Visual working memory representations bias attention more when they are the target of an action plan

Caterina Trentin *, Heleen A. Slagter, Christian N.L. Olivers

Institute for Brain and Behavior Amsterdam, Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Visual working memory
Action-oriented cognition
Action planning
Attentional biases
Prospective memory task
Eye-tracking

ABSTRACT

Attention has frequently been regarded as an emergent property of linking sensory representations to action plans. It has recently been proposed that similar mechanisms may operate within visual working memory (VWM), such that linking an object in VWM to an action plan strengthens its sensory memory representation, which then expresses as an attentional bias. Here we directly tested this hypothesis by comparing attentional biases induced by VWM representations which were the target of a future action, to those induced by VWM representations that were equally task-relevant, but not the direct target of action. We predicted that the first condition would result in a more prioritized memory state and hence stronger attentional biases. Specifically, participants memorized a geometric shape for a subsequent memory test. At test, in case of a match, participants either had to perform a grip movement on the matching object (action condition), or perform the same movement, but on an unrelated object (control condition). To assess any attentional biases, during the delay period between memorandum and test, participants performed a visual selection task in which either the target was surrounded by the memorized shape (congruent trials) or a distractor (incongruent trials). Eye movements were measured as a proxy for attentional priority. We found a significant interaction for saccade latencies between memorandum and test, participants performed a visual selection task in which either the target was

1. Introduction

Visual working memory (VWM) refers to the capacity of temporarily retaining visual information that is relevant to the task at hand. Much of VWM research has focused on how much and which aspects of the presented information are being remembered (e.g., for a review, see Luck & Vogel, 2013). Relatively few empirical studies have looked at how the eventual purpose for which the information is remembered affects VWM representations (Bae & Luck, 2018; Gunseli, Meeter, & Olivers, 2014; Myers, Stokes, & Nobre, 2017; van Driel, Gunseli, Meeter and Olivers, 2017). Here we investigate how a prospective goal, namely an action plan, modulates VWM.

It has long been argued that cognition is action-oriented and that the study of cognitive processes cannot prescind from considering bidirectional links with action control (Varela, Thompson, & Rosch, 1992; Clark, 1995; O’Regan & Noé, 2001; Engel, Maye, Kurthen, & König, 2013; Engel, Friston, Kragic, & Eds.), 2015; Cisek, 2019; van Ede, 2020). Indeed, over the past few years, an increasing number of studies suggests that the planning of future actions affects cognitive operations such as perception, working memory, and attention (Boettcher, Gresch, Nobre, & van Ede, 2021; Heuer, Crawford, & Schubö, 2017; Wykowska, Schubö, & Hommel, 2009). But already in the 1980s, Allport and Neumann advanced the idea that visual attention is fundamentally action-oriented (Allport, 1987; Neumann, 1987). Within their Selection-for-Action (Allport, 1987) and Functional View of Attention (Neumann, 1987) frameworks, they proposed that visual selective attention reflects the restricted range of targeted actions that one can perform on different objects, rather than some undefined limitation in cognitive resources. That is, there is a constraint in the mapping of sensory inputs to motor outputs (in what Wu (2011) calls ‘behavioural space’). The tight link between attention and action has since been recognized by multiple theoretical frameworks (e.g., Cisek, 2007; Cisek & Kalaska, 2010; Desimone & Duncan, 1995; Hommel, 1999; Hommel et al., 2019; Hommel, Müßeler, Aschersleben, & Prinz, 2001; Klein, 1980), some of which even claim that the attentional control and the action control systems completely overlap (Rizzolatti, Riggio, D’Ascoli, & Umiltà, 2020).
1987). Although these theories differ substantially in some aspects, they all conceptualize attention as an emergent property resulting from the coupling of sensory and motor representations (see also Krauzlis, Bol-limunta, Arcizet, & Wang, 2014). In these frameworks, planning or executing an action towards a visual stimulus means attending to it, which, as a consequence, enhances its perceptual representation. Experimental studies provide support for this notion. For example, the planning of movements towards a visual stimulus location biases attention towards it and enhances perception, a result that has been consistently found across many studies (Baldauf & Deubel, 2008; Han-nings, Jonkaitis, Deubel, & Szinte, 2016; Hoffman & Subramaniam, 1995; Ohl & Rols, 2017; Ohl & Rols, 2020; Rolfs & Carrasco, 2012), not just for eye-movements, but also for hand movements (i.e., grasping, pointing; Baldauf, Wolf, & Deubel, 2006; Baldauf & Deubel, 2010; Hanning & Deubel, 2018). In addition, for hand movements, the observed perceptual biases appear to favour the processing of visual features that are relevant for the upcoming action, for example, planning a grasping movement enhances memory for size, which is important to calibrate one’s hand aperture (Biosco, Daniele, & Fattori, 2017; Craigh-ero, Fadiga, Rizzolatti, & Umiltà, 1999; Fagioli, Hommel, & Schubotz, 2007; Guttingel, Kenemans, & Naggers, 2011; Guttingel, Park, Kenemans, & Naggers, 2013; Job, van Velzen, & de Fockert, 2017; Rolfs, Lawrence, & Carrasco, 2013; Song & Guo, 2019; Symes, Tucker, Ellis, Vainio, & Ottoboni, 2008; White, Kerzel, & Gegenfurtner, 2006).

Recently Olivers and Roelfsema (2020) proposed that also within VWM, attention emerges from sensory to motor mapping. There is now a vast body of research showing functional overlap between VWM and attention (for reviews see Lepsien & Nobre, 2006; Olivers, 2008; Chun, 2011). In addition, researchers have become increasingly interested in how VWM relates to prospective goals, including action plans (for reviews see Myers et al., 2017; van Ede, 2020). However, the link between VWM, attention, and action remains elusive. Olivers and Roelfsema (2020) hypothesized that sensory representations as maintained within VWM become strengthened by long-range recurrent connections between visual and motor areas, as the sensory information provides the necessary parameters for programming an action. As a consequence, performance for the strengthened items will improve over other items in memory – an improvement that is then typically interpreted as reflecting a “prioritized” or “attended” state.

In the present study we sought to investigate whether planning an action that is associated with an object in working memory indeed prioritizes that representation, as compared to when no direct action is associated with the memorized object. A number of previous studies have looked at the link between VWM and action. Within the spatial domain, Ohl and Rolfs (2018) and Heuer et al. (2017) presented participants with a memory array of stimuli of different orientation, located on eight possible placeholders. Right after the disappearance of the stimuli, and before they were tested on their memory, participants were cued to program and execute a saccade (Ohl & Rolfs, 2018) or a pointing movement (Heuer et al., 2017) towards the location of one of the eight placeholders, which were by then empty. Importantly, the cued location for the action was not predictive of which memory item would eventually be probed (and which participants responded to by means of a keypress). Yet, memory for items which shared their location with the action goal was better than for items that were presented at other locations, indicating that executing an action towards a location that is associated with a memory subsequently enhances that memory, as expected in a subsequent memory test.

To our knowledge only one experiment has looked at the effect of action plans on visual working memory and how such plans may already shape memory before the action is executed. Heuer and Schubo (2017, Experiment 2) asked participants to memorize an array of multiple stimuli of different size and color and then, in a subsequent memory test, indicate whether a probed stimulus had changed in either feature size or color. Before their memory was tested, participants were prompted to plan a pointing or a grasping movement, which was to be performed at the end of the trial, after the memory test, and regardless of whether the probe matched the memorized stimulus or not. The authors found that memory for size was better when observers planned a grasping rather than a pointing movement, while memory for color was more (although not significantly) accurate when observers planned a pointing movement, indicating that different features are relevant for different actions (e.g., grasping requires size information but not color, while color might have more use for pointing, though see Feldmann-Wüstefeld & Schubo, 2015, for inconsistent findings). While this study is informative in that it shows that memory representations may become specifically tuned for specific actions, it does not provide information on what planning an action per se does with working memory representations, compared to when such representations are not associated to an action plan.

Here we were interested in what happens to the memorandum when it becomes an integral part of the action goal. Does the status of a memory alter when it is the deliberate target of an upcoming action, compared to when it is equally task-relevant, but not the target of an action? We sought to answer this question and test the hypothesis that making an object in VWM the direct target of an action plan leads to a strengthening and therefore prioritization of its representation, compared to when the planned action is dissociated from it. To this end, we used a combined VWM and attentional capture paradigm (cf. Feldmann-Wüstefeld & Schubo, 2015; Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005), where the VWM representation varied in its functional relationship to the action plan. Participants were first asked to remember a geometric shape for a memory test at the end of the trial, in which they determined whether a probe stimulus matched the memorandum or not. Importantly, participants were instructed to respond differently to a match depending on the condition. In the action condition, they were asked to grip the probe stimulus if this matched the one held in memory. In the control condition, they were asked to perform the same matching operation, and the same action, but now the latter was directed towards a button-shaped stimulus that was presented together with the probe. For both conditions, if the probe stimulus did not match the memorized shape, participants had to altogether withhold their response. Note that this way, in both conditions the memory probe was equally relevant to the planned grip movement as such; indeed, in both conditions, to know whether they had to execute the movement or not, participants had first to evaluate whether the probe matched the item held in working memory or not. The only difference between the two conditions was therefore that, in one, the VWM object was also the target of the prospective action, whereas in the other it was not.

To test whether such prospective action-linking leads to strengthening of VWM representations and consequently, stronger attentional biases during the delay period between memorandum and test, we presented participants with a visual selection task display in which they had to select a target (i.e., the letter N) by means of a saccade (see Fig. 1). In the visual selection task, the memorized item was always present, either surrounding the target (congruent trials) or a distractor (incongruent trials). We hypothesized that in the action condition, the reciprocal link between the remembered shape and the action plan would strengthen its sensory representation, and thus render any matching stimuli in the visual selection task display more salient. If so, this should result in stronger attentional capture, as expressed in saccadic latencies and directions. We relied on eye movements rather than manual responses for measuring attentional biases, as this prevented the possible confound of motor interference between the memory and visual selection tasks. Moreover, these oculomotor measures are arguably more tightly linked to attention than manual responses, which may be influenced by post-selection processes.
2. Materials and methods

2.1. Participants

We recruited 52 healthy human participants (age range: 18–35 years, 16 males) with normal vision, who participated for either credits or monetary compensation. The sample size was established by means of a power simulation informed by piloting data, run with the aim of achieving a power > 0.8 for the interaction term between the variables action condition (action/control) and shape congruency (congruent/incongruent; see below for a more detailed explanation of the variables in the experiment). Participants provided digital consent for participation in the experiment. The study was approved by the Scientific and Ethics Review Board of the Faculty of Behavior and Movement Sciences at the Vrije Universiteit Amsterdam. Participants with <80% accuracy (i.e., <50% correct responses in the memory test), poor eye tracking data (i.e., <50% trials with proper fixation, defined as median eye position within 0.5° from the centre of the fixation cross), an average reaction time higher than 2.5 standard deviations of the group mean, or <30 trials per condition were excluded from the analysis. This led to the exclusion and replacement of three participants.

2.2. Apparatus

Participants were seated at 40 cm distance from a DELL P2418HT Touch Screen (1080 × 1920 pixels, 100 Hz sample rate). Participants positioned the elbow of their dominant arm on a wooden block, located at ~20 cm from the touch-screen. An Eyelink 1000 Tower eye-tracker was used to track participants’ right eye, at a rate of 1000 Hz. The experiment was coded in OpenSeame 3.3.8 (Mathôt, Schreij, & Theeuwes, 2012).

2.3. Stimuli and task

Each trial started with a fixation display, shown for a random time interval between 1200 and 1500 ms in duration. A shape outline was then presented at the centre of the screen for 1800 ms, and participants were asked to memorize it (see Fig. 1). Each shape was an irregular version of one of six basic geometric categories: a quadrilateral, a hexagon, a cross, a triangle, a heart, or a star. To promote the use of visual working memory, irregularities were introduced by jittering the vertices. At the memory test, a non-match would then consist of a shape of the same category, but with slightly differently jittered vertices. Two sets of shapes were thus created and assigned to each condition. The shapes were rendered in white on a grey background (RGB: 128,128,128), and subtended between 7.0 and 9.9 degrees of visual angle, depending on the shape. The encoding period was followed by a delay interval, randomly varying across trials between 900 and 1200 ms. On most trials a visual selection task display then appeared, presenting two geometric shapes arranged on the right and on the left of the fixation cross, at a distance of 7 visual degrees from fixation. These shapes always consisted of one circle (of diameter 7.37°) and another geometric shape, which matched the one held in working memory (having the same dimensions as the encoded item). Each geometric shape contained a white visual stimulus: either a horizontal hourglass or the letter N, both of dimensions 1.51° × 0.84° (length x height). Participants were instructed to find the letter N (i.e., the target) as fast as possible by performing a saccade towards it, while still retaining a high level of accuracy. The surrounding shapes were thus completely task-irrelevant. If the saccade landed within 2 visual degrees from the centre of the target or if >2000 ms passed since the onset of the visual selection task display, the display was terminated and replaced by a central fixation, for an interval that randomly varied between 900 and 1100 ms. Finally, a memory probe and a button-like shape were presented on the screen, each randomly presented either 7 visual degrees above or below the fixation cross. If the probe shape matched the one held in memory, participants had to perform a precision grip either towards the probe shape itself (action condition) or towards the button-like shape (control condition). Here too, participants were asked to respond as fast as possible, while remaining accurate. In both the action and control condition, if the probe did not match the memory template, the movement was to be withheld. In the no-match trials, the probe shape was from the same category as the stimulus held in memory, but with some variation. This was done to encourage participants to memorize the shapes visually rather than semantically. Independently of whether the action was executed, the next trial started 2200 ms from probe onset. In order to encourage an action-ready state in observers, we also inserted a number of catch trials in each block. On these trials the visual selection task was
skipped and the trial proceeded directly to the memory probes stage where the probe always matched the item held in working memory (i.e., catch trials).

2.4. Procedure

Participants performed one experimental block of the action condition and one experimental block of the control condition. Condition order was counterbalanced across participants. Each experimental block consisted of four mini-blocks of 32 trials each, thus resulting in 256 trials in total. In each experimental block, participants were presented with a different set of shapes, which were used throughout the whole block. Participants first practiced the task, either for two mini-blocks of 16 and 32 trials, before the first experimental block, or for one mini-block of 32 trials, before the second experimental block. More practice was given initially as then the task structure was still new.

Each mini-block included 8 catch trials (25% of trials) in which the visual selection task was left out altogether, and the memory probe was presented immediately after the encoding phase. In the remaining 24 trials (75%), participants also performed the visual selection task. Of these trials, 50% (i.e., 12 trials in each mini-block) were congruent trials that is, trials in which the letter N was presented within the stimulus held in working memory; 50% were incongruent trials, in which the letter N was located within the circle, rendering the other geometric shape (matching the one maintained in working memory) a distractor.

The final memory probe matched the item held in memory with a probability of 58.33% in trials in which the visual selection task was present. In trials in which the visual selection task was skipped, the memory probe always matched the shape held in memory – to promote action planning during encoding. Across trials, this resulted in ~68.75% match trials and ~31.25% no-match trials.

2.5. Data analysis

2.5.1. Primary analysis

We were primarily interested in testing our hypothesis that VWM representations which are the direct target of an action plan are strengthened compared to VWM representations which are not the target of an action plan. To do this, we looked at attentional biases induced by VWM representations during the visual selection task. We expected to find greater VWM-induced attentional capture in the action condition. We estimated attentional capture by looking at saccade latencies towards the target letter (N).

Practice blocks as well as catch trials were excluded from the analysis. Only trials with a correct response in the memory test were considered (69.10% of all trials). Saccades latencies were computed as those detected by the Eyelink software (i.e., saccades were detected as those eye movements with a velocity >30°/s and acceleration >8000°/s²). A saccade was classified as hitting the target if it fell within 2 visual degrees from the centre of the target. Trials with excessively fast or slow saccade latencies were removed independently for each participant. Outlier trials were defined as those data points falling above or below 2.5SD from each participant’s mean at each combination of factor levels (i.e., condition (action/control) x shape congruency (congruent/incongruent)) (1.70% of all trials).

Following Lo and Andrews (2015), saccade latencies on correct trials were analyzed by means of a generalized linear mixed-effects model (GLMM), with the data modelled as following an inverse Gaussian distribution to reflect the skew. The GLMM that best fit the data presented four fixed independent variables (see Table 1 for a comparison between the models considered): action condition, shape congruency, target position, and target repetition. All categorical values contained two levels: for the variable action condition these were labeled action and control; for the variable shape congruency, they were called congruent and incongruent; for the variable target position they were left and right, which referred to the target position with respect to the central fixation cross; finally, for the variable target repetition they were labeled repetition and switch, to indicate respectively those trials in which the target was located on the same side of the display as in the previous trial, and those trials in which the target position changed compared to the previous trial. Both main effects and interaction effects between the fixed variables action condition and shape congruency were estimated, whereas for the variables target position and target repetition only main effects were included in the model. The variables target position and target repetition were treated as covariates, to exclude variance in the data that could be driven by the position of the target or by inter-trial priming. Random effects included a random intercept for the variable subject. The GLMM model was defined by means of the glmer function of the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015). Planned two-tailed pairwise comparisons tests were conducted to uncover the source of any significant interactions. Given that the mapping out of an already significant interaction was the sole purpose of conducting these extra tests, they were not corrected for multiple comparisons.

2.5.2. Secondary analyses

Results from the primary analysis revealed a significant interaction in the hypothesized direction (see Results section). To better understand to what extent pre-selection attentional mechanisms contributed to the results of the primary analysis, we analyzed the proportion of trials in

| Model | Parameters of the model | AIC | BIC | P-value |
|-------|-------------------------|-----|-----|---------|
| M1    | Most complex model in terms of fixed effects models | Fixed: Action condition, Shape validity, Target repetition, and the interaction between Action condition and Shape validity; Random intercept: participants | 109,451 | 109,508 |
| M2    | Model without interaction term | Fixed: Action condition, Shape validity, Target repetition, and the interaction between Action condition and Shape validity; Random intercept: participants | 109,463 | 109,512 |
| M3    | Model without the covariate term target position | Fixed: Action condition, Shape validity, Target repetition, and the interaction between Action condition and Shape validity; Random intercept: participants | 109,478 | 109,508 |
| M4    | Model without the covariate term target repetition | Fixed: Action condition, Shape validity, Target repetition, and the interaction between Action condition and Shape validity; Random intercept: participants | 109,584 | 109,634 |

Table 1

Models fitted to the saccade latency data. As can be seen, the best model fit of the data is provided by the model which considers the main effects of Action condition, Shape congruency, Target position, and Target repetition together with the interaction between Action condition and Shape congruency. This can be seen by looking at the AIC and BIC values of M1, which are lower than those of M2, M3 and M4. Confirmation that M1 is the best model comes from a log-likelihood ratio test, the p-values of which are shown in the table.
which participants first looked at the target, the first saccade latencies to the
distractor in incongruent trials, and the first saccade latencies to the
target in congruent trials. Saccades were again computed as those
detected by the Eyelink software, and participants were considered to have
looked at the target/distractor when their saccades fell within 2 visual
degrees from the centre of the target/distractor.

The proportions of trials in which participants first looked at the
target were analyzed by means of a non-parametric repeated measures ANOVA, implemented with the art function of the arTool package in R (Wobbrock, Findlater, Gergle, & Higgins, 2011), with action condition and shape congruency as within-subject factors. To trace the source of any significant interaction, planned two-tail comparisons were again conducted without correcting for multiple comparisons.

Saccade latencies towards the distractor on incongruent trials and saccade latencies towards the target in congruent trials were analyzed to see whether there was any difference in oculomotor speed between the action and control condition. Differences in oculomotor speed would have pointed towards an effect of action condition on pre-selection attentional mechanisms. Saccade latencies were not found to follow any known distribution. Differences between conditions in saccade latencies were therefore tested on weighted means (where the weights were the number of trials per participant) by means of a Wilcoxon signed-rank test.

Finally, we asked to what extent the results of the primary analysis could be explained by differences in post-selection attentional mechanisms between the two conditions. We examined this by estimating the duration of fixations performed on the memorized shape in incongruent trials (i.e., when participants were distracted by the memorized shape in the visual selection task). Fixations were analyzed by means of a GLMM model, which allowed us to account for the correct distribution of the data (i.e., Gamma distribution) and for the different number of trials across subjects. The variable action condition was fixed and a random intercept was used for the variable subject.

2.5.3. Performance in the memory test

Accuracy in the memory test was computed as the percentage of trials in which participants responded correctly to the memory test (i.e., by gripping the shape in match trials and by withholding the response in no-match trials). Accuracy differences between the two conditions across all trials were computed by means of a GLMM, where data was indicated to follow a binomial distribution. In the model, accuracy was the dependent variable, while action condition was the fixed independent variable. The variable subject was used as a random intercept. Differences in accuracy were also tested separately for catch trials and trials with the visual selection task.

We also tested differences in grip reaction times between the two conditions. The analysis was run on those trials that were retained for the primary analysis, from which trials with reaction time outliers in the memory test were excluded. Outlier trials were defined as those data points falling above or below 2.5SD from each participant’s mean at each combination of factor levels (i.e., action condition (action/control) x shape congruency (congruent/incongruent)). Grip reaction times in the memory test were also analyzed for catch trials only. Data were modelled by means of a GLMM, and they followed an inverse gaussian distribution. The dependent variable was reaction times, and the independent variable was action condition. The variable subject was added as a random intercept.

2.5.4. Controlling for strategic perceptual resampling

One possibility is that participants were not captured by the memorized item in the visual selection task, but that they rather looked strategically at the memorized shape so as to refresh their memory of the item (i.e. engaging in strategic perceptual resampling (Woodman & Luck, 2007)). To check for this alternative explanation of our findings, we tested whether trials when participants first looked at the memorized shape in the visual selection task were also those with better memory performance and faster reaction times (a proxy for readiness and confidence in the decision made at test). We modelled accuracy by means of a GLMM, where accuracy followed a binomial distribution. Reaction times were also modelled with a GLMM, and they followed an inverse gaussian distribution. Action condition and First selected item were considered as the dependent variables in both models. The variable First selected item, counted two levels, namely memorized shape and circle, indicating which item was first selected in the visual selection task. The variable subject was the random intercept.

2.5.5. Analysis of eye movements during encoding

To determine whether action planning may have affected visual sampling during the encoding of the geometric shape, distance travelled by the eye, fixation count and fixation duration during the encoding phase were computed for each condition separately. These variables were chosen because each of them provides unique information about visual scanning behavior. Any difference between conditions in one of these three measures would indicate a difference in the way visual information is sampled (e.g., more fixations in one condition reflect the need to sample more information in preparation for the upcoming memory test). Differences in scanning behavior were statistically evaluated by means of non-parametric Wilcoxon signed-rank tests. Only fixations directed to the object outline were included in the analysis (those fixations at a distance larger than 1.5 visual degrees from the fixation cross). This was done to exclude fixations directed towards the fixation cross (most of them) as these could obscure potential differences in duration of fixations on the geometric shape.

3. Results

3.1. Primary analysis: saccade latencies towards the target

Fig. 2 shows the mean latency of the saccade that first landed on the target in the visual selection task. As can be seen in this figure, saccades to the target were generally faster when the target was surrounded by the shape held in visual working memory (i.e., in valid compared to invalid trials; main effect of shape congruency: t = −35.04, p < 2e-16, CI = [−150.12, −134.21]). Critically, in line with our hypothesis that planning an action towards an item strengthens its representation in VWM, this effect was more pronounced in the action compared to the control condition, as reflected in a significant cross-over interaction between action condition and shape congruency (t = 3.99, p = 6.57e-05, CI = [11.21, 32.84]).

This interaction effect was primarily driven by condition differences in the incongruent trials. Planned post-hoc comparisons revealed that in incongruent trials, saccade latencies towards the target were significantly slower (about 17 ms) in the action condition compared to the control condition, $t_{\text{action}, \text{invalid}} = -3.79, p = 0.0002, CI = [8.78, 27.59]$, indicating that in the action condition participants were distracted more by the shape held in working memory. In congruent trials, saccades towards the target were numerically faster in the action condition, but this difference was not significant $t_{\text{action}, \text{valid}} = -3.79, p = 0.0002, CI = [8.78, 27.59]$, most likely due to a floor effect. The GLMM furthermore showed a significant main effect of action condition ($t = 3.79, p = 1.51e-04, CI = [-27.59, -8.78]$), indicating generally shorter latencies in the control condition. This analysis also revealed an effect of the covariates target position ($t = 5.51, p = 3.44e-08, CI = [9.49, 19.95]$), reflecting overall faster saccade latencies when the target was located to the left of the fixation cross, and trial repetition ($t = 12.05, p < 2e-16, CI = [26.22, 36.40]$), showing slower saccade latencies in switch trials.
3.2. Secondary analyses: pre- and post-selection attentional processes

Our pattern of results so far suggests that VWM items towards which a future action is planned grab attention more so than VWM items that are not the target of an action plan. Such results could be driven by pre-selection mechanisms, post-selection mechanisms, or both. Therefore, to better characterize the pattern of saccade latencies observed, we ran several additional analyses. First, we examined differences in the proportion of trials in which participants directed their very first saccade towards the target. If linking an action plan to a VWM representation makes the represented item more salient, then one would expect participants to first look at the item matching the item held in working memory more often in the action condition. Such a difference in the proportion of first target saccades would thus signify a difference in pre-selection attentional processes between the two conditions. Fig. 3 shows the proportion of first saccades computed per action condition and shape congruency, separately. A non-parametric ANOVA showed a significant main effect of shape congruency \((F = 288.42, p < 2e-16, \eta^2_p = 0.65)\), but no main effect of action condition \((F = 0.09, p = 0.76, \eta^2_p = 0.001)\). Importantly, here too, we observed a significant interaction between shape congruency and action condition \((F = 5.71, p = 0.018, \eta^2_p = 0.036)\). Post-hoc two-sided Wilcoxon tests showed a statistically significant difference between the two conditions on congruent trials \((Z = 2.24, p = 0.025, \text{effect size } r = 0.22)\), as well as on incongruent trials \((Z = 2.77, p = 0.006, r = -0.35)\), reflecting the fact that in the action condition, participants were more likely to first look at the shape held in memory,
regardless of whether or not it contained a target or distractor. These findings thus confirm that the shape was rendered more salient when its memory representation was tightly linked to an action plan.

We next investigated how fast participants looked at the distractor when it matched the shape in VWM and how long they dwelled on it on incongruent trials. Faster saccade latencies towards the distractor in the action condition would provide further evidence for an effect of action planning on pre-selection attentional mechanisms, whereas longer dwell times on the distractor would argue for an additional effect on post-selection attentional mechanisms (i.e., after selecting the item, participants find it more difficult to disengage from it). Fig. 4A and B shows the means of these dependent measures together with 95% confidence intervals. The distribution of saccades towards the distractor was not fitted by any of the known distributions. A non-parametric Wilcoxon rank-test was therefore conducted on the latency means, weighted by the number of trials. The main effect of action condition was not significant ($Z = 0.08$, $p = 0.94$, $r = 0.01$), indicating no significant differences in saccade latencies to the distractor between the action and control condition. Therefore, although participants more often first looked at the item corresponding to the one held in VWM, even when it was a distractor, in the action condition, their eyes were not more quickly drawn towards it. Fixation duration data followed a gamma distribution and was fed into a GLMM, where the variables target position and target repetition were considered as covariates. This analysis showed a main effect of condition ($t = 2.88$, $p = 0.004$, CI $= [-13.99, -2.65]$), indicating longer dwell times on the distractor in the action condition. There was also a main effect of target repetition ($t = 5.30$, $p = 1.16e-07$, CI $= [9.92, 21.57]$), reflecting the fact that participants tended to fixate more on the distractor in switch trials, but there was no main effect of target position ($t = 0.85$, $p = 0.40$, CI $= [-3.31, 8.33]$). Collectively, these results show that when participants first looked at the distractor, their eyes dwelled on it longer in the action condition. This latter finding also suggests a difference in post-selection attentional mechanisms between the two conditions: participants not only first looked at the memory-matching distractor more often in the action condition, but they were also slower to disengage from this stimulus when it was the target of an action plan.

Finally, we determined if tightly linking an action plan to an item in VWM would speed up attentional capture by the target shape. To this effect of considered as covariates. This analysis showed a main effect of GLMM, where the variables corresponding to the one held in VWM, even when it was a distractor, in these trials, their eyes were not more quickly selected the item, participants were not faster in selecting the target in the action condition compared to the control condition. Overall, our findings thus indicate that both pre- and post-selection mechanisms contributed to the greater attentional capture observed in the action condition, but these effects were not accompanied by faster first saccade latencies towards the memorized item.

3.3. Performance in the memory test

We did not expect accuracy on the memory test to be sensitive to potential differences between the action and the control condition, since participants only had to memorize one item and were presented with the same six shapes throughout each condition (6 shapes presented in 128 trials), which made the task arguably easy. Indeed, participants scored high on average on the memory test ($\mu_{\text{acc}}$, action $\pm \sigma_{\text{acc}}$, action $= 0.927 \pm 0.048$; $\mu_{\text{acc}}$, control $\pm \sigma_{\text{acc}}$, control $= 0.929 \pm 0.043$), suggesting that they understood and correctly performed the task. Their performance also did not differ between the action condition and the control condition.
This was true both when estimating accuracy across all trials ($z = 0.409, p = 0.682, CI = [-0.104, 0.160]$), and when computing performance separately for catch trials ($z = 1.667, p = 0.096, CI = [-0.45, 0.560]$) and for trials with the visual selection task ($z = 0.342, p = 0.732, CI = [-0.173, 0.121]$). Accuracy scores in both conditions were not influenced by whether, in the visual selection task, participants first selected the memorized shape. Indeed, when we looked at how behavior in the visual selection task related to accuracy in the memory test, we found neither a main effect of First selected item ($t = 0.947, p-value = 0.361, CI = [-0.159, 0.435]$) nor an interaction between First selected item and action condition ($t = 0.601, p = 0.548, CI = [-0.562, 0.298]$), suggesting that the attentional biases displayed in the visual selection task did not reflect participants' attempt to sample once more, before the memory test, the memorized item to maximize their memory accuracy (i.e., strategic perceptual resampling).

A comparison of the reaction times in the memory test showed that participants were significantly faster in the action compared to the control condition ($t = 5.557, p < 0.001, CI = [0.012, 0.024]$), both in trials with the visual selection task ($t = 5.476, p < 0.001, CI = [0.014, 0.029]$) and in catch trials ($t = 2.41, p = 0.016, CI = [0.0024, 0.0235]$). Differences in reaction times between the two conditions were expected given that in the action condition participants directly gripped the memory-matching probe, whereas in the control condition they first looked at the shape and then, in case of a match, they gripped the button-like shape, conceivably introducing an additional switch cost. Reaction times in the memory test, as accuracy scores, did not vary depending on whether participants first selected the memorized item in the visual selection task. Again, we found no main effect of First selected item ($t = 1.568, p = 0.117, CI = [-0.025, 0.0028]$) and no interaction between First selected item and action condition ($t = 0.994, p = 0.395, CI = [-0.021, 0.019]$), indicating that no strategic perceptual resampling was adopted, and no differences in strategy between the two conditions was employed.

3.4. Eye movements at encoding

The above results leave unclear whether the observed effects of directly linking an action plan to an object in VWM are related to differences in the encoding and/or maintenance of that object in VWM. The action plan was already known at the moment of encoding, so could have also affected how the shape was sampled and thus encoded in VWM. To examine this, we explored whether there were any differences between the action and the control condition in terms of eye movements during the encoding of the geometric shape. Specifically, we examined condition differences in distance travelled by the eye, as well as fixation count and durations during the presentation of the stimulus, while considering the distance of the eyes from the fixation cross. A non-parametric Wilcoxon signed-rank test showed that there were no differences between the action and control condition in saccade count ($z = 0.70, p = 0.48, r = 0.10$), travelled distance ($z = 0.24, p = 0.81, r = 0.03$), fixation count ($z = 0.31, p = 0.75, r = 0.04$), or fixation duration ($z = 1.028, p = 0.20, r = 0.18$). Thus, we could not discern any differences in eye movement patterns at encoding between the action and the control condition. This finding may suggest that the observed difference in attentional allocation during search between conditions cannot be attributed to differences in perceptual sampling between conditions.

4. Discussion

In the present study, we aimed to determine whether VWM representations of items that are the target of an action plan are strengthened, and therefore lead to prioritization, compared to VWM representations of items that – although nominally equally relevant for solving the task – are dissociated from a prospective action. We found that when items in VWM were tightly linked to an action plan, they more often biased attention towards that item when it appeared in a subsequent visual selection task (i.e., it became more salient), as indicated by a greater attentional capture effect. Indeed, the difference in latencies of saccades that first landed on the target between trials in which the memorized item surrounded the distractor (incongruent trials) and trials in which the memorized item surrounded the target (congruent trials) was larger when an action plan was directly associated with a VWM representation (action condition) than when it was not (control condition). In particular, in the action condition, the item held in memory was more distracting to participants, as indicated by significantly slower saccade latencies towards the target in incongruent trials. Moreover, the memory-matching shape attracted more first saccades when it was the target of a future action, indicative of attentional capture. Our results critically expand past work on the effects of action planning on VWM by showing that planning an action further prioritizes those items that are the target of action plan compared to non-target stimuli. Our study therefore provides novel support for the idea that VWM is action-oriented (Engel et al., 2015), and that attentional biases emerge from establishing an association between sensori-motor memory representations and action plans, as proposed by Olivers and Roelfsema (2020).

So far, previous studies have investigated action-induced weighting of VWM representations by means of dual-task paradigm designs where the link established between memory representations and action plans has been rather loose. In these studies, working memory and action tasks are typically unrelated to each other, and visual properties that are relevant for the correct performance of one task may be irrelevant or even inconducive for the performance of the other task. For example, participants are prompted to plan and execute an action towards a certain location in space, which is not predictive of where a subsequent memory probe will appear (Heuer et al., 2017; Ohl & Rolfs, 2018), de facto biasing attention in a manner that could negatively interfere with overall performance in the memory task. Or participants are asked to plan a grasping or pointing movement towards a certain stimulus, while such differential actions are not relevant to the memory task and need to be performed independently of whether the probed stimulus matches or not a memorized one (Heuer & Schubö, 2017). That is, in these studies, the memory stimulus is not univocally paired with an action, and the action is not conditional on some properties of the stimulus. While such “indirect” designs have been tremendously valuable in showing automatic associations between action, attention, and working memory, in real life, the objects that we hold in working memory often inform and guide our actions directly: cognitive and action operations are rarely unrelated, unless we are attempting to do two different things at the same time. So far, few studies have directly examined the inherent functional role that action plays in VWM or the extent to which VWM is action-oriented (Olivers & Van der Stigchel, 2020; van Ede, 2020). By directly manipulating whether an action had to be planned on an item in VWM itself or not, and by assessing how this affected attentional biases, our findings thus expand our understanding of the relationship existing between action, VWM, and attention.

Moreover, while previous studies showed that action planning towards action-relevant properties of an external stimulus induces attentional biases to those same properties within VWM (i.e., action-induced external attention leads to biases in internal attention), our study reveals the opposite sequence of mechanisms: Planning an action towards an item held in VWM influences the extent to which external attention is then biased towards matching stimuli – presumably mediated by a stronger internal representation. In line with the idea of VWM as an action-oriented cognitive operation, results from our study clearly show that mnemonic information that is the target of an action plan preferentially guides behavior, by being prioritized compared to information that is not so tightly linked to an action plan. Combining previous studies with our study thus demonstrates the full spectrum of interactions between action, attention, and working memory. Crucially, the greater interference between VWM and perception in the action condition likely was not due to the employment of a strategy aimed at
refreshing the memory for the upcoming memory test (cf. Woodman & Luck, 2007). In fact, we showed that when participants first looked at the memorized item in the visual selection task, they did not perform better on the memory test, neither in terms of accuracy nor of grip reaction times. Moreover, there was no difference between the action and the control conditions in the way attentional capture in the visual selection task related to performance on the memory test.

Note that there was no spatial overlap between the locations of the memorandum (center), the search items (left and right from fixation), and the memory test items (above and below fixation) – with the location of the memorandum in the latter two displays moreover being randomized. Therefore, the greater attentional biases during search observed when an action was planned towards the item currently held in VWM (i.e., in the action condition) cannot be explained in terms of increased spatial attention, but must reflect enhanced feature-based (or object-based) attention. Previous work has demonstrated a close relationship between attention and action planning within the spatial domain (see Heuer, Oih, & Rolfs, 2020 for a review). Yet, there is less evidence for a similar link in the feature domain. While one study has shown that planning different movements, namely grasping and pointing, leads to improved VWM for different visual features (size and color, respectively) (Heuer & Schub, 2017), this study left unresolved if and how these findings related to changes in feature-based attention. Our results clearly show that when a visual stimulus of unknown location is the goal of an action plan, it biases attention more strongly compared to when that same stimulus is not the primary target of an action. As such, they suggest a tight relationship between action planning, VWM and feature-based attention.

It remains to be determined what neural mechanisms underlie these behavioural effects. One likely possibility is that the activation of feedback loops projecting from motor to early visual brain regions strengthens VWM representations, as predicted by Olivers and Roelfsema’s (2020) framework. Indeed, studies in the domain of visual perception have shown that activity in early visual cortex can be modulated by higher-order areas involved in action planning by means of top-down connections (Gallivan, Chapman, Gale, Flanagan, & Culham, 2019). Moreover, this modulation was found to be highly selective and to involve only areas which retinotopically represented the designated target’s features, in accordance with their action relevance (Guttingel et al., 2015). A few studies (e.g., Monaco, Gallivan, Figley, Singhal, & Culham, 2017; Singhal, Monaco, Kaufman, & Culham, 2013) also observed that areas involved in action execution can modulate activity in visual areas in a top-down manner, even in the case of memory-guided actions, when visual information is no longer available. This work together supports the notion that VWM representations that are linked to an action plan can be strengthened by means of feedback signals from motor to visual areas, although future studies are necessary directly to test this idea.

Regardless of the precise neural mechanisms involved, we found that functionally, planning an action towards an item in VWM influenced both early and late attentional processes. During the visual selection task, more first saccades went to the memory-matching shape when the shape was part of the action plan. These effects point towards an effect of action planning on early attentional processes. In addition, we found that once observers had fixated the memory-matching object, when it was a distractor, they were slower to disengage from it when it was part of the action plan, as reflected in longer fixation durations. This latter finding indicates that the action orientation also affected post-selection processes. We did not find evidence for an overall effect of action planning on oculomotor speed, since first saccade latencies, both in valid and invalid trials, did not significantly differ in the action vs control condition.

Although our study shows that tightly linking an action plan to a VWM representation leads to increased attentional biases, it remains unclear if this effect is due to effects of action planning during VWM encoding, on VWM maintenance, or both. An exploratory analysis of eye movements at encoding did not reveal differences in eye patterns between the action and control condition, suggesting that the observed differences may reflect differences in covert attention, in maintenance processes, or both. Of importance, both Ohl and Rolfs (2018) and Heuer and Schubo (2017) showed that action planning effects were present both when the action plan was specified at encoding and when it was specified during maintenance. The same may hold true for the situation in which an action is planned directly on an item in VWM, as in our study, but this awaits empirical testing.

It also remains to be established to what extent the action-induced strengthening of VWM representations varies depending on how well-defined the action-VWM coupling is. Note that in our study, as is arguably the case for any task, in the end an action had to be performed even in the control condition. So it was not the case that in that condition there was no memory-action link at all – it was just less direct. This suggests that the tighter the link between an action plan and a VWM representation the greater the enhancement of the representation. Future studies could further test this idea, by systematically varying the strength of the motor-VWM link. For example, one could compare representations that are unequivocally linked to a concrete pre-defined action plan (i.e. one the observer knows what to do) with representations that are associated with multiple possible action plans that have yet to be defined (i.e. the observer knows he or she has to act, but does not know the specific action yet). Such studies will help to further characterize the relationship that exists between action planning, VWM, and attention.

5. Conclusions

To conclude, the VWM representation of a stimulus that is directly coupled to an action plan renders this stimulus more salient. This novel finding adds to an emerging body of work showing that VWM is a fundamentally action-oriented cognitive operation, and may be better understood as a memory system that evolved to guide future behavior rather than to passively store past information, and hence, that VWM should also be examined as such. Our results also support the notion that attention is better conceived as an emergent property of the coupling of motor to sensory representations, in line with attention’s crucial role in goal-directed behavior.

Supplementary information

The raw data are available at OSF (https://osf.io/z5vc4/).

Declaration of Competing Interest

We have no known conflict of interest to disclose.

Data availability

Data is available at OSF (https://osf.io/z5vc4/).

Acknowledgments

We thank Freek van Ede for inputs and suggestions on the original draft of the paper. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. Perspectives on Perception and Action, 15, 395–416.

Bae, G. Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. Journal of Neuroscience, 38 (2), 469–422.
Baldau, D., & Deubel, H. (2008). Properties of attentional selection during the preparation of sequential saccades. Experimental Brain Research, 18(4), 411–425.

Baldau, D., & Deubel, H. (2010). Attentional landscapes in reaching and grasping. Vision Research, 50(1), 999–1013.

Baldau, D., Wolf, M., & Deubel, H. (2006). Deployment of visual attention before sequences of goal-directed hand movements. Vision Research, 46(26), 4535–4574.

Bates, D., Mächler, B., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48.

Boettinger, S. E., Gresch, D., Nobre, A. C., & van Ede, F. (2021). Output planning at the input stage in visual working memory. Science Advances, 7(13), Article eabe8212.

Bosco, C., Daniele, F., & Fattori, P. (2017). Reaching and grasping actions and their context shape the perception of object size. Journal of Vision, 17(12), 1–19.

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. Neuropsychologia, 49(6), 1407–1409.

Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. Philosophical Transactions of the Royal Society, B: Biological Sciences, 362(1485), 1585–1599.

Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. Attention, Perception, & Psychophysics, 81(7), 2265–2287.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. Annual Review of Neuroscience, 33, 269–298.

Clark, A. (1995). Moving minds: Situating content in the service of real-time success. In O. Helm, P. A., & R. I. Schubotz (Eds.), Action planning mediates guidance of visual attention. Memory & Cognition, 4(3), 468–469.

Feldmann-Wüstefeld, T., & Schubotz, R. I. (2007). Intentional control of attention: Action preparation of sequential saccades. Journal of Vision, 7(11), 22–29.

Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. Psychological Research, 71(1), 9–14.

Flanagan, J. R., & Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. Annual Review of Neuroscience, 18(1), 193–222.

Engel, A. K., Maye, A., Kurthen, M., & Kasten, L. E. (2010). The role of the prefrontal cortex in guiding visual search. Nature Reviews Neuroscience, 11(5), 377–389.

Feldmann-Wüstefeld, T., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. Psychological Research, 71(1), 9–14.

Gallivan, J. P., Chapman, C. S., Gale, D. J., Flanagan, J. R., & Culham, J. C. (2019). Selective modulation of early visual cortical activity by movement intention. Cerebral Cortex, 29(11), 4662–4678.

Gusseit, E., Meeter, M., & Olivers, C. N. (2014). Is a search template an ordinary working memory? Journal of Experimental Psychology: Human Perception and Performance, 35(4), 1675–1689.

Gutteling, T. P., Kemenes, J. I., & Naggerts, S. F. (2011). Grasping preparation enhances orientation change detection. PLoS One, 6(5), Article e17675.

Gutteling, T. P., Park, S. Y., Kemenes, J. I., & Naggerts, S. F. (2013). TMS of the anterior intraparietal area selectively modulates orientation change detection during action preparation. Journal of Neurophysiology, 110(1), 33–41.

Gutteling, T. P., Petridou, N., Dumolin, S. O., Harvey, B. M., Aarnoutse, E. J., Kemenes, J. I., & Naggerts, S. F. (2013). Action preparation shapes processing in early visual cortex. Journal of Neuroscience, 35(16), 6672–6480.

Hannigan, M. N., & Deubel, H. (2018). Independent effects of eye and hand movements on visual working memory for object size. Journal of Experimental Psychology: Human Perception and Performance, 44(12), 37.

Hannigan, M. N., Komikaitis, D., Deubel, H., & Szinte, M. (2016). Oculomotor selection underlies feature retention in visual working memory. Journal of Neurophysiology, 115(2), 1071–1076.

Heuer, A., Crawford, J. D., & Schubotz, R. I. (2017). Visually driven selection of visual stimuli. Attention, Perception, & Psychophysics, 79(6), 2265–2287.

Heuer, A., Ohl, S., & Rolfs, M. (2020). Memory for action: A functional view of selection. Cognition, 230, 105274.