Eye movement suppression interferes with construction of object-centered spatial reference frames in working memory

Mikkel Wallentin [a,b,*], Line Burholt Kristensen [a,c], Jacob Hedeager Olsen [b], Andreas Højlund Nielsen [a,d]

[a] Center of Functionally Integrative Neuroscience, Aarhus University Hospital, Nørrebrogade, 8000 Aarhus C, Denmark
[b] Center for Semiotics, Aarhus University, Denmark
[c] Department of Scandinavian Studies and Linguistics, University of Copenhagen, Denmark
[d] Department of Anthropology, Archaeology and Linguistics, Aarhus University, Denmark

ARTICLE INFO

ABSTRACT

The brain's frontal eye fields (FEF), responsible for eye movement control, are known to be involved in spatial working memory (WM).

In a previous fMRI experiment (Wallentin, Roepstorff & Burgess, Neuropsychologia, 2008) it was found that FEF activation was primarily related to the formation of an object-centered, rather than egocentric, spatial reference frame. In this behavioral experiment we wanted to demonstrate a causal relationship between eye movement control and manipulation of spatial reference frames. Sixty-two participants recalled either spatial ("Was X in front of Y?") or non-spatial ("Was X darker than Y?") relations in a previously shown image containing two to four objects, each with an intrinsic orientation and unique luminance. During half of all recall trials a moving visual stimulus was presented, which participants had to ignore, thus suppressing eye movement. Response times were significantly slower for spatial relations with distraction while there was no effect on non-spatial relations. There was no effect on accuracy, i.e. WM maintenance. This is consistent with the hypothesis that in spatial representations the FEFs are involved in WM content manipulation, such as establishing an object-centered spatial frame of reference.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Building on work by Baddeley (1986), it has previously been found that voluntary eye movements selectively interfere with spatial working memory (Hale, Myerson, Rhee, Weiss, & Abrams, 1996; Lawrence, Myerson, & Abrams, 2004; Postle & Hamidi, 2007; Postle, Idzikowski, Della Sala, Logie, & Baddeley, 2006). Hale and co-workers (Hale et al., 1996) found that maintenance of short-term memory for locations was disrupted by looking and pointing at irrelevant locations, but not by verbal responses, mental rotation, or color- or shape-discrimination processes. Lawrence and co-workers (Lawrence et al., 2004) found that both eye movements and attention shifts in the absence of eye movements interfered with memory for spatial location. Postle and co-workers in two papers (Postle & Hamidi, 2007; Postle et al., 2006) reported data from participants performing delayed recognition of either the identity or location of geometric shapes. Voluntary eye movements during the retention interval were found to interfere

with the memory of where the object was, but not the memory for shape. One feature of these studies was that it was conducted in complete darkness. Locating the object in space thus entirely depended on mapping it according to an egocentric coordinate system whereas most natural spatial processes involve some kind of mapping of the relative locations of the observer and two or more objects.

The present experiment aimed at investigating the relationship between the brain's system for eye movement control and working memory for spatial relations that are not egocentric in nature, i.e. independent of retinal or head centered coordinate systems. These non-egocentric relations are often referred to as "allocentric" (e.g. Burgess, 2006; O'Keefe & Nadel, 1978). However, they can be broken down further into absolute coordinate systems (e.g. north/south) and coordinate systems related to the relations between objects (e.g. "in front of"), where one object has some intrinsic property (e.g. orientation/line of sight) anchoring the spatial relation (e.g. Kemmerer, 2006; Kessler & Rutherford, 2010; Levinson, 2003). The present experiment only addresses the object-centered version of these two allocentric coordinate systems.

Neuroimaging experiments have shown that spatial working memory processing involves a distributed fronto-parietal network

* Corresponding author at: Center for Semiotics, Aarhus University, Denmark.
Fax: +45 89494400.
E-mail address: mikkel@cfin.au.dk (M. Wallentin).

of brain regions including the posterior parietal cortex and frontal eye fields (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Wallentin et al., 2008; Wallentin, Roepstorff, Glover, & Burgess, 2006; Wallentin, Weed, Østergaard, Mouridsen, & Roepstorff, 2008). Neuroimaging studies (Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000) and studies using transcranial magnetic stimulation (Hamidi, Tononi, & Postle, 2008) have suggested that frontal brain regions may be more involved in manipulation tasks in working memory whereas parietal regions may be more involved in maintenance tasks.

Spatial working memory (Baddeley & Hitch, 1974) and imagery supporting it may be inherently egocentric in nature (Byrne, Becker, & Burgess, 2007; Milner, Djickerman, & Carey, 1999; Schenk, 2006). Accessing information from an object-based allocentric representation therefore involves adopting a different "view point" or frame of reference (Coventry & Garrod, 2004; Kemmerer, 2006; Levinson, 2003) and thus transforming the representation into an imaginary egocentric framework. This supposedly relies on imaginary movement, similar to that found in studies of mental object rotation, where it has been shown that the time it takes to perform the rotation increases linearly with the size of the rotated angle (Shepard & Metzler, 1971). Investigating a scene relative to an object-centered frame of reference, however, is probably accomplished by several strategies, depending on the task, involving not only imaginary object rotation, but also imaginary self rotation and motion (Kessler & Rutherford, 2010; Kessler & Thomson, 2010; Michelon & Zacks, 2006). Michelon and Zacks (2006) found that the increase in RT was observed with increasing angle when participants were asked to judge whether an object was left or right of the new view point, but not when participants were asked to judge whether the object was visible from the new view point, suggesting that there might be two independent processes at play. These findings were replicated by Kessler and co-workers in two papers (Kessler & Rutherford, 2010; Kessler & Thomson, 2010), but with the extension that for the left–right judgments an interaction with the bodily position of the participant was observed so that RT was further delayed if the body position was incongruent with the new viewpoint, suggesting that the notion of "left" and "right" is coded in coordinates relative to the body. The same was not true for judgments related to whether two objects were "in front of" or "behind" one-another as seen from the new perspective, suggesting that this is not necessarily coded in coordinates relative to the body. These findings seem to suggest that adopting a perspective other than the egocentric could take at least two forms, one based on a simulated body rotation/translation into the target perspective, resulting in body-centered object locations that subserve left/right judgments and one relying on simulating a line-of-sight relations between the reference perceptive and the target object. One hypothesis could be that the latter process is coded in eye-centered coordinates.

The frontal eye fields (FEFs) are known to be involved in eye movements and in both overt and covert shifting of spatial attention to present stimuli (Corbetta & Shulman, 2002; Greenberg, Esterman, Wilson, Serences, & Yantis, 2010; Moore & Fallah, 2004; Pessoa, Kastner, & Ungerleider, 2003) and during the preparation of anti-saccades, i.e. when you are preparing to suppress the natural tendency to move your gaze towards an appearing target (Connolly, Goodale, Menon, & Munoz, 2002; Munoz & Everling, 2004). The FEFs are thought to contain a map of visual space in oculomotor coordinates (Bruce, Goldberg, Bushnell, & Stanton, 1985), and lesions to the FEFs cause impairments of voluntary oculomotor control, revealed as an inability to suppress reflexive eye movements towards a distracting stimulus (Milner, 1982; Pierrot-Deseilligny, Rivaud, Gaymard, & Agid, 1991).

Few neuroimaging experiments have studied changes in spatial frames of reference (David et al., 2006; Vogeley et al., 2004; Wallentin et al., 2008). In a recent BOLD-fMRI experiment Wallentin and co-workers (2008) investigated working memory (WM) probed by linguistic cues for object-centered and egocentric spatial relations in a visual scene. Participants were shown images depicting three characters, denoted "he", "she" and "it", and were subsequently probed about their spatial and non-spatial relations (e.g. "Was he in front of her?"/"Was he older than her?") using either an allocentric, object-centered reference frame or an egocentric reference frame (e.g. "Was he in front of you?"/"Was he older than you?"). Allocentric spatial recall for the "in front of" relationship was found to yield an increased BOLD activation in a network of brain regions including the FEFs compared to recall of both non-spatial and egocentric spatial image content. Furthermore, this difference was modulated by performance, such that good performers exhibited greater activation in FEFs than poor performers. The authors suggest that these effects may be due to a beneficial frontal eye field involvement during spatial recall caused by participants moving their internal gaze during relocation of reference frames in representational space.

BOLD fMRI, however, is a strictly correlational measure. A stronger support for the hypothesis that the FEFs are involved in the construction of spatial reference frames would involve actively interfering with the system. We therefore conducted a dual task interference experiment. Similar to the Wallentin et al. (2008) study, we asked participants to remember both spatial and non-spatial object-centered relations within an image and subsequently probed the working memory content by asking questions. Similar to the previous study the recall of spatial relations was assumed to involve the FEFs. To further manipulate the activity of the FEFs a visual distractor was added, designed to evoke involuntary eye movements. This was displayed coinciding with the recall questions. Studies examining the interplay between voluntary and involuntary saccades have found that an abrupt onset of an irrelevant visual stimulus causes a reliable distraction effect (Irwin, Colcombe, Kramer, & Hahn, 2000; Ludwig & Gilchrist, 2002; Schreij, Owens, & Theeuwes, 2008; Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999). On 20–50% of such trials distractors cause an involuntary saccade towards the distractor (Godijn & Theeuwes, 2002; Irwin et al., 2000; Ludwig & Gilchrist, 2002), although some habituation has also been observed (Godijn & Kramer, 2008). Motion and motion onset has also been shown to be a potent capturer of attention in object detection paradigms (Abrams & Christ, 2003; Al-Aidroos, Guo, & Pratt, 2010; Franconeri & Simons, 2005; Guo, Abrams, Moscovitch, & Pratt, 2010). Together, these results suggest that a distractor consisting of an appearing object that moves fast, disappears and reappears on a new location should act as a strong elicitor of involuntary eye movements that need to be suppressed during fixation. Participants in the present experiment were asked to ignore this distractor under the assumption that this would force participants to perform voluntary eye movement suppression, which again would involve the FEFs (see above).

We hypothesized that if, on the one hand, the system for suppressing eye movements is also involved in maintaining spatial working memory content, then we would expect a decline in accuracy as a function of interfering with the system and we would expect this effect to increase with an increase in working memory load. If, on the other hand, this system is primarily involved in the manipulative reconstruction of the memory content during inquiry, such as relocating the frame of reference, then we would only expect a delay of spatial processing observable as increased response time. Since this process involves the same manipulation and the same number of objects regardless of how many there may be stored in WM, it may be relatively independent of WM load.

2. Methods and materials

2.1. Participants

62 participants took part in the experiment, 30/32 male/female, 59/3 right/left handed; median age: 24 years, range 19–58 years; median years of education: 17, range 11–22 years. Participants were kept naïve to the detailed hypotheses of the study, but instructed about their rights as volunteers prior to testing and reminded that they could leave the experiment at any given point in time. They received 100 DKK for participation.

2.2. Stimuli and procedure

Participants were instructed that during the experiment they were to remember the configuration of a number of "mice" (see Fig. 1). There were four different mice made from simple geometrical figures (a circle, a square, a triangle and a pentagon), each equipped with two eyes and a snout (see Fig. 1). Each mouse was denoted by one of the four Danish third person personal pronouns. The circle was called "Hun" [she], the square was called "Han" [he], the triangle was called "Den" [it1 – third person common gender] and the pentagon was called "Det" [it2 – third person neuter]. Participants were informed about these names and instructed that they were merely conventional and used because they, due to the multi-referential ability of pronouns (e.g. the pronoun "she" is not only used to refer to female individuals but also to inanimate objects, such as ships), were hypothesized to yield as little interference with memory processing as possible. The mice differed from round to round in luminance (different shades of gray), but participants were instructed that this had no effect on their names.

The experiment consisted of 90 trial rounds. Each experimental round contained an encoding phase and two recall trials. During
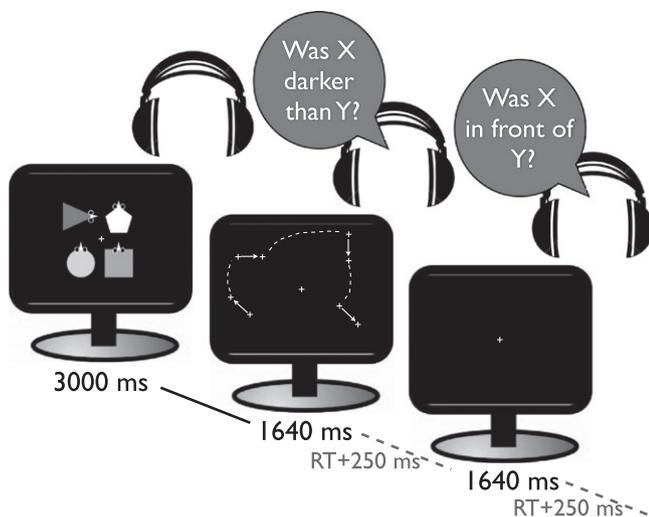

Fig. 1. The experiment had 90 rounds. A round in the experiment consisted of an encoding phase (3000 ms) and two recall trials, a spatial and a non-spatial (order randomized). During encoding participants saw two, three or four "mice" named "Han" ("He" – square), "Hun" ("She" – circle), "Den" ("It1"– common gender" – triangle) and "Det" ("It2" – neuter" – pentagon). The mice differed in position and luminance from round to round and the task was to remember both the spatial, (e.g. "Was he in front of her?" – correct response: "no"), and the non-spatial (e.g. "Was he darker than her?" – correct response "yes") relations. Questions lasted 1640 ms and the first was presented immediately after the offset of the image. The second question was presented 250 ms after the participant's response. "Yes/No" responses were given by button presses using the right index or middle finger. Coinciding with half of the questions (order randomized) a distractor was presented on the screen. The distractor moved, then jumped to a new location and moved again. Participants were asked simply to try to ignore the distractor and keep their gaze focused on the screen.

encoding, an image depicting two, three or four mice was displayed for 3000 ms. The task was to remember both the spatial relations among the mice and their relations in terms of differences in luminance (see Fig. 1). Two questions were asked immediately after each image had been displayed, a spatial question which always had the form "Var X foran Y?" [Was X in front of Y?] and a non-spatial question which always had the form "Var X mørkere end Y?" [Was X darker than Y?]. All questions were played through headphones so as not to interfere with visual processing. The order of questions was randomized across rounds. Participants were instructed to respond as fast and accurately as possible.

Each mouse had a size of approximately 120×120 pixels in all stimulus images and the mouse configurations were placed within the central 300×300 pixel quadrant of the screen. A mouse could be oriented either upwards, downwards, leftwards or rightwards. Participants were instructed that "in front of" was defined as being located immediately in front of the snout of another mouse. In the example in Fig. 1 "Den" (triangle) is in front of "Hun" (circle) whereas "Det" (pentagon) is not in front of "Hun".

30 rounds probed an image with two mice, 30 rounds had three mice and 30 rounds had four mice (Fig. 1 depicts an example with four mice), i.e. there were three levels of increasing working memory load.

All questions were recorded in mono (22,050 Hz sample rate, 16 Bit) and edited in Sony Sound Forge 8.0 so that all questions had a duration of 1640 ms. All questions consisted of the template questions "Was X in front of Y?" and "Was X darker than Y?" with the personal pronouns (He/Him, She/Her, It1, It2) copy/pasted from other recordings (after correcting for small differences in duration) in order to match everything in the sounds except the pronouns.

Participants responded with a button press (right index finger for "yes" and right middle finger for "no") using the computer keyboard with a green tag indicating the key used for "yes" and a red tag indicating the key used for "no". Half of all trials required a "yes" response and half required a "no" response. During half of all trials a distractor was displayed coinciding with the recall question. The order and question/distractor pairings were randomized within and between participants with the constraint that half of each of the two question types had to be coupled with a distractor for each participant.

During distraction trials the distractor was displayed for 1800 ms. It consisted of a moving "+" (Font "Times", size 55). During the on-period the distractor moved across the screen in steps of 30 pixels per 20 ms. It moved five steps in one direction (up, down, left, right or along diagonals) and then jumped to a new position on the screen where it started a new movement. Starting positions were all placed outside the region of the screen where the stimulus was presented (i.e. outside the central 300×300 pixel quadrant). If the distractor reached the border of the screen it changed direction and moved back. The resulting impression was that of fast and erratic movement. Participants were instructed to keep their eyes focused on the fixation cross in the middle of the screen and simply ignore the distractor, but to refrain from doing so by closing their eyes, moving their gaze and/or by "zooming out", i.e. moving the point of visual focus to a point behind (or in front of) the screen. Ignoring the distractor was thought to involve active suppression of involuntary eye movements, which again was thought to involve the FEFs (see introduction).

37 participants were tested using a laptop computer with a 14" screen in a quiet room. The rest were tested in groups of 3–8 people in a computer room with standard Windows PCs with 17" screens (see below for an analysis of data including test site as covariate). Participants in the group sessions were asked to remain quiet if they finished the test before their fellow participants. For all participants, the experiment was conducted using a 1024×768 pixel screen resolution. No head holder was used but

Table 1
Descriptive statistics for all conditions, e.g. Space2: spatial questions probing images with 2 elements, NSpace4: non-spatial question probing image with 4 objects, etc.

	Space2	Space3	Space4	NSpace2	NSpace3	NSpace4
Percentage correct (s.e.)						
No Distraction	93.5 (1.0)	87.5 (1.3)	80.7 (1.3)	94.8 (0.9)	86.8 (1.2)	75.3 (1.3)
Distraction	94.5 (0.8)	86.9 (1.2)	77.9 (1.3)	94.3 (0.9)	86.7 (1.2)	74.5 (1.5)
Difference	−1.0 (0.8)	0.5 (1.0)	2.7 (1.3)	0.6 (0.8)	0.1 (1.1)	0.8 (1.4)
Response time – ms (s.e.)						
No Distraction	2109 (40)	2507 (55)	2864 (57)	2091 (30)	2539 (46)	2833 (58)
Distraction	2161 (36)	2599 (58)	2931 (76)	2116 (38)	2549 (42)	2809 (59)
Difference	−52 (26)	−91 (38)	−68 (36)	−25 (26)	−10 (26)	24 (33)

participants were placed with a distance of approximately 50 cm from the screen and asked to remain in that position during the whole experiment. At this distance the 300 × 300 pixel region containing the stimuli subtended 9.4° on the 14" screen and 11.7° on the 17" screen. Stimuli and responses were shown and logged using Cogent 2000 executed in MatLab.

Before the main experiment, participant completed a computerized training session in which they were shown images of the mice one at a time and presented with a "name" through headphones. They then responded with a "yes" or "no" button press depending on whether they believed the name was correct or not. Upon each response they were given feedback. This procedure continued until the participant was capable of answering correctly on 12 training trials in a row. Participants completed at least two of these training sessions and were encouraged to continue with more if they were not yet comfortable with the names of the mice.

In summary, the experiment had a 2 × 2 × 3 factorial design: the question could be either spatial or non-spatial (two levels), either coupled with a visual distractor or not (two levels), and there were three levels of WM load, i.e. remembering the relations between two, three or four mice. Each participant received 30 trials of each condition during two experimental sessions lasting approximately 40 min in total.

Repeated measures analyses of variance (ANOVA) were conducted in MatLab for both accuracy and response time using Greenhouse-Geisser correction of degrees of freedom and all reported P-values were Bonferroni corrected for multiple comparisons. Initial analyses included test-site as a between-participants covariate, but since this did not yield any effect, the covariate was removed. One-sample t-tests (reporting uncorrected P-values) were used to explore the effects found.

3. Results

Table 1 displays mean accuracy for each condition as well as means of the median response time latencies.

A 2 × 2 × 3 repeated measures ANOVA conducted on median response time measures for each condition and participant yielded no effect of space/non-space question type ($F(1,61) = 3.36$, n.s.), a main effect of distractor ($F(1,61) = 8.3$, $P < 0.04$), a main effect of WM load ($F(1.3, 79.8) = 294.1$, $P < 0.0007$) and a space/non-space question type × distractor interaction ($F(1,61) = 9.36$, $P < 0.03$). No other interactions reached significance. The main effect of distractor was due to response latencies being significantly greater when the distractor was present than when it was absent ($t(61) = −2.88$, $P < 0.006$), however, this effect was driven solely by the spatial questions where there was a significant effect of the distractor when collapsing across the WM load conditions ($t(61) = −4.1$, $P < 0.0002$). This was not the case for the non-spatial questions ($t(61) = −0.21$, n.s.). This difference revealed itself in the interaction between question type and distractor condition reported above. Participants were significantly more interrupted by

the distractor with the spatial questions than with the non-spatial questions ($t(61) = 3.06$, $P < 0.005$) – see Fig. 2(top).

Another 2 × 2 × 3 repeated measures ANOVA was conducted on condition-specific intra-individual response time variability measures (standard deviations). This revealed a main effect of WM load ($F(1.9, 114.3) = 28.1$, $P < 0.0001$). No other effects or interactions were significant at the Bonferroni corrected level ($P > 0.05$ for all tests). This means that the distractor effects on the spatial recall task observed using median response times were not accompanied by significant changes in the distribution of response times.
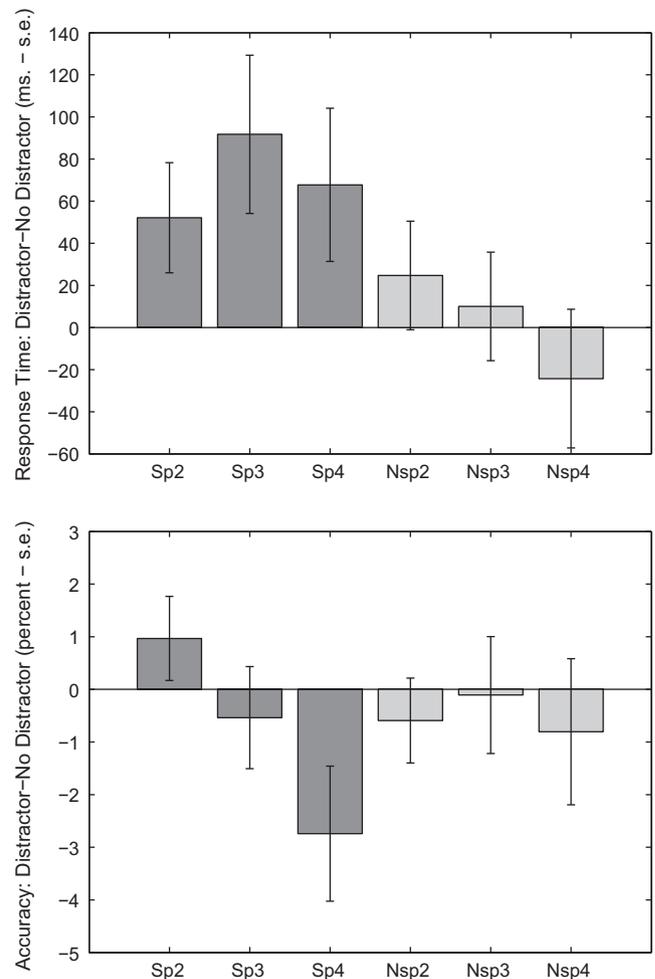

Fig. 2. TOP: Response time differences in milliseconds, BOTTOM: Accuracy differences in percent, between Distraction and No Distraction tasks for spatial and non-spatial questions across the three working memory load levels, e.g. Sp2: spatial questions probing images with 2 elements, NSp4: non-spatial question probing image with 4 objects, etc.

Participant accuracy for each condition was also investigated using a 3-way ANOVA. This showed no effect of space/non-space question type ($F(1,61) = 2.29$, n.s.), no effect of distractor ($F(1,61) = 4.4$, n.s.), a significant effect of WM load ($F(1.86,113.24) = 273.62$, $P < 0.0001$), and a significant interaction between question type and WM load ($F(1.81,110.56) = 13.06$, $P < 0.0002$). No other interactions reached significance. As can be seen from Table 1, the interaction effect was due to lower accuracy for the non-spatial question in the four items condition (74.9%), compared to the spatial question with four items (79.3%). A paired t-test comparing the spatial and non-spatial four items conditions shows this (data collapsed across the distractor/no distractor conditions; $t(61) = 4.7$, $P < 0.001$), while both the two items and three items conditions yielded non-significant differences across the two question types ($t(61)<1$, n.s. for both comparisons).

4. Discussion

In this experiment we investigated whether a secondary eye movement suppression task interfered with a spatial working memory task involving object-centered reference frames. We found that participants were significantly slower during recall of spatial aspects of an image in the presence of the distractor task whereas this was not the case for non-spatial recall.

Accuracy was not affected by the distractor. One interpretation of this is that suppression of eye movements, contrary to real eye movements, does not interfere with the memory representation as such, but rather interferes with the process of manipulating the representation with the purpose of recalling specific aspects of the spatial configuration. This would support the hypothesis that brain regions involved in eye movement control are also involved in establishing a spatial frame of reference in object-centered allocentric representational coordinates (Wallentin et al., 2008) whereas maintenance may be more related to other brain regions, perhaps located in posterior parietal cortex (Wallentin et al., 2006). An alternative interpretation would be that the interference task simply was not strong enough to elicit measurable accuracy deficits. Individual levels of eye movement suppression by definition could not be measured in the present experiment with external means, since the outcome of suppression is a lack of observable movement. But an interpretation saying that the lack of an accuracy effect was due to weak interference would not be inconsistent with the hypothesis that the frontal eye fields are also involved in the establishment of frame of reference. In favor of this interpretation is the post hoc observation that when looking at the 4-object images alone, the distractor does seem to have an effect on spatial accuracy (paired t-test: $t(61) = 2.13$, $P < 0.05$) – see Fig. 2(bottom). This, however, is a biased observation and thus can only be used to inform future research.

The eye movement suppression effect on spatial recall was not caused by the act of simply comparing two or more objects in working memory, e.g. by an imaginary shift of gaze back and forth between the objects. If so, then we would have expected an equal effect on the non-spatial trials, since these were equally relational.

The effect was independent of task difficulty since it could be seen across the different levels of working memory load and even though the 4-object, non-spatial questions unexpectedly turned out to be more difficult than the 4-object, spatial questions, this did not cause any interference to occur. We tentatively interpret the difference in difficulty between the 4-object conditions as reflecting the possibility of grouping the spatial relations. In some of the 4 objects images two or more mice were placed in parallel positions. This might have allowed for chunking strategies to lift

some of the working memory load on the spatial task whereas the number of independent relations was constant across the four object trials for the luminance task.

The reported result of probing spatial working memory using linguistic cues is consistent with a broader framework of embodied cognition (Barsalou, 2008) where knowledge of goals can be studied using eye movement paradigms (Eshuis, Coventry, & Vulchanova, 2009) and where perception and language constantly interact in a predictive manner (Richardson, Dale, & Kirkham, 2007; Spivey, Tyler, Eberhard, & Tanenhaus, 2001; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) and where linguistic information is also partly remembered using spatial indexing in terms of ocular coordinates, i.e. when we remember what was said partly by remembering where it was said (Richardson & Spivey, 2000). It is compatible with the hypothesis that multiple frames of reference may be at play in more complex spatial reasoning, such as projecting the egocentric notions of "left" and "right" out onto an allocentric frame of reference (Kessler & Rutherford, 2010; Kessler & Thomson, 2010; Michelon & Zacks, 2006). Left and right may be coded in body-centered coordinates while probing intrinsic front/back relations, as seen in this experiment, may to some extent rely on eye-centered coordinates.

5. Conclusion

In line with previous neuroimaging results suggesting a role for the brain's frontal eye fields in establishing an object-centered allocentric frame of reference (Wallentin et al., 2008) we found that eye movement suppression selectively interfered with processing in allocentric spatial working memory. This happened at the level of manipulation (RT) rather than maintenance (accuracy). We suggest that this is because the FEFs play an active role in establishing a new object-centered spatial frame of reference.

References

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science, 14*(5), 427–432.

Al-Aidroos, N., Guo, R. M., & Pratt, J. (2010). You can't stop new motion: Attentional capture despite a control set for colour. *Visual Cognition, 18*(6), 859–880.

Baddeley, A. D. (1986). *Working Memory*. Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.). *Recent advances in learning and motivation* (vol. 8, pp. 47–90). New York: Academic Press.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Bruce, C. J., Goldberg, M. E., Bushnell, M. C., & Stanton, G. B. (1985). Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. *Journal of Neurophysiology, 54*(3), 714–734.

Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences, 10*(12), 551–557.

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psychological Review, 114*(2), 340–375.

Connolly, J. D., Goodale, M. A., Menon, R. S., & Munoz, D. P. (2002). Human fMRI evidence for the neural correlates of preparatory set. *Nature Neuroscience, 5*(12), 1345–1352.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience, 3*(3), 201–215.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerleider, L. G., & Haxby, J. V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science, 279*(5355), 1347–1351.

Coventry, K. R., & Garrod, S. C. (2004). *Saying, seeing, and acting: The psychological semantics of spatial prepositions*. Hove, UK: Psychology Press.

David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., et al. (2006). Neural representations of self versus other: Visual-spatial perspective taking and agency in a virtual ball-tossing game. *Journal of Cognitive Neuroscience, 18*(6), 898–910.

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychological Science. A Journal of the American Psychological Society/APS, 20*(4), 438–440.

Franconeri, S. L., & Simons, D. J. (2005). The dynamic events that capture visual attention: A reply to Abrams and Christ (2005). *Perception and Psychophysics, 67*(6), 962–966.

Godijn, R., & Kramer, A. (2008). Oculomotor capture by surprising onsets. *Visual Cognition, 16*(2–3), 279–289.

Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1039–1054.

Greenberg, A., Esterman, M., Wilson, D., Serences, J., & Yantis, S. (2010). Control of spatial and feature-based attention in frontoparietal cortex. *Journal of Neuroscience, 30*(43), 14330.

Guo, R. M., Abrams, R. A., Moscovitch, M., & Pratt, J. (2010). Isoluminant motion onset captures attention. *Attention Perception and Psychophysics, 72*(5), 1311–1316.

Hale, S., Myerson, J., Rhee, S. H., Weiss, C. S., & Abrams, R. A. (1996). Selective interference with the maintenance of location information in working memory. *Neuropsychology, 10*(2), 228–240.

Hamidi, M., Tononi, G., & Postle, B. R. (2008). Evaluating frontal and parietal contributions to spatial working memory with repetitive transcranial magnetic stimulation. *Brain Research, 1230*, 202–210.

Irwin, D. E., Colcombe, A. M., Kramer, A. F., & Hahn, S. (2000). Attentional and oculomotor capture by onset, luminance and color singletons. *Vision Research, 40*(10–12), 1443–1458.

Kemmerer, D. (2006). The semantics of space. Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia, 44*(9), 1607–1621.

Kessler, K., & Rutherford, H. (2010). The two forms of visuo-spatial perspective taking are differently embodied and subserve different spatial prepositions. *Frontiers in Psychology, 1*(213), 1–12.

Kessler, K., & Thomson, L. A. (2010). The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference. *Cognition, 114*(1), 72–88.

Lawrence, B. M., Myerson, J., & Abrams, R. A. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin and Review, 11*(3), 488–494.

Levinson, S. C. (2003). *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Ludwig, C. J. H., & Gilchrist, I. D. (2002). Stimulus-driven and goal-driven control over visual selection. *Journal of Experimental Psychology Human Perception and Performance, 28*(4), 902–912.

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. *Perception and Psychophysics, 68*(2), 327–337.

Milner, B. (1982). Some cognitive effects of frontal-lobe lesions in man. *Philos Transactions of the Royal Society of London, B Biological Science, 298*(1089), 211–226.

Milner, A. D., Djickerman, H. C., & Carey, D. P. (1999). Visuospatial processing in a pure case of visual-form agnosia. In N. Burgess, K. J. Jeffery, & J. O'Keefe (Eds.), *Spatial functions of the hippocampal formation and the parietal cortex* (pp. 442–466). Oxford: Oxford University Press.

Moore, T., & Fallah, M. (2004). Microstimulation of the frontal eye field and its effects on covert spatial attention. *Journal of Neurophysiology, 91*(1), 152–162.

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews. Neuroscience, 5*(3), 218–228.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. *Journal of Neuroscience, 23*(10), 3990–3998.

Pierrot-Deseilligny, C., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of reflexive visually-guided saccades. *Brain, 114*(Pt 3), 1473–1485.

Postle, B. R., & Hamidi, M. (2007). Nonvisual codes and nonvisual brain areas support visual working memory. *Cerebral Cortex, 17*(9), 2151–2162.

Postle, B., Idzikowski, C., Della Sala, S., Logie, R., & Baddeley, A. (2006). The selective disruption of spatial working memory by eye movements. *The Quarterly Journal of Experimental Psychology, 59*(1), 100–120.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychological Science. A Journal of the American Psychological Society/APS, 18*(5), 407–413.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: Looking at things that aren't there anymore. *Cognition, 76*(3), 269–295.

Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. J., & Passingham, R. E. (2000). The prefrontal cortex: Response selection or maintenance within working memory? *Science, 288*(5471), 1656–1660.

Schenk, T. (2006). An allocentric rather than perceptual deficit in patient D.F. *Nature Neuroscience, 9*(11), 1369–1370.

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Perception and Psychophysics, 70*(2), 208–218.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(972), 701–703.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science. A Journal of the American Psychological Society/APS, 12*(4), 282–286.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology Human Perception and Performance, 25*(6), 1595–1608.

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., & Fink, G. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *Journal of Cognitive Neuroscience, 16*(5), 817–827.

Wallentin, M., Roepstorff, A., & Burgess, N. (2008). Frontal eye fields involved in shifting frame of reference within working memory for scenes. *Neuropsychologia, 46*(2), 399–408.

Wallentin, M., Roepstorff, A., Glover, R., & Burgess, N. (2006). Parallel memory systems for talking about location and age in precuneus, caudate and Broca's region. *Neuroimage, 32*(4), 1850–1864.

Wallentin, M., Weed, E., Østergaard, L., Mouridsen, K., & Roepstorff, A. (2008). Accessing the mental space-spatial working memory processes for language and vision overlap in precuneus. *Human Brain Mapping, 29*(5), 524–532.