Research Article

Massive Effects of Saliency on Information Processing in Visual Working Memory

Martin Constant[1,2] and **Heinrich R. Liesefeld**[1,3]
[1]Department of Psychology, Ludwig-Maximilians-Universität München; [2]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München; and [3]Munich Center for Neurosciences–Brain & Mind, Ludwig-Maximilians-Universität München

Abstract

Limitations in the ability to temporarily represent information in visual working memory (VWM) are crucial for visual cognition. Whether VWM processing is dependent on an object's saliency (i.e., how much it stands out) has been neglected in VWM research. Therefore, we developed a novel VWM task that allows direct control over saliency. In three experiments with this task (on 10, 31, and 60 adults, respectively), we consistently found that VWM performance is strongly and parametrically influenced by saliency and that both an object's relative saliency (compared with concurrently presented objects) and absolute saliency influence VWM processing. We also demonstrated that this effect is indeed due to bottom-up saliency rather than differential fit between each object and the top-down attentional template. A simple computational model assuming that VWM performance is determined by the weighted sum of absolute and relative saliency accounts well for the observed data patterns.

Visual working memory (VWM) is a crucial hub in the processing of visual information, and its limitations are strongly related to general cognitive ability (Fukuda et al., 2010). Variation in VWM performance is typically interpreted in terms of some limited commodity (slots or resources; Cowan, 2001; Liesefeld & Müller, 2019a; Luck & Vogel, 2013; Ma et al., 2014), but alternative interpretations exist (Emrich et al., 2017; Oberauer & Lin, 2017). Identifying influences on VWM performance is of high applied and theoretical relevance because of its central role in theories of visual cognition.

It has been extensively demonstrated that how well an object is memorized hinges on its behavioral relevance, that is, on the explicit intention to favor one or several objects (top-down influences; Emrich et al., 2017; Souza & Oberauer, 2016). How VWM processing might differ for equally relevant objects because of contextual features of these objects themselves (bottom-up influences) has been largely neglected. In fact, all current models assume that, apart from random variation, all equally relevant objects within a display are

processed equally well or have the same chance of being processed. This assumption seems reasonable for highly controlled, abstract stimuli but might not hold for somewhat more naturalistic stimuli and for the everyday use of VWM in complex real scenes.

It is well known from the visual attention literature that factors other than top-down goals influence the allocation of processing resources (Awh et al., 2012; Liesefeld et al., 2018; Wolfe & Horowitz, 2017). A particularly strong influence on object processing that is largely neglected in the VWM literature is bottom-up saliency. An object is salient if at least one of its features stands out, such as the blackness of a black sheep in a flock of white sheep. More technically, saliency is largely determined by local feature contrast (Nothdurft, 1993): Via lateral inhibition (i.e., at the same hierarchical level

Corresponding Author:
Heinrich R. Liesefeld, University of Bremen, Department of Psychology
E-mail: Heinrich.Liesefeld@uni-bremen.de

of visual processing), neurons with overlapping tuning curves (i.e., coding similar features) mutually suppress each other (*lateral iso-feature suppression*; Li, 2002); the resulting net activity is highest for features that differ maximally from their immediate surroundings, because the respective neuronal activity receives little suppression. Because saliency has a strong and parametric influence on object processing in visual search (Liesefeld et al., 2016), it seems likely that salient objects are also prioritized for VWM processing.

In the rare cases in which the influence of object saliency on VWM processing has been studied, the design did not allow manipulating each object's saliency independently (Rajsic et al., 2016) or confounded saliency with the discriminability of the to-be-remembered feature. Klink et al. (2017), for example, had participants remember the orientation of Gabor gratings and manipulated saliency by varying the Gabor contrast (see Fig. 1a; see also Knops et al., 2014). In line with an effect of saliency, the lower the contrast, the worse the VWM performance. However, varying the contrast also influences the discriminability of the to-be-remembered orientation because the Gabor grating increasingly merges with the background for lower contrasts. In fact, in psychophysical studies, Gabor contrast is often used to scale discrimination difficulty (e.g., Alvarez & Cavanagh, 2008). These and other confounds also affect studies using quasinatural stimuli (which are by definition not well-controlled; for a review, see Santangelo, 2015). Nevertheless, these studies indicate that saliency has some influence on VWM processing.

To study the influence of saliency on VWM encoding under controlled conditions, we developed a task that deconfounds the saliency of target objects and discriminability of to-be-remembered features and allows the researcher to manipulate each object's saliency continuously and independently (see Fig. 1b). With this novel task, we conducted three experiments in which participants had to remember the color of three target objects. These three targets were always equally likely to be probed but differed in saliency either within or across displays. Our results show a strong impact of bottom-up saliency on how well equally relevant objects are stored in VWM.

Experiment 1

Method

In many VWM studies, participants hold the colors of a bunch of isolated objects in mind for a short retention period and then have to decide whether one of the objects changed color in a second display (*change detection*) or reproduce the color of a probed object
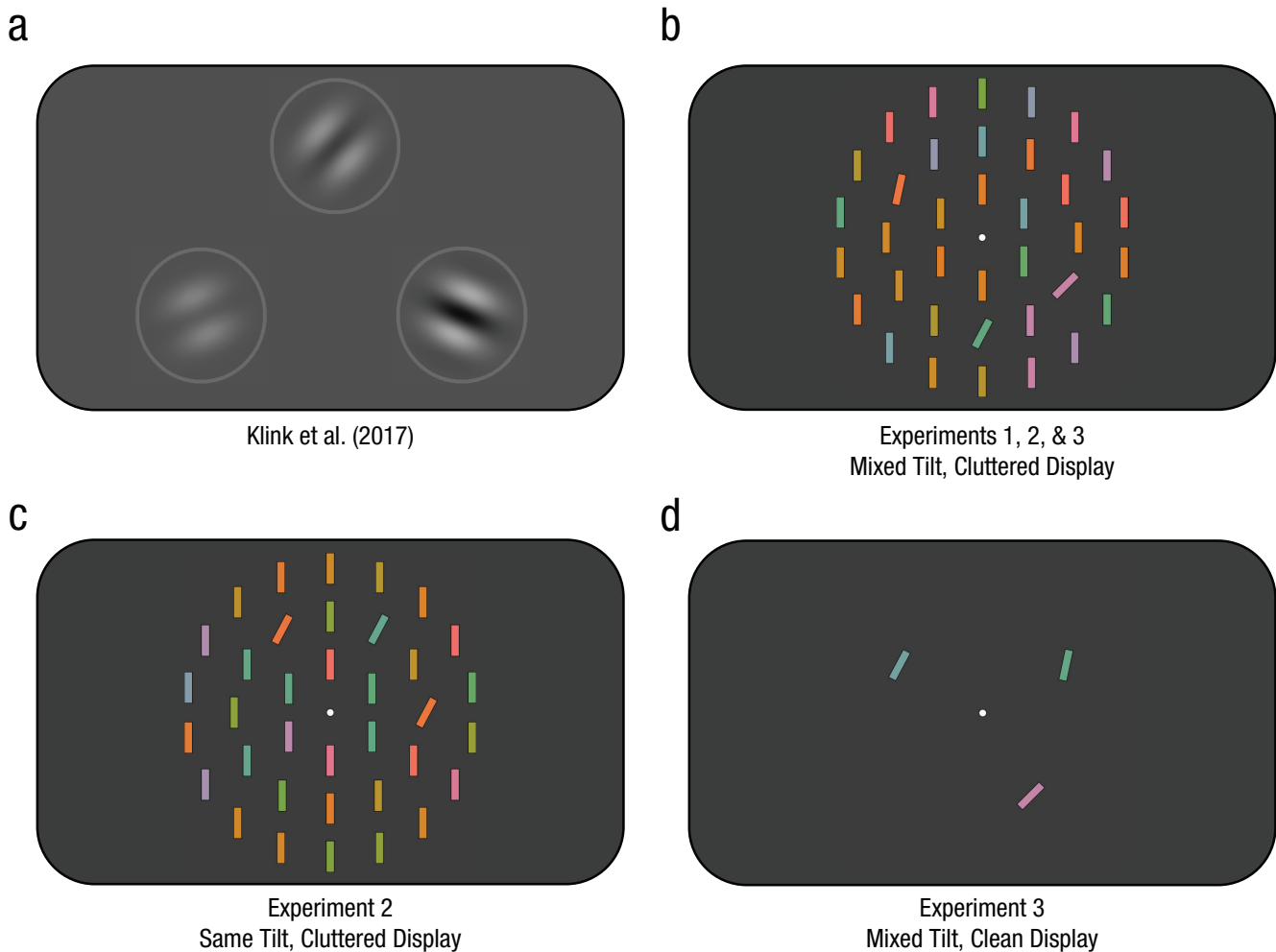
Statement of Relevance

The amount of visual information arriving each moment from our eyes is impossible to process to any reasonable extent by any limited system, and human visual processing abilities are severely limited indeed; the major bottleneck for visual processing is called visual working memory (VWM). Using a novel task design, we demonstrated that the selection problem is solved in part by preferably processing the most prominent objects within a scene. How well an object is processed in VWM is determined both by how much it stands out and by how strong the other competitors in the scene are. This study brings VWM research one step closer to the highly complex real world and reveals that saliency has a major impact on VWM processing that is easily overlooked in the traditionally very abstract VWM paradigm.

(*continuous report*). A wide variety of versions of this basic design exist, but the focus on isolated (i.e., highly salient) objects is common to virtually all of them (see Fig. 1d). To open up the VWM paradigm to the well-controlled examination of saliency effects, we developed a novel VWM task in which we can directly, gradually, and independently manipulate each object's saliency while keeping the discriminability of the to-be-remembered features and the objects' behavioral relevance untouched. This design also enables the use of modern computational models and neuroimaging methods.

We built on our previous experience from visual attention research to develop the task. In particular, Liesefeld et al. (2016) devised a visual search task that allowed a gradual manipulation of the search target's saliency (see also Nothdurft, 1993) and showed that search becomes faster as a continuous function of target saliency. By placing a tilted target bar into a dense array of vertical nontarget bars and adapting the tilt of the target bar (and therefore the contrast between target and nontargets), we were able to control target saliency to any desired precision. Liesefeld et al. (2017) showed that in this design, processing priority (measured by the order of attention allocations) is almost perfectly determined by object saliency.

Here, we translated this design to the study of VWM by employing memory displays featuring a dense array of vertical nontarget bars into which three differently tilted and randomly colored target bars were placed (see Fig. 1b). Participants had to remember the target bars' colors in order to later reproduce one of them. In

a

Klink et al. (2017)

b

Experiments 1, 2, & 3
Mixed Tilt, Cluttered Display

c

Experiment 2
Same Tilt, Cluttered Display

d

Experiment 3
Mixed Tilt, Clean Display

Fig. 1. A typical example of previous manipulations of saliency and design of the present memory displays. Klink et al. (2017) had participants remember the orientations of Gabor gratings and manipulated saliency via the gratings' contrasts (a); note how the contrast also influences the discriminability of the to-be-remembered orientations. In our novel task design (b and c; Experiments 1–3), participants have to remember the colors of three tilted target bars to later reproduce one of these colors, and saliency is manipulated via target tilt. Using the same tilt for all three of the target bars in Experiment 2 (c) equated the bars' relative saliency within each display. Removing the vertical nontarget bars in Experiment 3 (d) rendered all target bars highly salient (leaving only the isolated colored objects that are often used in studies of visual working memory).

order not to make color dominate the contrast (and therefore determine saliency), we also drew the nontarget bars in random (completely irrelevant) colors.

The critical deviation from previous research on VWM is that our displays are cluttered with irrelevant vertical nontarget bars. As explained above, this is necessary to control the saliency of the relevant bars because saliency of an object depends on its relationship to its immediate surroundings. This is not an artificial change to the task, though. It mimics a feature of the real world: Hardly any natural environment consists of well-isolated relevant objects, but the real world is cluttered with many objects that are irrelevant for the task at hand (e.g., Hollingworth, 2008). Also note that in Liesefeld et al.'s (2016) study, even targets with a 12°

tilt (the smallest tilt employed in the present study) produced clear pop-out, that is, participants were able to almost exclusively process the target bar and completely ignore the vertical nontarget bars. Thus, the vertical bars are sufficiently less salient than even the 12°-tilted bars, so that they likely do not significantly compete for VWM processing as distractors in other designs would (Liesefeld et al., 2014; Vogel et al., 2005; for a review, see Liesefeld et al., 2020).

Participants. For each experiment, sample size was determined via sequential testing with Bayes factors (BFs), following the recommendations by Schönbrodt and Wagenmakers (2018; for details, see the Supplemental Material available online). The critical tests determining

the stopping rule were directional (and thus conducted one-tailed; see https://osf.io/ktp6n for the preregistration). These tests examined whether VWM performance (the mean absolute angular distance between correct and selected response, i.e., *recall error*) would decrease with object saliency (tilt). This resulted in a sample of 10 adults with normal or corrected-to-normal vision (age: $M = 26.3$ years, $SD = 3.37$; four female; all right-handed).

Procedure and design. After a 1-s fixation dot, the memory display (see Fig. 1b) was presented for 350 ms. This display consisted of a dense array of vertical nontarget bars among which were three differently tilted (12°, 28°, 45°) target bars. The colors of all bars (both target and nontarget) were randomly chosen. Participants task was to remember the target bars' color. The memory display was followed by another 1-s fixation dot (delay period). A response display was then presented; this display contained a color wheel and outlined placeholder bars at the locations of each bar from the memory display. One of the placeholders was filled in black to indicate which bar to report (hereafter, *probe*), and participants were instructed to report the color they remembered for that bar by using the computer mouse to select a point on the color wheel. After each response, a feedback line appeared at the correct location on the color wheel to show the participant the correct response (and, by implication, how far off the actual response was).

Each participant completed a total of 600 trials divided into blocks of 30 trials each. Each condition (i.e., tilt of the probe) was randomly presented 200 times (10 times per block).

Data analysis. In addition to the *t* tests discussed in the Participants section, we report Bayes factors quantifying the evidence for the alternative over the null hypothesis (BF_{10}, BF_{+0}, or BF_{-0}) or the null over the alternative hypothesis (BF_{01}). For directional tests, we report the corresponding BF_{+0} or BF_{-0} (which place zero probability on negative or positive effects, respectively), rather than the undirectional BF_{10}.

Results

As expected, our manipulation of saliency had a huge and reliable impact on VWM performance (see Fig. 2): Despite all three objects being equally relevant, recall error was higher for 12° probes ($M = 59.07°$, between-participants 95% confidence interval [CI] = [50.04, 68.10]) than for 28° probes ($M = 41.84°$, 95% CI = [32.81, 50.86]), $t(9) = 6.56$, $p < .001$, $d_z = 2.07$, 95% CI = [0.93, 3.19], $BF_{+0} = 551.51$, and higher for 28° probes than for 45° probes ($M = 28.14°$, 95% CI = [19.12, 37.17]), $t(9) = 4.66$, $p < .001$, $d_z = 1.47$, 95% CI = [0.54, 2.37], $BF_{+0} = 70.6$. Effect sizes were so huge that despite the
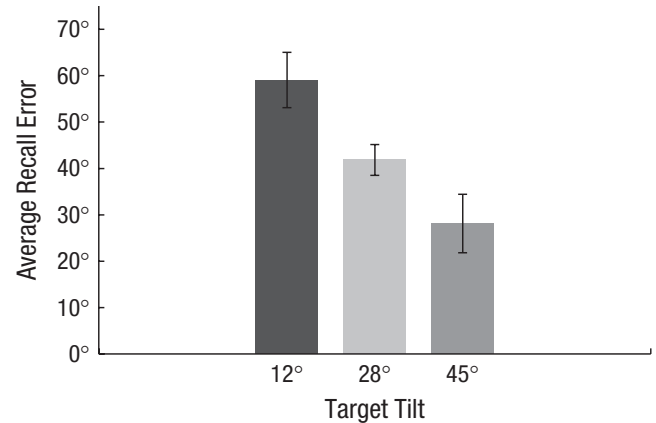
Fig. 2. Results from Experiment 1: average recall error (i.e., mean absolute difference between correct responses and given responses) for each of the three target tilts. Error bars represent 95% within-participants confidence intervals.

relatively small sample size (which we had defined as the minimum in our preregistration), the BFs indicated overwhelming evidence for both differences. This finding demonstrates that VWM performance is strongly and parametrically dependent on saliency.

Fitting the data to the Zhang and Luck (2008) model revealed that the probability that the probed item was in memory (p_{mem}) differed significantly between 12° probes ($M = 44.08\%$, 95% CI = [32.25, 55.89]) and 28° probes ($M = 68.89\%$, 95% CI = [57.06, 80.71]), $t(9) = 6.37$, $p < .001$, $d_z = 2.01$, 95% CI = [0.89, 3.10], $BF_{10} = 227.57$, and between 28° and 45° probes ($M = 86.41\%$, 95% CI = [74.58, 98.23]), $t(9) = 4.10$, $p = .003$, $d_z = 1.30$, 95% CI = [0.42, 2.14], $BF_{10} = 18.18$ (see the Supplemental Material for details on the model and model parameters). However, the memory precision (as estimated in terms of the standard deviation of a von Mises distribution) did not significantly differ between 12° probes ($M = 26.93°$, 95% CI = [22.31, 31.55]) and 28° probes ($M = 25.99°$, 95% CI = [21.37, 30.61]), $t(9) = 0.315$, $p = .760$, $d_z = 0.10$, 95% CI = [–0.52, 0.72], $BF_{01} = 3.10$, or between 28° and 45° probes ($M = 23.91°$, 95% CI = [19.29, 28.53]), $t(9) = 1.29$, $p = .230$, $d_z = 0.41$, 95% CI = [–0.25, 1.04], $BF_{01} = 1.68$. Even though this evidence for the absence of an effect on memory precision is only moderate or indecisive, respectively, it is clear that potential effects on precision cannot explain the overwhelming evidence for an effect of saliency on recall error ($BF_{+0} = 551.51$ and $BF_{+0} = 70.6$).
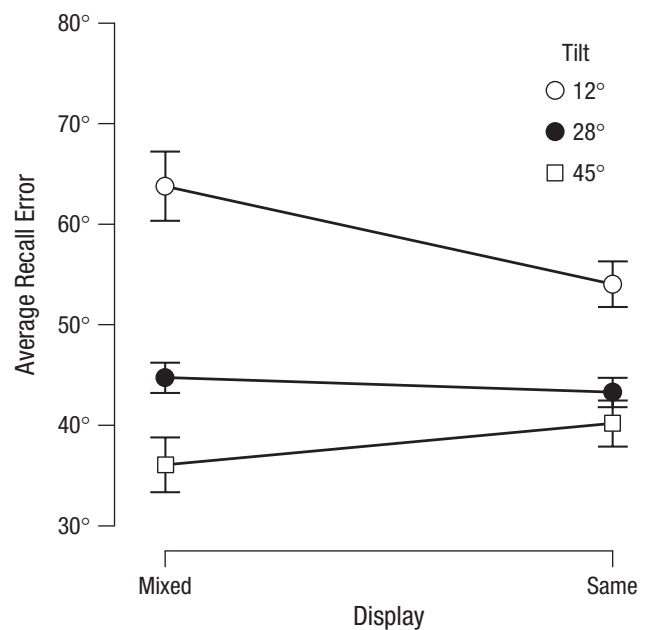
Experiment 2

Saliency might influence VWM processing in two non-exclusive ways. First, objects compete for VWM processing, so the most salient object within a display is eventually remembered best. This effect depends on the object's relation to other objects in the display, and we therefore refer to it as an effect of *relative saliency*.

Second, processing of more salient objects might be enhanced regardless of what else is in the display—an effect of *absolute saliency*. In visual search, the absolute saliency of a single target affects processing difficulty (Liesefeld et al., 2016; Nothdurft, 1993), but little is known about the effects of relative saliency with multiple target objects.

Method

To disentangle the two potential effects of saliency, we ran an experiment that compared the mixed-tilt displays of Experiment 1 with displays containing three bars of the same tilt. An effect of absolute saliency would predict that even in displays with only 12°-tilted bars (12° same-tilt displays), each 12°-tilted bar is remembered worse than each 45°-tilted bar in 45° same-tilt displays. If relative saliency contributed to the effect of saliency observed in Experiment 1, the 45°-tilted object was processed particularly well (beyond the effect of absolute saliency) by virtue of the other two tilted bars being less salient. Correspondingly, the 12°-tilted object then was processed particularly poorly because the other two tilted bars were more salient. By contrast, when all targets within a display are equally salient, the degree of VWM processing should be equal for all of them. This means that each 45°-tilted object in a display with only 45°-tilted objects among vertical bars would be remembered less well than the 45°-tilted object competing with the 28°- and 12°-tilted object in mixed-tilt displays. Conversely, each 12°-tilted object in a display with only 12°-tilted objects would be remembered better than the 12°-tilted object competing with the 28°- and 45°-tilted objects in Experiment 1. Thus, demonstrating that performance decreases from mixed- to same-tilt displays for 45°-tilted objects and increases for 12°-tilted objects would constitute proof of an influence of relative saliency on VWM performance.

In Experiment 2, the preregistered tests determining the stopping rule for the sequential testing procedure examined whether recall error would decrease with object saliency (as in Experiment 1) for both same- and mixed-tilt displays and also whether recall error would differ between same- and mixed-tilt displays even for the same probe tilt (see https://osf.io/d8t62 for the preregistration). We predicted an increase for 45° probes and a decrease for 12° probes. This resulted in a sample of 31 adults with normal or corrected-to-normal vision (age: $M = 26.4$ years, $SD = 5.44$; 25 female; four left-handed). Experiment 2 was modeled after Experiment 1, with the crucial difference being that one of two types of memory displays could be presented on each trial. Mixed-tilt displays were identical to the displays of Experiment 1 (see Fig. 1b) in all relevant

Fig. 3. Results from Experiment 2: average recall error (i.e., mean absolute difference between correct responses and given responses) as a function of display condition and target tilt. Error bars represent 95% within-participants confidence intervals.

aspects. Same-tilt displays were similar to mixed-tilt displays except that the tilted bars all shared the same tilt (12°, 28°, or 45°).

Each participant completed a total of 600 trials divided into blocks of 30 trials each. Each condition (i.e., Type of Display × Tilt of the Probe) was randomly presented 100 times.

Results

The mixed-tilt condition of Experiment 2 replicated the results of Experiment 1 (see Fig. 3): Recall error was higher for 12° probes ($M = 63.77°$, 95% CI = [58.94, 68.59]) than for 28° probes ($M = 44.74°$, 95% CI = [39.92, 49.57]), $t(30) = 10.57$, $p < .001$, $d_z = 1.90$, 95% CI = [1.30, 2.49], $BF_{+0} = 1.44 × 10^9$, and higher for 28° probes than for 45° probes ($M = 36.06°$, 95% CI = [31.24, 40.89]), $t(30) = 5.83$, $p < .001$, $d_z = 1.05$, 95% CI = [0.60, 1.48], $BF_{+0} = 1.68 × 10^4$.

Crucially, and as expected, performance was better for 12° probes, $t(30) = 6.02$, $p < .001$, $d_z = 1.08$, 95% CI = [0.63, 1.52], $BF_{+0} = 2.69 × 10^4$, and worse for 45° probes, $t(30) = −2.88$, $p = .004$, $d_z = −0.52$, 95% CI = [−0.89, −0.13], $BF_{−0} = 11.56$, in same-tilt compared with mixed-tilt displays. This difference was weak and unreliable only for 28° probes (for which we had no specific hypotheses, as mentioned in our preregistration; see https://osf.io/d8t62), $t(30) = 1.57$, $p = .128$, $d_z = 0.28$, 95% CI = [−0.08, 0.64], $BF_{01} = 1.75$. Indeed, VWM recall performance for a particular object depends on the

object's relative saliency with respect to the other objects in the scene.

Even though the effect of probed-target tilt was weaker for same- than for mixed-tilt displays, it was still present, indicating an effect of absolute saliency on top of the effect of relative saliency. In particular, recall error was higher for 12° same-tilt displays ($M = 54.02°$, 95% CI = [49.20, 58.85]) than for 28° same-tilt displays ($M = 43.29°$, 95% CI = [38.47, 48.12]), $t(30) = 7.79$, $p < .001$, $d_z = 1.40$, 95% CI = [0.90, 1.89], $BF_{+0} = 2.39 \times 10^6$, and higher for 28° same-tilt displays than for 45° same-tilt displays ($M = 40.19°$, 95% CI = [35.37, 45.02]), $t(30) = 3.10$, $p = .002$, $d_z = 0.56$, 95% CI = [0.17, 0.93], $BF_{+0} = 18.85$ (see Fig. 3). Replicating Experiment 1, results from the Zhang and Luck (2008) mixture model again showed that saliency mainly influenced p_{mem} in both same- and mixed-tilt displays (see the Supplemental Material).

Computational Modeling

One might argue that the observed effects of target tilt are not due to differential bottom-up saliency but, rather, to differential fit between each object and the top-down attentional template used to select target objects (e.g., Duncan & Humphreys, 1989; Geng & Witkowski, 2019). In particular, when participants look for tilted objects, their attentional template in our study might be matched best by the 45°-tilted object, followed by the 28°-tilted object, despite all objects being equally relevant. Such an attentional template might be optimal because it minimizes the match between the search template and the vertical (0°) nontarget objects, thus potentially minimizing interference (Geng & Witkowski, 2019).

Method

To test how well the two conflicting explanations account for the data from Experiment 2, we used computational modeling. In particular, we devised two novel models that implement the two potential accounts for the observed data pattern. First, the saliency model attempts to account for the data by a mixture of absolute and relative saliency; the degree to which relative saliency has an influence is a free parameter estimated from the data (w_{rel}). Second, the alternative optimal-template model posits that the different target bars differentially match the top-down template. Importantly, rather than deciding a priori on the value of the template, we included template tilt as a free parameter so that the optimization algorithm could estimate the unobservable template tilt from the observed behavioral data (for a detailed description of both models, see the Supplemental Material).
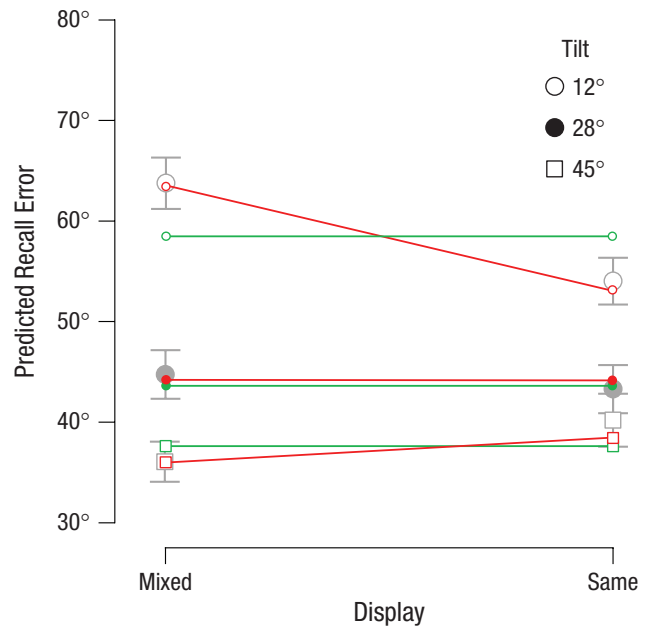
Fig. 4. Predictions of our preferred saliency model (red) and the alternative optimal-template model (green) as a function of display condition and target tilt. For comparison, mean empirical data are plotted in gray (error bars indicate ±1 *SE*).

Results

Comparing the fit of both models with the data of Experiment 2 (see Fig. 4), we found that the saliency model well outperformed the optimal-template model. In particular, the optimal-template model failed to account for the difference between same- and mixed-tilt displays. Thus, performance in Experiment 2 is best explained by variation in saliency. Notably, to account for the data, the saliency model has to assume a positive influence of relative saliency ($w_{rel} > 0$), thus providing further support for this novel assumption.

Experiment 3

The model was devised after Experiment 2 was conducted and analyzed to rule out the possibility (brought forward by a reviewer) that our data might also be explained by participants employing some attentional template. To additionally provide an empirical test with a priori hypotheses, we preregistered and ran Experiment 3 (see https://osf.io/f9c72 for the preregistration). We reasoned that if differential fit between the objects and an attentional template explains our results, the effect of tilt should remain when the vertical bars are removed (clean displays[1]; see Fig. 1d) because the tilted bars still differentially fitted to this assumed attentional template. By contrast, our explanation in terms of saliency predicts that removing the task-irrelevant

vertical bars renders all tilted bars highly and almost equally salient because local feature contrast is high for all three bars when presented in isolation (see the Method section for Experiment 1). In contrast to the cluttered displays of Experiment 1, clean displays should result in a strong decrease or even a complete absence of the effect of tilt.

Method

Experiment 3 was conducted online (for details, see the Supplemental Material). The preregistered *t* tests determining the stopping rule for the sequential testing procedure examined whether recall error would decrease with object saliency in displays with vertical nontarget bars (cluttered displays; mixed-tilt targets as in Experiments 1 and 2) and whether the effect of tilt was lower in clean displays compared with cluttered displays. A third noncritical hypothesis was that the effect of tilt might be fully absent in clean displays. This sequential testing procedure (for details, see the Supplemental Material and the preregistration) resulted in a sample of 60 adults with normal or corrected-to-normal vision (age: $M = 25.6$ years, $SD = 6.20$; 23 female; four left-handed).

Experiment 3 was modeled after Experiment 1, with the critical difference being that one of two types of memory displays could be presented on each trial. Cluttered displays were identical to the displays of Experiment 1 (see Fig. 1b) in all relevant aspects. Clean displays contained only the three tilted bars (i.e., the task-irrelevant vertical nontarget bars were removed) but were otherwise identical to cluttered displays. Each participant completed a total of 150 trials divided into blocks of 50 trials. Each condition (i.e., Type of Display × Tilt of the Probe) was randomly presented 25 times.

Results

For cluttered displays, we replicated the results of Experiments 1 and 2 (mixed-tilt displays; see Fig. 5): Recall error was higher for 12° probes ($M = 71.64°$, 95% CI = [68.10, 75.18]) than for 28° probes ($M = 48.56°$, 95% CI = [45.02, 52.10]), $t(59) = 11.74$, $p < .001$, $d_z = 1.52$, 95% CI = [1.14, 1.88], $BF_{+0} = 2.63 \times 10^{14}$, and higher for 28° probes than for 45° probes ($M = 35.30°$, 95% CI = [31.76, 38.84]), $t(59) = 6.11$, $p < .001$, $d_z = 0.79$, 95% CI = [0.50, 1.08], $BF_{+0} = 3.18 \times 10^5$. Crucially, and as expected, the effect of tilt decreased in clean displays compared with cluttered displays for 12° probes compared with 28° probes, $t(59) = -10.01$, $p < .001$, $d_z = -1.29$, 95% CI = [−1.63, −0.95], $BF_{-0} = 2.69 \times 10^4$, and for 28° probes compared with 45° probes, $t(59) = -5.06$, $p < .001$, $d_z = -0.65$, 95% CI = [−0.93, −0.37], $BF_{-0} = 8024.60$. Finally, there was no significant effect of tilt,
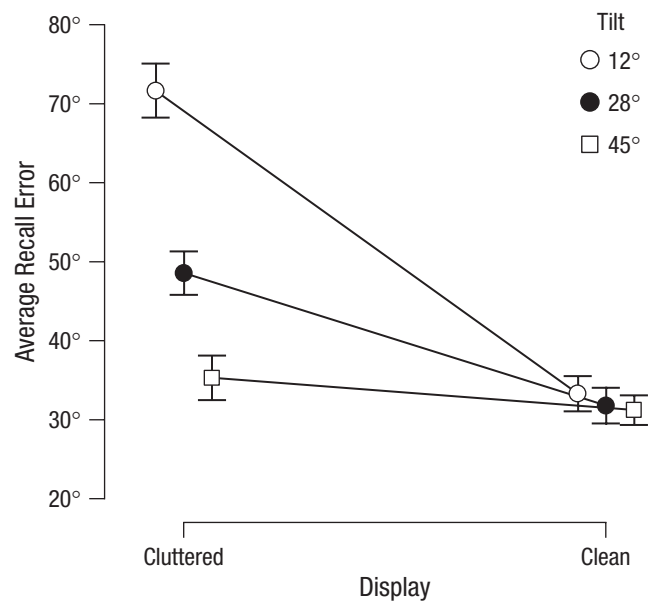
Fig. 5. Results from Experiment 3: average recall error (i.e., mean absolute difference between correct responses and given responses) as a function of display condition and target tilt. Error bars represent 95% within-participants confidence intervals.

and there was even some evidence for the absence of this effect, in clean displays for 12° probes ($M = 33.31°$, 95% CI = [29.78, 36.85]) compared with 28° probes ($M = 31.79°$, 95% CI = [28.25, 35.33]), $t(59) = 1.17$, $p = .247$, $d_z = 0.15$, 95% CI = [−0.10, 0.40], $BF_{01} = 3.71$, and moderate evidence for the absence of an effect for 28° probes compared with 45° probes ($M = 31.23°$, 95% CI = [27.70, 34.77]), $t(59) = 0.46$, $p = .650$, $d_z = 0.06$, 95% CI = [−0.19, 0.31], $BF_{01} = 6.41$. This pattern indicates that the effect of target tilt is not due to differential match between the objects and an attentional template but, rather, due to variation in saliency.

General Discussion

We set out to demonstrate an influence of saliency on performance in a VWM task, an influence that has not yet been acknowledged in any current model of VWM processing. Experiment 1 indeed provided overwhelming evidence for the existence of this effect by showing that how well an object's color is remembered is largely determined by how much it differs in tilt from its immediate surroundings (local feature contrast). Experiment 2 demonstrated that both relative and absolute saliency contribute to the effect of saliency. Finally, a newly devised computational model and Experiment 3 demonstrated that the effect of target tilt is indeed explained by saliency rather than differential fit between each object and some attentional template. How saliencies of multiple relevant objects interact has—to the best of

our knowledge—not yet been systematically examined, and an observation of an effect of relative saliency is therefore new not only for the VWM community but also for the visual cognition community in general.

Many theories of visual search (e.g., Duncan & Humphreys, 1989; Fecteau & Munoz, 2006; Liesefeld & Müller, 2019b, 2020; Wolfe, 2021) assume a preattentive spatial representation of the visual scene coding for relevance at each location and informing a second, attentive-processing stage. This assumption is needed to explain how second-stage focal attention can be allocated to the most promising objects in view without analyzing each object in detail first. This preattentive *priority map* is thought to be influenced by task goals and experiences (top-down) as well as saliency (bottom-up). We propose that the very same priority map supporting visual search might also determine VWM processing (Bundesen et al., 2011; Liesefeld et al., 2020). Findings from the present study and those manipulating each object's relevance (e.g., Emrich et al., 2017) can be integrated using the priority-map concept: Although previous studies manipulated top-down influences, we are the first to systematically manipulate bottom-up contributions (i.e., saliency) to preattentive-priority-map activations in a VWM task.

There are many potential mechanisms by which first-stage priority (and, thus, saliency) could theoretically impact second-stage VWM processing. First, it might influence VWM encoding directly (in particular without the allocation of focal attention) or via the allocation of an attentional resource (Emrich et al., 2017). Second, encoding and attention allocation could be conceived of as serial (one object is processed or attended after the other) or parallel (all objects are processed or attended at once; Bundesen et al., 2011; Sewell et al., 2014). Third, priority might affect how much (if any) information about each object is processed or how much of a limited (quantized or continuous) VWM resource it receives (Ma et al., 2014; Vogel et al., 2006). Fourth, priority might additionally influence third-stage postselective and postencoding processes, such as how fast attention is disengaged from a processed object to continue with the next object in line (see Sauter et al., 2021) or how well a processed object is stored (e.g., by attaining a special state; Oberauer & Lin, 2017; Olivers et al., 2011). All kinds of combinations between these mechanisms seem theoretically possible, and we will speculate on some in turn.

Our exploratory Zhang and Luck (2008) mixture-model analysis indicated that saliency mainly affects whether an item is encoded (p_{mem}) rather than the precision of encoding (standard deviation of the von Mises distribution). If mixture modeling is valid (for a critical view, see Ma, 2018), this finding somewhat constrains the range of potential mechanisms by which saliency (as represented on a first-stage, preattentive priority map) is translated into VWM performance: If, at the second stage, all objects are processed in parallel, one would assume that information on each object accrues continually with a slower rate for less salient objects (e.g., Moran et al., 2016). The mixture-modeling finding would then indicate that an object is stored in full when a certain amount of information is accumulated (Bundesen et al., 2011) because, otherwise, we should have observed an effect on memory precision. Alternatively, second-stage encoding might proceed serially, starting at the most salient target object and sometimes not reaching the least salient target object (Wolfe, 2021; e.g., because focal attention needs to be allocated sequentially to encode each object).

Another implication from our study is that previous studies might have unintentionally induced and misinterpreted disguised effects of saliency. Data from the same-tilt condition of Experiment 2 indicate that less information was remembered in low-saliency compared with high-saliency displays (the effect of absolute saliency). One could easily misinterpret this effect as a decrease in VWM capacity from high- to low-saliency displays. However, this would be theoretically awkward because a fixed limit is the core assumption behind both slot theories and flexible-resource theories of VWM alike (for an alternative, see Oberauer & Lin, 2017). Actually, this effect recalls other findings that processing difficulty of an object class correlates with how many objects of that class can be held in VWM: Manipulating object complexity, Alvarez and Cavanagh (2004) showed that visual search rate (as a measure of processing difficulty) predicts VWM capacity for the respective object class. They argued that search rate and VWM capacity were related by the objects' informational content, which would affect how long it takes to process each item in visual search and how much of the limited VWM capacity it consumes. In light of the present results, it seems equally likely, though, that the two measures are more directly related by the saliency-dependent ease of processing each object. For example, processing of the first low-saliency/high-complexity objects might take so long (see the third-stage mechanisms discussed above) that on some trials, no time is left to process the remaining objects in the display (e.g., in our same-tilt displays, only two of the three 12° objects might have been processed on some trials). Crucially, in our study, this cannot be explained by the to-be-encoded informational content (which was the same for each object) but must be due to the saliency of the object carrying that information. Thus, effects of object complexity on VWM performance observed in earlier studies might alternatively be explained as effects of saliency. More complex objects might be less salient in their respective displays and

therefore take longer to process (irrespective of their informational content). Along similar lines, our findings might trigger reevaluations of further influential findings from VWM studies using relatively complex stimuli.

ORCID iDs

Martin Constant https://orcid.org/0000-0001-9574-0674
Heinrich R. Liesefeld https://orcid.org/0000-0003-4551-8607

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620975785

Note

1. We thank Nelson Cowan for suggesting these displays at the Virtual Working Memory 2020 Symposium.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111. https://doi.org/10/b88gs7

Alvarez, G. A., & Cavanagh, P. (2008). Visual short-term memory operates more efficiently on boundary features than on surface features. *Perception & Psychophysics*, *70*(2), 346–364. https://doi.org/10/dbs6k7

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10/f34nps

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2011). A neural theory of visual attention and short-term memory (NTVA). *Neuropsychologia*, *49*(6), 1446–1457. https://doi.org/10/b8863z

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10/ddq83h

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10/ck4p2t

Emrich, S. M., Lockhart, H. A., & Al-Aidroos, N. (2017). Attention mediates the flexible allocation of visual working memory resources. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(7), 1454–1465. https://doi.org/10/gbn3xj

Fecteau, J., & Munoz, D. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. https://doi.org/10/fb66m9

Fukuda, K., Vogel, E., Mayr, U., & Awh, E. (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. https://doi.org/10.3758/17.5.673

Geng, J. J., & Witkowski, P. (2019). Template-to-distractor distinctiveness regulates visual search efficiency. *Current Opinion in Psychology*, *29*, 119–125. https://doi.org/10/gg5vtp

Hollingworth, A. (2008). Visual memory for natural scenes. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 123–162). Oxford University Press.

Klink, P. C., Jeurissen, D., Theeuwes, J., Denys, D., & Roelfsema, P. R. (2017). Working memory accuracy for multiple targets is driven by reward expectation and stimulus contrast with different time-courses. *Scientific Reports*, *7*(1), Article 9082. https://doi.org/10/gbtwp7

Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. *The Journal of Neuroscience*, *34*(30), 9857–9866. https://doi.org/10/f6fskg

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9–16. https://doi.org/10/bjs32m

Liesefeld, A. M., Liesefeld, H. R., & Zimmer, H. D. (2014). Intercommunication between prefrontal and posterior

brain regions for protecting visual working memory from distractor interference. *Psychological Science*, *25*(2), 325–333. https://doi.org/10/f5344z

Liesefeld, H. R., Liesefeld, A. M., Pollmann, S., & Müller, H. J. (2018). Biasing allocations of attention via selective weighting of saliency signals: Behavioral and neuroimaging evidence for the dimension-weighting account. In T. Hodgson (Ed.), *Processes of visuospatial attention and working memory* (Vol. 41, pp. 87–113). Springer International Publishing. https://doi.org/10/dc5q

Liesefeld, H. R., Liesefeld, A. M., Sauseng, P., Jacob, S. N., & Müller, H. J. (2020). How visual working memory handles distraction: Cognitive mechanisms and electrophysiological correlates. *Visual Cognition*, *28*(5–8), 372–387. https://doi.org/10/gg5vsv

Liesefeld, H. R., Liesefeld, A. M., Töllner, T., & Müller, H. J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *NeuroImage*, *156*, 166–173. https://doi.org/10/gbsjqj

Liesefeld, H. R., Moran, R., Usher, M., Müller, H. J., & Zehetleitner, M. (2016). Search efficiency as a function of target saliency: The transition from inefficient to efficient search and beyond. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 821–836. https://doi.org/10/ggbnjc

Liesefeld, H. R., & Müller, H. J. (2019a). Current directions in visual working memory research: An introduction and emerging insights. *British Journal of Psychology*, *110*(2), 193–206. https://doi.org/10/gfvm2p

Liesefeld, H. R., & Müller, H. J. (2019b). Distractor handling via dimension weighting. *Current Opinion in Psychology*, *29*, 160–167. https://doi.org/10/ggbnjd

Liesefeld, H. R., & Müller, H. J. (2020). Modulations of saliency signals at two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology: General*. Advance online publication. https://doi.org/10.1037/xge0000970

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10/39v

Ma, W. J. (2018). *Problematic usage of the Zhang and Luck mixture model*. bioRxiv. https://doi.org/10/ggbpq3

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10/gd8cbs

Moran, R., Zehetleitner, M., Liesefeld, H. R., Müller, H. J., & Usher, M. (2016). Serial vs. parallel models of attention in visual search: Accounting for benchmark RT-distributions. *Psychonomic Bulletin & Review*, *23*(5), 1300–1315. https://doi.org/10/gbgdk6

Nothdurft, H.-C. (1993). The role of features in preattentive vision: Comparison of orientation, motion and color cues.

Vision Research, *33*(14), 1937–1958. https://doi.org/10/d6fw5n

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review*, *124*(1), 21–59. https://doi.org/10/f9mhb2

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*(7), 327–334. https://doi.org/10/fcfxg6

Rajsic, J., Sun, S. Z., Huxtable, L., Pratt, J., & Ferber, S. (2016). Pop-out and pop-in: Visual working memory advantages for unique items. *Psychonomic Bulletin & Review*, *23*(6), 1787–1793. https://doi.org/10/f9hf36

Santangelo, V. (2015). Forced to remember: When memory is biased by salient information. *Behavioural Brain Research*, *283*, 1–10. https://doi.org/10/f67dgp

Sauter, M., Hanning, N. M., Liesefeld, H. R., & Müller, H. J. (2021). Post-capture processes contribute to statistical learning of distractor locations in visual search. *Cortex*, *135*, 108–126. https://doi.org/10.1016/j.cortex.2020.11.016

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10/gcsk2r

Sewell, D. K., Lilburn, S. D., & Smith, P. L. (2014). An information capacity limitation of visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(6), 2214–2242. https://doi.org/10/gg5fn7

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retrocue effect. *Attention, Perception, & Psychophysics*, *78*(7), 1839–1860. https://doi.org/10/f83nvs

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/10/drbpg5

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1436–1451. https://doi.org/10/c2vrp7

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*. Advance online publication. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), Article 0058. https://doi.org/10/gftcnc

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10/c693sk