hand movements towards targets on the contra-lateral side of the body. Based on these findings, the model predicts right hand choice as a function of the before mentioned factors.

While all the factors followed from hand choice experiments in body-stationary situations, they may also play a role in hand selection when we are in motion. Consider a scenario where you want to make a reaching movement toward a target, while you are horizontally accelerated and the target is stationary relative to the body. The inertia of the arms has an effect on the forces required to make the movement. Since the biomechanical effects of the whole body acceleration are different for each arm, the costs of the movement do not increase equally. Thus, it may be more efficient to use the hand you would not normally use to reach for this target under stationary conditions (Cos, Belanger, & Cisek, 2011).

How do we sense body translations? When the motion is passively imposed there is a key role for the vestibular system, in particular the otoliths, which detect linear accelerations. Furthermore, body acceleration can be registered by body sensors like pressure on the skin. For instance, during takeoff in a plane you will feel increased skin pressure on your back as it is pressed against the chair. A combination of these signals will lead to a sense of acceleration which in turn can be used to estimate the instantaneous inertial forces on the body or predict the upcoming inertial forces.

In this study we investigated the effects of body acceleration on hand choice. Subjects were continuously sinusoidally translated along the inter-aural axis using a vestibular sled. While in motion, they had to perform a reaching movement to a body-fixed target, presented at different directions. Subjects were free to choose either their right or their left hand to perform the movement on each trial. Targets were presented at different phases during the motion of the sled. As described above, hand choice in stationary conditions is determined by effort, success rate and handedness. In our paradigm we only manipulated effort, by presenting targets during different phases of the sled motion, while both success rate and handedness remained equal across all phases of the sled motion. Therefore, our hypothesis was that hand choice would be most affected (i.e., biased) by the sled motion when the target was presented at maximum acceleration as opposed to maximum velocity, since acceleration has an effect on effort. Furthermore, we expected that acceleration towards the right would result in an increase in usage of the right hand as the inertia of the hands would make movements in the direction of the acceleration more difficult. For instance, during rightwards acceleration, the inertia of the hands will push them slightly to the left. Since the hands are being pushed leftwards relative to the body, we expect an increase in right hand usage as the threshold for picking the right hand will also shift to the left. During maximum velocity we expect that hand choices are less affected since, velocity does not cause changes in inertial forces. Opposite effects during maximum acceleration may be indicative of an anticipatory mechanism because the sled is in continuous sinusoidal motion and biomechanical costs change when the arm is undergoing passive acceleration. We expect a person to select the hand for which the lowest cost is anticipated. Hence, the choice may depend on the anticipated biomechanical costs rather than the costs at the moment when the choice is made.

2. Methods

2.1 Participants

Nine healthy subjects (six female, ages 21–54, one left handed) with normal or corrected to normal vision and free of any known sensory or motor disorders participated in our study. All participants gave their written informed consent in accordance with the institutional guidelines of the ethics committee of the Social Sciences Faculty of the Radboud University Nijmegen. One (male, age 28, right handed) out of nine participants was excluded because of failure to comply with the task. Participants were orally instructed for the task and performed practice trials. Subjects performed two sessions of approximately 75 minutes each, tested on two different days. Due to technical failure, we lost a significant number of trials (136 trials) during the first session of one participant; however, the participant was not excluded.

2.2 Experimental setup

Participants were seated in a motion sled (Fig. 1A) that performed continuous sinusoidal motion with an amplitude of 0.15 meters and period of 1.6 seconds. The targets were 15 LEDs integrated in a table fixed to the chair (Fig. 1B). The targets were placed on a semicircle with a radius of 28 cm. They ranged from -61° to 61°, with the middle target at 0°, and were spaced 10° apart, except for the four targets closest to the middle target which were spaced 5° apart. The starting positions for the hands were located at 2.5 cm from the centre of the semicircle. Furthermore there was a fixation LED placed at 15 cm from the centre of the semicircle, for eye fixation. Hand movement was recorded using Optotrak (100 Hz). An infrared light was attached to the right and left index finger. Finger positions were expressed relative to the subject's body.

2.3 Paradigm

Participants were only allowed to use a single hand to move to the target during a trail. However, they were free to choose whichever hand they preferred. We would register their choice and observe how it changed over the range of targets, by use of psychometrics. Before the target was presented, they had to wait for a series of cues. The cues consisted of four high frequency tones, spaced at 500 ms intervals of which the last one coincided with the onset of the target. They served to make the stimulus presentation more predictable and hence, they might improve the reaction time. While waiting they had to keep their index fingers on the starting positions and fixate their eyes on the fixation light (Fig. 1B). After the final cue, participants were to move one of their hands to the target as fast as possible, with an accuracy of 5 cm to the target. When the target was

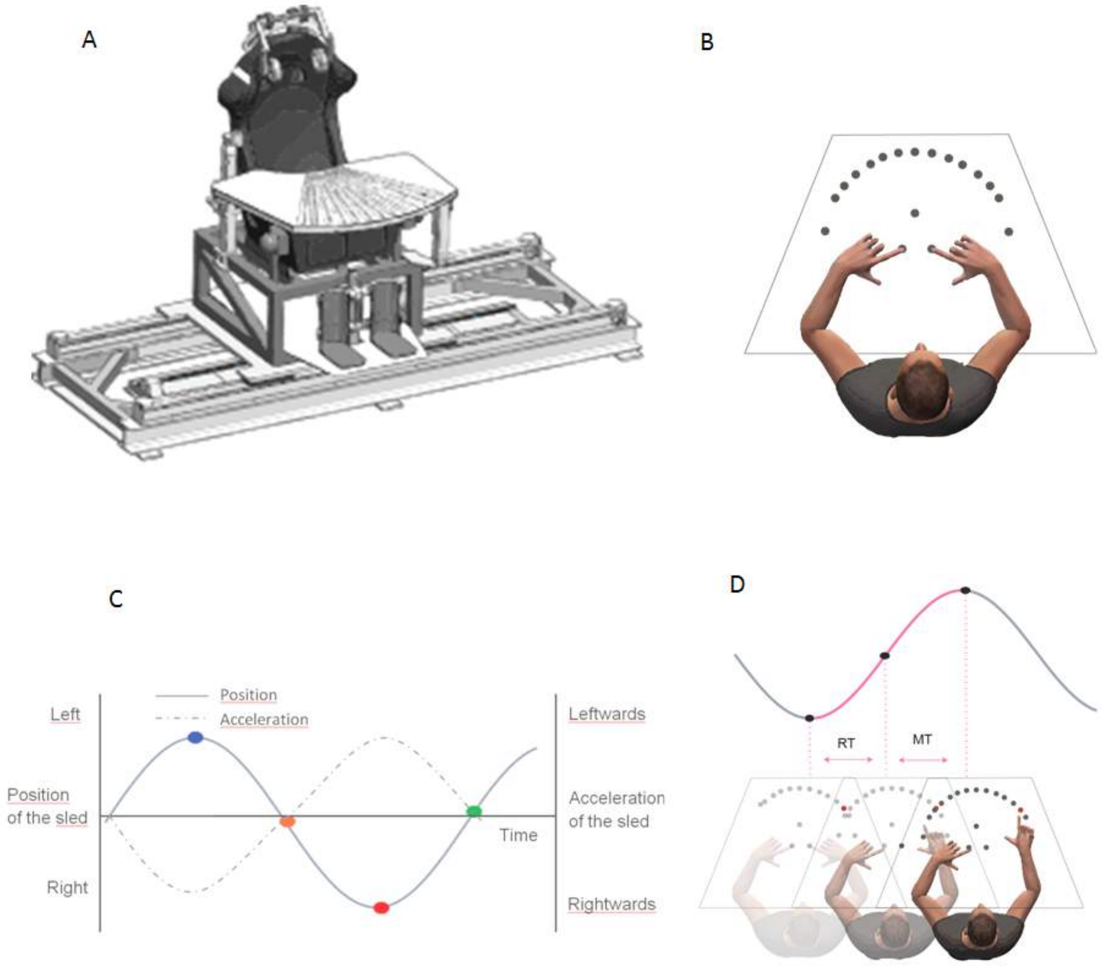

Fig. 1. Setup and paradigm. **A.** Motion sled. During the experiment the sled moved sinusoidally with an amplitude of 0.15 m. **B.** Table with integrated LEDs. 15 LEDs were positioned in a semicircle (radius 28 cm). At the beginning of each trial, subjects had to position their hands at the starting locations as shown in the figure. **C.** Sled motion over time. Colored ovals represent the four different phases on which target presentation was possible. Acceleration profile is shown as a dashed line **D**. Typical trial. A typical trial from target presentation until the target was reached is shown in pink. (RT = reaction time, MT = movement time)

reached within 1.1 sec a high frequency sound would be played, otherwise, a low frequency sound would be played. There were four different phases of sled motion at which a target could be presented (Fig. 1C). Targets were presented in blocks of 136 trials with a total of 12 blocks. In six blocks the target was presented at maximum velocity, either leftward or rightward (Fig. 1C, green and orange ovals) and in the other six block it was presented at maximum acceleration leftward or rightward (Fig. 1C, blue and red ovals). The blocks followed an interleaved pattern which was counterbalanced across participants. On half of the trials in a block, the target was presented at leftwards velocity or acceleration, on the other half it was presented at rightwards velocity or acceleration. One out of 15 trials was a catch trial during which two targets were presented and the subject had to use two hands to reach to them, this to prevent a pre-determined hand-choice. Figure 1D shows a typical trial from target presentation until the target is reached.

2.4 Data analysis

Analyses were performed in Matlab and SPSS. First, we analysed choice behaviour, (i.e., which hand the subject chose to reach for the target). Secondly, we analysed the kinematic data to determine how fast the responses were.

2.4.1 Choice behaviour

To see if the different target presentation phases had an effect on hand preference, we calculated the

Point of Subjective Equality (PSE) for each phase of the sled motion. A positive PSE reflects preference for left hand usage. The PSE was calculated, for each participant by using a cumulative Gaussian function to summarize the observed hand choice over the 15 presented targets, the angles of which ranged from −61° to 61°:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(y-\mu)^2}{2\sigma^2}} dy$$

in which x represents the target angle. The mean of the cumulative Gaussian, μ, represents the PSE. Parameter σ, which is inversely proportional to the slope of the psychometric curve, shows the variability of the participant's responses. Hence, a low σ corresponds with low variability. y is an integration variable that runs within the target range. Parameters were based on maximum likelihood estimates. The PSE represents the target direction for which there is equal probability of choosing the right or left hand and could therefore be located in between of two targets despite the fact that we never actually presented a target at that location. We compared the difference in PSE between the maximum acceleration phases and the maximum velocity phases. A Wilcoxon rank sum test was used to compare this difference ($\alpha = .05$), since the distribution was not normally distributed.

2.4.2 Reaction times

Hand movements were analysed using the kinematic data from Optotrak. Unfortunately, not every trial could be used for hand movement analysis since the finger markers were sometimes invisible to the cameras. Hand trajectories were computed from the Optotrak data. We differentiated these position traces to obtain a velocity trace. We defined the start of the movement (RT) at the time when the speed of a hand exceeded 0.1 m/s after target presentation. In case a participant changed his initial hand choice, we only calculated RT for the hand that ultimately reached the target. We analysed each trial for completeness to select trials that were suitable for data analyses. Supplementary Figure 1 contains a flowchart used for data selection.

First, we tested if there was a difference in RT between trials in which the right or the left hand was chosen, using a Wilcoxon rank sum test since the RTs were not normally distributed. We expected RT to be slower for targets for which there was high competition between choosing the left or the right arm, which are the two targets around the PSE. Targets at more peripheral locations, the two most extreme targets, have a defined choice. We also tested for the effect of sled motion and hand usage on RT, by using a repeated measures ANOVA. In total there were three factors with different levels: Right or left hand (2), targets at PSE or peripheral (2) and the position of the sled when the LED was shown (4). We choose to use a repeated measures ANOVA despite the violation of the normality distribution since, it was only marginally violated and repeated measures ANOVA is known to be robust for this violation. Therefore, this approach allowed us to look for interaction effects as well. Furthermore, we expected that RT and the time it took to complete the reaching movement might have an influence on hand choice, since the inertial forces on the hands still change after stimulus presentation. Figure 1D shows that during a trial the position of the sled changes. Hence, on a slow trial, the inertial forces on the hand will be different than on a fast trial. To test if the duration until a trial was completed had an effect on hand choice, we performed a logistic regression analysis within each participant where we determined how much of the hand choice was explained by the speed at which the trial was completed (quickness) while also accounting for the location of the target to which a reach was made. If the probability of choosing the right hand would significantly increase or decrease with quickness, this would be evidence that the brain predicts the acceleration of the sled such that it can use this information to choose the optimal hand. Quickness was determined as the time it took from stimulus presentation to the time that the hand movement reached maximum velocity. We chose the time to maximum hand velocity because this point is a good approximation of the average sled motion the hand undergoes during the reach. Additionally, we expected an interaction effect of quickness and stimulus presentation phase. During trials where the stimulus is presented at maximum acceleration, the direction of the acceleration switches while the participant is performing the trial as can be seen from Figure 1D. It switches because the time it takes the participant to move causes the sled to accelerate in the opposite direction. During the trials where the stimulus is presented at maximum velocity, this should happen less often since the acceleration of the sled remains in the same direction for a longer period of time hence, during quick trials the acceleration does not go towards the opposite direction. To test for a possible interaction effect of quickness and target presentation phase, we

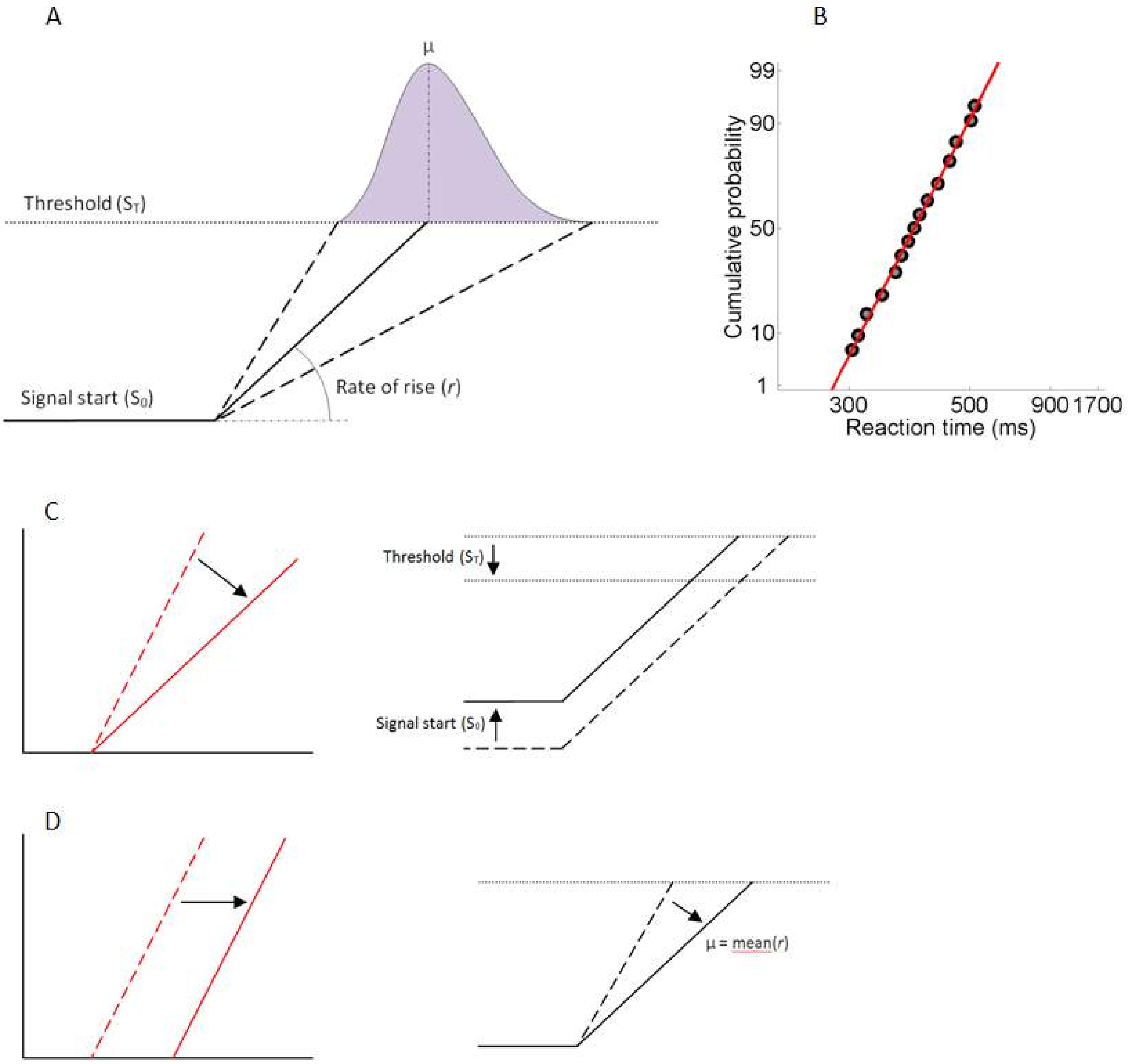

Fig. 2. LATER model **A.** LATER model and RT distribution. Starting at Signal start (S_0) evidence will accumulate according to the rate of rise (*r*) until it reaches a threshold (S_T) upon which a movement is made. Hence, the distribution of *r* (purple) is the RT distribution. **B.** Reciprobit line fitted to RT data. The predicted cumulative probability of RT is a line on transformed axes; the horizontal axis is RT^{-1}, the vertical axis is z-transformed probability. **C.** Example of steepness change of the reciprobit. Changes in steepness of the reciprobit line is reflected as a change in S_0 or S_T. **D.** Example of reciprobit shift. Reciprobit shift is reflected as a change in *r*.

performed a similar logistic regression in which we included the interaction effect to the model.

2.4.3 LATER model

Finally, we tried to quantify how reaction times, hand choices and different phases of the whole body motion are linked. To this end we used the LATER model (Carpenter, Reddi, & Anderson, 2009). The LATER model is a race model, in which two decision processes accumulate evidence until one of them reaches threshold and a decision is gated (Gold & Shadlen, 2007). The first process to reach the threshold determines the decision outcome. Here two decision processes are running in parallel, setting up a competition to use the right hand versus the left hand. The rate at which the decision signal rises for each process (i.e., the rate at which the evidence accumulates) is expressed as *r* which varies from trail to trial as a Gaussian with mean μ and variance σ^2. The signal starts at S_0 and ends at a threshold S_T (Fig. 2).

The parameters of the LATER model have a fixed relationship with the parameters that describe the reciprobit plot of the corresponding RT data. Therefore, by fitting a reciprobit line to the RT data, we may get a better understanding of how a decision is made (Fig. 2B). For instance, a change in the steepness of the reciprobit line will lead to a shift in starting point of the decision process or a shift of the threshold (S_0 or S_T; Fig. 2C), while a horizontal shift of the reciprobit line leads to a change in evidence accumulation (*r*; Fig. 2D; Carpenter, 1981; Reddi & Carpenter, 2000). Using linear regression we fitted a reciprobit line to the RT data. As before we could discriminate between three factors that influenced the slope of the line; hand choice (2), target location (2) and motion phase (4). Therefore we fitted 16 different lines. The parameters of these fitted lines were then compared, using an ANOVA, and used to gain estimates of the parameters of the LATER model. In order to obtain sufficient trials per condition we pooled the data from the subjects. To this end we performed a mean correction, subtracting the average RT per subject to be able to get an across subject distribution of RTs without adding additional variance because of combining distributions with different means. The mean correction was performed because our analysis depended on the spread of the RT distribution, which without a mean correction becomes larger due to the individual differences between participants in mean RT. We expected that thresholds or evidence accumulation would differ in different phases of whole body motion, as the forces during maximum acceleration could influence the decision process.

3. Results

3.1 Choice behaviour

We studied hand preference during whole body motion. To this end we recorded hand choices over a range of target directions and found the direction for which there is equal probability of choosing left or right. Figure 3 shows the psychometric curves of a typical participant when the target was presented at maximum velocity, both left- and rightward, (left panel) and maximum acceleration, both left- and rightward, (rightward panel) of the sled motion. Data represent the proportion of trials on which the right hand was chosen. We found the PSE for all target presentation phases; a more negative PSE implies an increase in right hand usage. For the target presentation phases where velocity was maximal to the left and to the right, the PSE was at −2.40° and −3.07°, respectively. For targets presented at maximum acceleration, the PSEs were −0.61° and −5.04°, respectively. This was opposite to our initial hypothesis, as we expected rightwards acceleration to be associated with an increase in right hand usage.

This pattern of PSE shifts is consistent across participants: Figure 3 shows that the difference in PSE between left- and rightward acceleration was larger than the difference between PSE at maximum

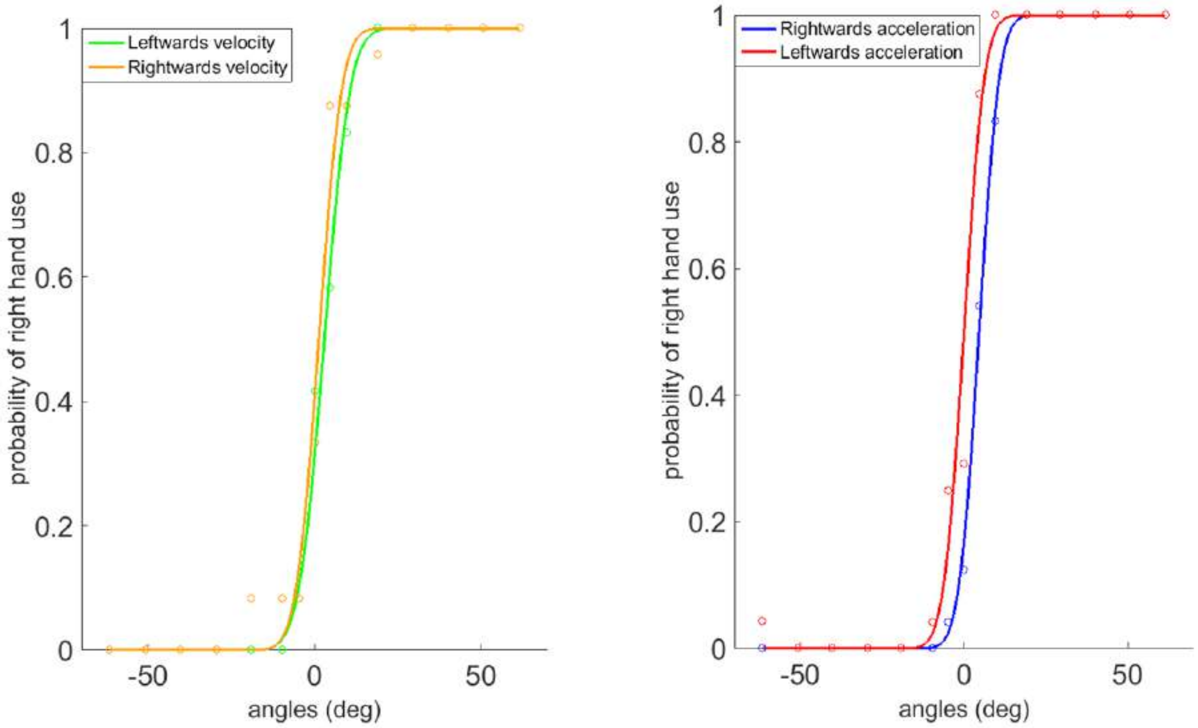

Fig. 3. Psychometric curves of a typical participant (participant 7). Circles represent the proportion of trials on which the right hand was used.

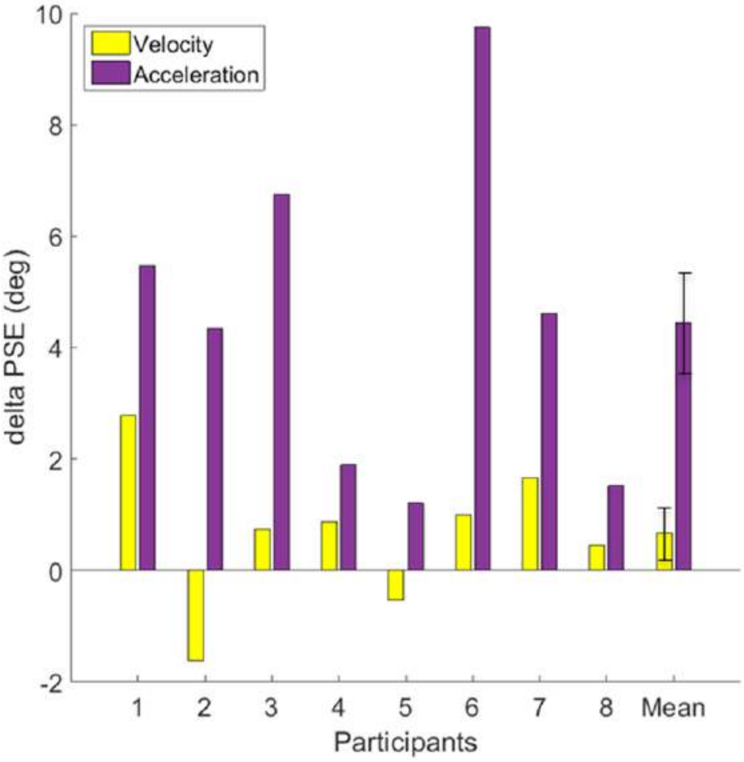

Fig. 4. Delta PSE for each participant in the velocity and acceleration phases. Delta PSE was calculated as the difference between PSE at the two phases of target presentation during velocity and acceleration. The final set of bars contains the mean over all participants with standard error of the mean (SEM).

left- and rightward velocities. Figure 4 shows this shift in PSE for all participants for maximum velocity and maximum acceleration phases.

When stimulus presentation occurred at maximum velocity the mean difference between the PSEs at the two different stimulus presentation phases was 0.67° (*SD* = 1.33°), which is non-significant ($t(7) = -1.43$, $p = .20$). On the other hand the difference between the PSEs at maximum acceleration was 4.44° (*SD* = 2.93°), which is significant ($t(7) = -4.29$, $p = .004$).

3.2 Reaction times

In order to gain a better understanding of why this observed shift in PSE was present, we analysed the RTs of the hand choices of selected trials (see methods). In total about 13% of trials were discarded, due to missing markers, with a minimum of 7% and a maximum of 21% per participant. First we looked at the difference between the RTs of the right and the left hand. Across subjects, the mean RTs for the left hand and right hand were not significantly different: 378 ms (*SD* = 72 ms) and 376 ms (*SD* = 71 ms), respectively ($z = 1.35$, $p = .025$). RT was slower when targets were presented during maximum velocity (385 ms, *SD* = 75 ms) than during maximum acceleration (373 ms, *SD* = 68 ms; $z = 8.58$, $p < .001$). The time until maximum velocity of the reaching movement occurred was also shorter when the target was presented at maximum velocity (143 ms, *SD* = 48 ms) than at maximum acceleration (139 ms, *SD* = 48 ms; $z = 4.30$, $p < .001$).

Furthermore, we tried to find proof for a higher competition between the two arms in the decision process for targets around the PSE. Therefore, we looked at the difference in RT between targets around PSE and the peripheral targets. Figure 5 shows the RT for each of the three factors, as defined in the methods section; phase (4), hand (2) and target location (2). The different colours in the graphs represent the four phases at which a stimulus was presented. The outer bars of each set of four bars represent the RT for the peripheral targets; the inner bars represent RT for targets around the PSE. The two left bars of each set of four bars represent RT for trials made with the left hand; the two right bars represent RT for trials made with the right hand. We found a main effect of phase on RT ($F(3) = 10.89$, $p = .012$) and an interaction of phase and target position ($F(3) = 9.35$, $p = .017$).

A logistic regression was performed to ascertain the effects of target presentation phase, target location and time until maximum velocity of the reaching movement on the likelihood that a

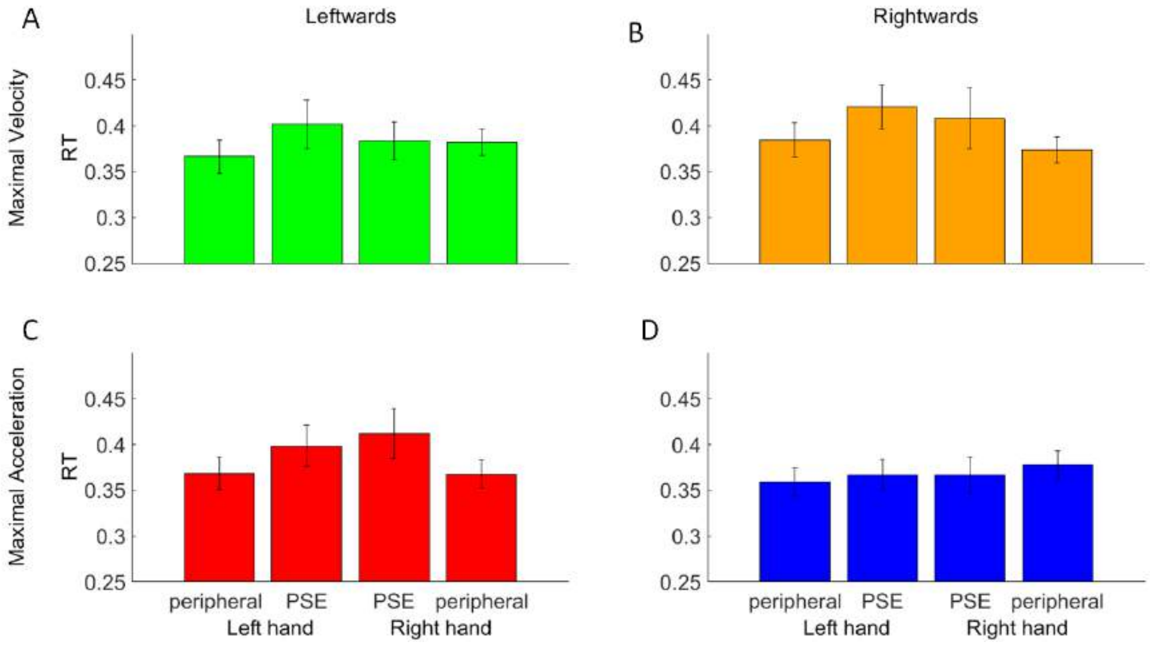

Fig. 5. Mean RT with SEM for all 16 conditions. **A.** Stimulus presentation at maximum leftwards velocity. **B.** Stimulus presentation at maximum rightwards velocity. **C.** Stimulus presentation at maximum leftwards acceleration. **D.** Stimulus presentation at maximum rightwards acceleration.

participant chose the left hand. Logistic regressions were statistically significant for all participants (Table 1). The models explained between 80% and 93% (Nagelkerke R^2) of the variance in choice behaviour and correctly predicted between 88% and 96% of hand choices. These high values were to be expected as most of the choice behaviour can be explained by the location of the target. Table 2 shows that more rightward positioned targets were associated with decreased left hand use. Furthermore, quickness did not add significantly to the model for half of the participants and varied in effect direction for the others. A separate logistic regression analysis, which included an interaction effect between quickness and stimulus presentation phase, showed no significant effect of the interaction factor for any of the participants. Thus, we could find no proof for an effect of the time it took participants to complete the trial on hand choice preferences.

3.3 LATER model

Finally, we analysed the RT distribution using a LATER modelling approach (REF). After taking the inverse of the RT data (promptness), only one of 16 conditions was not normally distributed (left hand trials towards targets around the PSE during leftward acceleration). Figure 6 shows that the reciprobit fits for two arbitrarily picked conditions seem to have a different slope. Similar fits were made for the other conditions as well. Figure 7 shows the slopes for the reciprobit fits with 95% confidence interval (C.I.) for all conditions. Within each phase it can be seen that the slope for the reciprobit lines of the left hands are less steep as for the right hand when a reach is made for targets around the PSE, indicating that the RT distribution of the left hand is wider, hence the RT of the left hand is more variable than the RT of the right hand. This result was reflected by the ANOVA which showed a significant interaction effect of target presentation phase, target position and hand choice on the slope of the reciprobit line ($F(3) = 11.59$, $p = .001$). When relating this to the LATER model, it shows that the left hand should have a higher starting position or a lower decision

Table 1
Model fit per participant

Participant	$\chi^2(2)$
1	1110,803**
2	1247,780**
3	1361,011**
4	1120,862**
5	1315,981**
6	1217,808**
7	1441,947**
8	1725,076**

Note. Likelihood ratio $\chi^2(\text{df})$, ** significant at $p < .000$.

Table 2
Model variables per participant

				95% C.I. for Exp(B)	
Participant	Factor	Wald(1)	Exp(B)	Lower	Upper
1	Quickness**	12.845	112.505	8.501	1488.951
	Target position**	141.738	0.811	0.784	0.840
	Constant**	16.644	0.049		
2	Quickness**	22.001	896.339	52.331	15352.616
	Target position**	195.453	0.848	0.828	0.868
	Constant**	30.128	0.016		
3	Quickness	0.469	0.431	0.039	4.792
	Target position**	177.054	0.817	0.794	0.842
	Constant	1.904	2.973		
4	Quickness	3.653	0.030	0.001	1.093
	Target position**	163.346	0.834	0.812	0.858
	Constant*	5.098	7.419		
5	Quickness*	8.565	0.003	0.000	0.149
	Target position**	151.157	0.788	0.759	0.818
	Constant*	8.738	18.143		
6	Quickness	0.061	1.582	0.042	59.365
	Target position**	255.176	0.870	0.856	0.885
	Constant*	7.823	0.070		
7	Quickness	0.190	0.424	0.009	20.024
	Target position**	159.357	0.757	0.725	0.790
	Constant	1.081	3.195		
8	Quickness*	5.834	0.005	0.000	0.366
	Target position**	124.111	0.608	0.557	0.663
	Constant*	4.453	13.320		

Note. Wald tests were used to determine the statistical relevance of each individual variable to the model. Exp(B) higher than one indicates a decreased odds of using the left hand when that particular variable increased. * Significant at $p < .050$, ** Significant at $p < .000$.

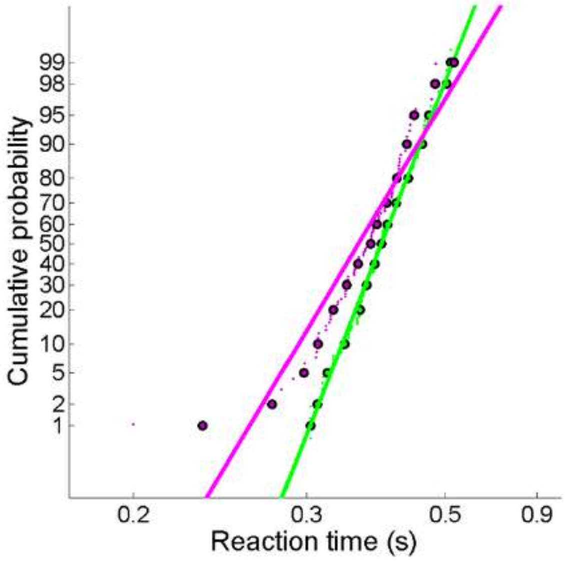

Fig. 6. Reciprobit plots. Reciprobit plots of RT for trials at maximum leftward velocity for targets around the PSE with the right hand (green) and maximum leftwards acceleration for targets around the PSE with the left hand (purple). Line is the fitted line to the data, black circles represent quantiles, and small dots represent cumulative mean corrected RT across participants.

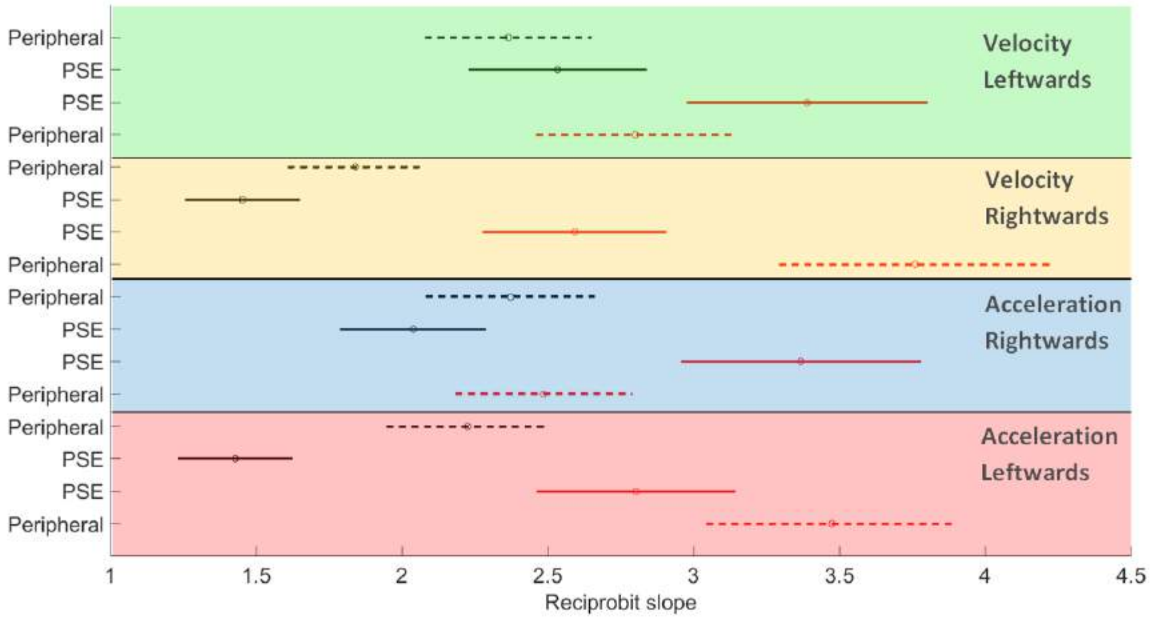

Fig. 7. Slopes of the reciprobit lines with 95% C.I. The different background colors represent the different phases of stimulus presentation. Dashed lines represent peripheral targets, while straight lines represent targets at PSE. Black lines belong to trials made with the left hand, red lines to trials made with the right hand.

threshold for trials towards targets around the PSE since, the left hand has a smaller reciprobit slope than the right hand for these trials.

4. Discussion

In this study we aimed to gain a better understanding of how hand choices are influenced by body motion. To this end, we performed an experiment where we translated people sideways while we studied their choice behaviour for hand reaches towards a range of targets. We found that hand choice was most severely affected when we presented the targets while the body was subjected to high acceleration. Our hypothesis was that the inertial forces of the arm generated by the movement of the sled would cause an increase in the usage of the arm corresponding to the direction of the acceleration: For instance, during rightward sled acceleration we would have expected an increase of right hand usage. The inertia of both hands will cause them to be pushed to the left. Therefore, we would expect that the threshold for choosing your right hand would also move a bit to the left. Contrary to our hypothesis, we found an increase in the usage of the hand that was on the opposite side of the direction of the acceleration. Hence, we found the threshold of our previous example to be moved to the right instead of to the left.

Initially, we assumed that the brain would base the arm choice on the forces experienced during target presentation. However, the reaching movement is not executed directly following the onset of this presentation. The time it takes for a response to be initiated and completed means that the experienced inertial forces during the reach are different from those derived from the instantaneous acceleration during target presentation. The PSE patterns observed in Figure 3 may be explained by such a predictive mechanism of the inertial forces actually experienced during the reach. It has already been shown that the brain can make such predictions of the biomechanical costs of a movement based on intrinsic factors (Cos et al., 2011; Cos, Medleg, & Cisek, 2012). The brain gains an estimate of external forces by use of the vestibular system and skin pressure senses. Since the sled moved in a predictive sinusoidal pattern it would be possible to predict the path of the sled motion (Prsa, Jimenez-Rezende, & Blanke, 2015). Predictive mechanisms have also been shown in studies investigating smooth pursuit (Barnes, 1993; Han et al., 2005).

The average time from target presentation until maximum hand velocity was reached for maximum velocity trials was 528 ms (*SD* = 87 ms), while for maximum acceleration trials it was on average

]

511 ms (SD = 81 ms). The period of the chair movement was 1600 ms. Hence, irrespective of the phase of the sled motion, the actual arm reach would take place more than a quarter phase later (Fig 1D). This would mean that the inertial forces on the hands were in the opposite direction as at target presentation. To test if the time it took participants to complete the trial had an effect on hand choice we performed a logistic regression analyses. However, the results of the analysis remained inconclusive as it only showed a significant effect of time for half of the participants amongst whom the effect was both positive and negative. An explanation might be that the variation in the time it took to complete a trial was too large. By taking a look at Figure 1C we can see that if a participant were to move within ¼ period, acceleration would keep decreasing, when he or she would have reached between a ¼ and ½ period it would increase towards the other side, while if he or she would have moved between ½ and ¾ the acceleration would decrease again. Logistic regression assumes a linear relationship between the covariate and the dependent variable. For future research it would be interesting to study how acceleration is used in the decision process by making the chair movement maintain an acceleration profile towards the same direction for the entire trial length.

We could not find that competition between hand choices was higher for targets around the PSE than for the most extreme targets, based on RTs. Previous literature did show this effect in stationary conditions (Oliveira et al., 2010). It could be that the larger variation of RTs in our study obscures the effect and that a change in the experimental paradigm is needed to gain more precise measurements of the RT. Oliveira et al. (2010) reported standard errors of approximately 10 ms while we found the standard error of the mean (SEM) to be approximately 50 ms. Since the effect Oliveira et al. observed is close to 10 ms, we believe that a larger number of trials in combination with a larger number of subjects would reveal a similar competition effect as observed by Oliviera et al. (2010).

By studying the differences between the reciprobit lines of the different conditions we found that for the left hand the starting level was higher or the threshold was lower as for the right hand when the target was located close to the PSE. This implies that the decision process of the left hand was significantly more variable in time than that of the right hand. The underlying cause of the variation of the left hand might be that this hand was the non-dominant hand for all but one of our participants. Hence, either decreased confidence or less developed motor skills with these movements might have resulted in higher reaction time variability. It is possible that these variations only show themselves if task complexity is high enough, hence only for targets close to the PSE. Harris and Wolpert (1998) showed that movement variability depends on the size (complexity) of the control signal, strengthening our hypothesis. Further evidence can be found in a study investigating reaction time variability in ADHD patients. They showed that larger RT variability correlated with decreased motor skill development (Klotz, Johnson, Wu, Isaacs, & Gilbert, 2012). Finally, larger RT variability cannot be found when both hands need to be used simultaneously (Mickeviciene, Skurvydas, & Karanauskiene, 2015).

5. Conclusion

Our study has shown that hand choices can be affected by whole body translation and that acceleration seems to have a major role in the decision process. However, it remains unclear how the brain incorporates this factor in its decision.

References

Barnes, G. R. (1993). Visual-vestibular interaction in the control of head and eye movement: the role of visual feedback and predictive mechanisms. *Progress in Neurobiology, 41*(4), 435–472.

Bryden, P. J., Pryde, K. M., & Roy, E. A. (2000). A performance measure of the degree of hand preference. *Brain and Cognition, 44*(3), 402–414.

Carpenter, R. H. (1981). Oculomotor Procrastination. In D. F. M. Fisher, R.A. Senders, J.W. (Ed.), *Eye Movements: Cognition and Visual Perception* (pp. 237–246). Hillsdale: Lawrence Erlbaum.

Carpenter, R. H., Reddi, B. A., & Anderson, A. J. (2009). A simple two-stage model predicts response time distributions. *The Journal of Physiology, 587*(Pt 16), 4051–4062.

Cos, I., Belanger, N., & Cisek, P. (2011). The influence of predicted arm biomechanics on decision making. *Journal of Neurophysiology, 105*(6), 3022–3033.

Cos, I., Duque, J., & Cisek, P. (2014). Rapid prediction of biomechanical costs during action decisions. *Journal of Neurophysiology, 112*(6), 1256–1266.

Cos, I., Medleg, F., & Cisek, P. (2012). The modulatory influence of end-point controllability on decisions between actions. *Journal of Neurophysiology, 108*(6), 1764–1780.

Fisk, J. D., & Goodale, M. A. (1985). The organization of eye and limb movements during unrestricted reaching to targets in contralateral and ipsilateral visual space. *Experimental Brain Research, 60*(1), 159–178.

Gabbard, C., Tapia, M., & Helbig, C. R. (2003). Task complexity and limb selection in reaching. *International Journal of Neuroscience, 113*(2), 143–152.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*, 535–574.

Han, Y. H., Kumar, A. N., Reschke, M. F., Somers, J. T., Dell'Osso, L. F., & Leigh, R. J. (2005). Vestibular and non-vestibular contributions to eye movements that compensate for head rotations during viewing of near targets. *Experimental Brain Research, 165*(3), 294–304.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature, 394*(6695), 780–784.

Klotz, J. M., Johnson, M. D., Wu, S. W., Isaacs, K. M., & Gilbert, D. L. (2012). Relationship between reaction time variability and motor skill development in ADHD. *Child Neuropsychology, 18*(6), 576–585.

Mickeviciene, D., Skurvydas, A., & Karanauskiene, D. (2015). Is intraindividual variability different between unimanual and bimanual speed-accuracy movements? *Perceptual and Motor Skills, 120*(1), 125–138.

Oliveira, F. T., Diedrichsen, J., Verstynen, T., Duque, J., & Ivry, R. B. (2010). Transcranial magnetic stimulation of posterior parietal cortex affects decisions of hand choice. *Proceedings of the National Academy of Sciences of the United States of America, 107*(41), 17751–17756.

Prsa, M., Jimenez-Rezende, D., & Blanke, O. (2015). Inference of perceptual priors from path dynamics of passive self-motion. *Journal of Neurophysiology, 113*(5), 1400–1413.

Reddi, B. A., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience, 3*(8), 827–830.

Schweighofer, N., Xiao, Y., Kim, S., Yoshioka, T., Gordon, J., & Osu, R. (2015). Effort, Success, and Non-use Determine Arm Choice. *Journal of Neurophysiology, 114*(1), 551–559.

Stoloff, R. H., Taylor, J. A., Xu, J., Ridderikhoff, A., & Ivry, R. B. (2011). Effect of reinforcement history on hand choice in an unconstrained reaching task. *Frontiers in Neuroscience, 5*, 41.

Abstracts

Proceedings of the Master's Programme Cognitive Neuroscience is a platform for CNS students to publish their Master thesis. Given the number of submissions, we select the articles that received the best reviews, under recommendation of our editors, for the printed edition of the journal. The abstracts of the other articles are provided below, and for interested readers a full version is available on our website: www.ru.nl/master/cns/journal.

Decoding of Concepts Within and Across Semantic Categories

Evgenia Bednaya, Irina Simanova, Marcel A. J. van Gerven

Currently, there is no agreement about how conceptual knowledge is represented in the brain. Studies on semantic dementia suggest that the anterior temporal lobe (ATL) functions as the semantic hub, combining information distributed throughout the brain and constructing abstract semantic representations. However, functional neuroimaging studies argue that the semantic hub is located more posteriorly, suggesting that the posterior medial temporal gyrus (pMTG) might be a better candidate for the semantic hub than the ATL. The goal of this functional magnetic resonance imaging (fMRI) study was to examine the temporal lobe's role in the organisation of semantic knowledge in the brain. We expected to find a posterior-to-anterior gradient of specificity of semantic representations, where more coarse categorical information is processed in the pMTG and more detailed representations are computed in the ATL. We used spoken and written Dutch words representing either basic or subordinate names for categories of animals and buildings (e.g., dog/house, retriever/chihuahua). The combined results from the whole brain standard univariate analysis and searchlight-based multivariate pattern analysis (MVPA) revealed distributed brain activation associated with the semantic processing and within-/across-category discrimination of individual written and spoken words. However, our results were challenged by high variability across subjects, and possible limitations in the experimental design. Therefore, we were not able to obtain an expected activation in the ATL.

Electrophysiology of spoken phrase production: Exploring effects of complexity and switching on the P300

Caitlin Coughler, Ardi Roelofs, Herbert Schriefers, Natalia Shitova

P300 is one of the most studied event-related brain potentials (ERPs) in cognitive neuroscience related to attention and memory. However, little is known about P300 modulations in spoken language production. Previous research outside the language domain has shown that P300 amplitude is sensitive to task complexity and switching. The current study examined the influence of these two factors in P300 amplitude in phrase production through a within-language switching paradigm using differential noun phrase length. Participants switched every second trial between describing pictures with one adjective (size only; short condition) or two adjectives (size and colour; long condition) in the format determiner + adjective(s) + noun. Response times (RTs) were longer for long- than short-phrase trials (length effect). Moreover, length and sequence interacted: RTs were longer on switch than on repeat trials for the short phrases (switch cost) but shorter on switch than on repeat trials for the long phrases (switch benefit). This caused the length effect to be smaller on switch than on repeat trials. The ERPs showed an effect of length but not of sequence, as well as no interaction. The P300 amplitude was greater on long trials than on short trials. Thus, whereas RTs reflected both length and an interaction between length and sequence effects, P300 amplitude reflected only length. These results suggest that P300 in language production is sensitive to task complexity but not to switching, unlike what is found outside the language domain.

Are Visual Representations Phase-Coupled by Alpha Oscillations?

Mats van Es, Tom Marshall, Eelke Spaak, Marcel van Gerven, Ole Jensen

Visual representations in the cortex are thought to be modulated by alpha oscillations. We used decoding techniques to verify this hypothesis. Attentional modulation is characterized by changes in alpha oscillations and it modulates visual processing. Since orientation information can be reliably decoded from neural activity with magnetoencephalography (MEG), this technique can be used to study modulations of visual processing. In this study, we investigated whether attention modulates decoding performance in the alpha band (8–12 Hz), such that decoding performance shows ~10 Hz rhythmicity. Although no conclusive answer was achieved, the study provides insightful data on the used methodology in studying modulation effects on decoding. Moreover, we have shown that the orientation of peripheral, drifting gratings can be decoded from MEG data.

Ketamine on Working Memory: What are the Underlying EEG Correlates?

Constanze von Randow, Wilhelmus H. I. M. Drinkenburg, Gillies van Luijtelaar

Neurodegenerative disorders are associated with a decline in working memory (WM) and thought to be accompanied by dysfunctional connectivity and altered theta/gamma amplitude within the hippocampal-prefrontal circuit (HC-PFC). The hippocampus and prefrontal cortex both have dense populations of N-methyl-D-aspartate (NMDA) receptors. Ketamine, a NMDA receptor antagonist, is of interest as a mechanistic model of glutamatergic dysfunction mimicking cognitive impairments in animal and human studies. In this study, we sought to identify an EEG fingerprint of WM under normal and impaired conditions of functional connectivity and amplitude in the theta and gamma frequency bands. Long-Evans rats received a baseline saline injection followed by an acute and repeated sub-anaesthetic doses of ketamine (10 mg/kg, s.c.) 30 min prior to performing a Delayed-non-match to position task, and amplitude as well as functional phase based connectivity changes were studied in the retrosplenial, frontal association, lateral parietal association and cingulate cortex, with the cingulate serving as the seed. The task showed to be measuring working memory, yet ketamine did not influence performance neither acutely nor after repeated exposure. The EEG did not reveal a specific effect of ketamine on WM either, but we did identify an EEG fingerprint of WM, which showed a dissociation between amplitude and network connectivity for the different brain regions and frequency bands. The main effects of WM were found in the higher theta band in the network, whereas no changes were occurring in the gamma bands. Ketamine did not show an effect in the low theta band possibly owing to compensatory mental effort. Working memory in this study did not show to be impaired by acute or repeated ketamine, which also was reflected in the behavioural data, not necessarily ruling out ketamine as a good model for degenerative diseases, as ketamine effects were visible, but might bear evidence for the task having been learned too well and/or the exposure to ketamine not being sufficient to functionally disrupt the system.

Why is a raven like a writing desk? How insights from parameterized complexity theory predict human analogical reasoning

Colby Tibbets, Tobias Winner, Iris van Rooij, Ivan Toni

We possess the remarkable capacity to identify and understand relational similarities between the constituent parts of disparate wholes. The analogical mapping process underlying this capacity allows us to draw inferences about objects, actions, and events that we see as analogous to one another. This is believed to be a fundamental aspect of intelligence, found in language, creativity, problem solving, and reasoning. A better understanding of how the brain supports the analogical mapping process carries the potential to better understand the domains where it manifests. However, the most well-known model of analogical mapping, called Structure-Mapping Theory, has been shown to be computationally intractable. This is problematic because, assuming that the brain is limited by finite computational resources, brain computation is constrained to be tractable. A solution to this problem has been proposed by van Rooij et al. (2008), who have proven that the computations postulated by SMT are tractable provided that the model parameter o (denoting the number of objects in the analogical match) is relatively small. This proposal yields the prediction that humans can quickly determine good analogical matches in situations where o is small. Moreover, it predicts that performance should deteriorate as o grows because of the inherent intractability of the postulated computations. In this thesis, we set out to test these predictions in a behavioural experiment. Participants were instructed to identify squares that correspond with one another on opposite sides of a divided screen. The results demonstrated that increasing the number of squares resulted in longer response times and less optimal analogical mappings. These findings are consistent with the model's predictions and provide support for the FPT-Cognition thesis.

Grey and White Matter Correlates of Human Place Learning Competence

Sanne Tops, Janneke van Ekert, Nils Müller, Gabriele Janzen

Two different ways of learning have been suggested to support successful navigation. Response learning, through which an organism learns to associate an individual landmark with a goal location, relies primarily on the striatal system. Place learning, on the contrary, relies on the hippocampal system and refers to a type of learning where multiple landmarks are integrated and together form a cognitive map of the environment. The current training study investigates how the grey and white matter correlates of the brain structures involved in place learning, relate to the different strategies. In a virtual environment participants had to pick up objects and then relocate it to their original position. The anatomical and diffusion-weighted scans allowed us to investigate how the striatal and hippocampal system contribute to the different learning strategies. We used voxel-based morphometry and fractional anisotropy to examine grey and white matter differences of the hippocampal and striatal system. The results showed increased grey matter volume of the right caudate nucleus for better baseline performance and bigger training effects when all spatial cues were present (standard condition). White matter anisotropy of the right anterior limb of the internal capsule correlated negatively with training effects in the standard condition. These findings indicate that people biased towards using a response strategy have increased grey matter volume of the right caudate nucleus and also increased white matter in the right anterior limb of the internal capsule. Grey matter volume and white matter anisotropy of the caudate nucleus can predict response strategy use.

The relationship between heart rate variability and the neural mechanisms of freezing in police officers

Naomi de Valk, Mahur Hashemi, Floris Klumpers, Karin Roelofs

Freezing, defined as a complete absence of movement, is a defensive response observed after threat detection in both humans and animals. It is characterized by heart rate deceleration, also called bradycardia, caused by the parasympathetic autonomic nervous system. The periaqueductal grey (PAG) is involved in the neural mechanism underlying freezing behavior. Furthermore, studies in humans suggest that differences in freezing responses may have consequences for the development or maintenance of psychopathology. Resting heart rate variability (HRV) reflects the ability of the parasympathetic nervous system to rapidly modulate heart rate, which is necessary in order to flexibly respond to the constantly changing environment with physiological and emotional arousal. Therefore resting HRV is thought to be an index of someone's emotion regulation capacity and has been linked to psychopathological vulnerability. Here we study the relationships between resting HRV, trait anxiety, and freezing during threat anticipation in police officers using an active shooting paradigm. Replication of the correlation between BOLD responses in the PAG and bradycardia during threat anticipation confirmed the involvement of the PAG in freezing behavior. This is the first study to show a relation between resting HRV and stimulus-induced bradycardia, which suggests a common neural mechanism. Furthermore, BOLD responses in the PAG correlated negatively with reaction time, confirming the link between freezing and action preparation processes. Identifying the link between resting HRV, as a psychopathologic vulnerability factor, and defensive responses such as freezing will shed light on the involvement of differences in defensive responding in the development and maintenance of fear-related psychopathology.

Learning to Predict Others′ Behaviour: Neural Mechanisms of Social-Categorical Knowledge Acquisition and Utilisation

Suhas Hassan Vijayakumar, Egbert Hartstra, Harold Bekkering

As social animals, we are constantly required to speculate about other individuals' behaviour. Given the complexity and highly variable nature of human behaviour, we tend to group people into certain categories from as early as infancy in order to make better predictions about others' behaviour. The ability to reason about others' intentions and mental states is called having a theory of mind, and this network has been studied elaborately. It is further established that we make use of social-categorical knowledge stored in the temporal cortex while making such predictions. However, the exact neural mechanisms underlying prediction of other's behavior based on social-categorical knowledge still remains unknown.

To answer this question, we designed a task in which participants were asked to predict the behaviour of individual agents based on prior social-categorical knowledge that they learn during the experiment. Behavioural findings show that the participants learnt to successfully perform the task progressively better with higher prediction efficiency indices throughout the experiment. While the medial prefrontal cortex (mPFC) and the anterior temporal lobe showed an increase in activation upon acquiring social-categorical information towards the latter half of the experiment, areas like temporoparietal junction (TPJ), superior temporal sulcus (STS), middle occipital gyrus, and fusiform gyrus showed an increase in activation upon witnessing unexpected outcomes. Furthermore, we show that there was increased functional connectivity between mPFC and TPJ, as well as mPFC and STS, while utilizing social-categorical knowledge to predict others' actions. Our study thus outlines the key brain regions involved in behavior prediction by social-categorization. We further discuss our results under the light of hierarchical predictive coding theory, which has been proposed to be a potential candidate to explain the process of mentalising and open up a venue for further exploration.

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/donders/

Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD Nijmegen

P.O. Box 310
6500 AH Nijmegen
http://www.mpi.nl

Radboudumc
Geert Grooteplein-Zuid 10
6525 GA Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.umcn.nl/

Baby Research Center
Montessorilaan 3
6525 HR Nijmegen

P.O. Box 9101
6500 HB Nijmegen
http://www.babyresearchcenter.nl