The role of saliency for visual working memory in complex visual scenes

Martin Constant & Heinrich R. Liesefeld

Scan for more information

Department of Psychology, Ludwig-Maximilians-Universität München
Correspondence: Martin.Constant@psy.lmu.de

Highlights

- **Saliency enhances VWM performance with salient items being recalled more precisely than non-salient items.**
- **This effect scales with the saliency of items: the more salient an item, the better it is recalled, especially when among less salient items.**
- **Increasing the encoding time decreases but does not fully erase the effect.**
- **In the right conditions, the effect of saliency can be counteracted by top-down control**

Background

We know from the visual search literature that salient targets are typically found and processed faster than non-salient targets. Target orientation with regards to the background has proven to be a reliable way to scale saliency. In typical **visual working memory** (**VWM**) studies, targets are often isolated and therefore highly salient; but more complex displays can contain big saliency variations. Finally, VWM literature has not yet evaluated the **influence of saliency** in a systemic manner.

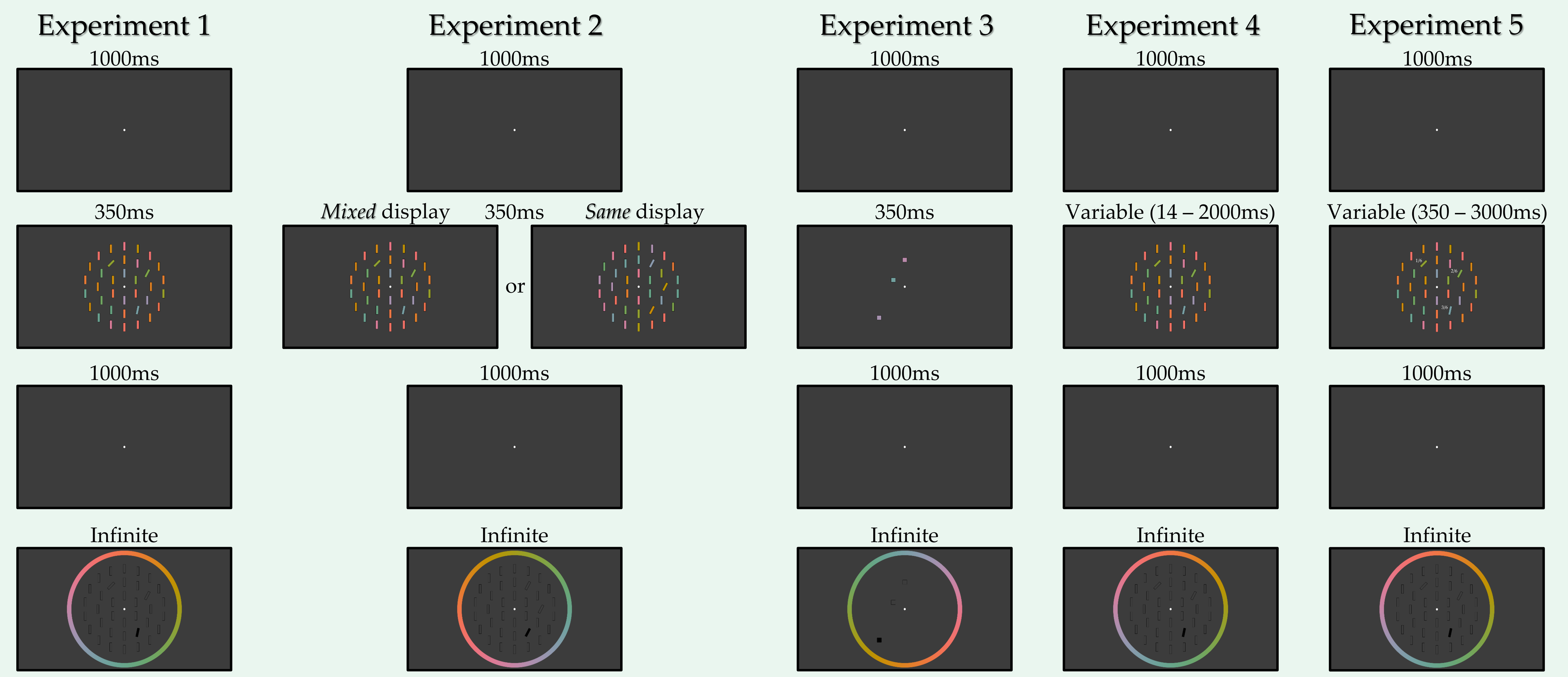
Abstract

Given its severe capacity limitations, visual working memory can process only a tiny fraction of the complex visual world. While selection of relevant information from cluttered scenes is a main topic of research on visual attention, it has not received much research efforts in the VWM community. Based on knowledge from visual-attention research, we developed a task which approaches the complexity of real-world scenes while maintaining tight experimental control over stimulation. Participants were presented with an array of 33 vertical bars and 3 tilted target bars (12°, 28° and 45°). After presentation, one of the targets was probed and participant had to recall its color (continuous report). In a first experiment, we provide evidence that VWM performance is parametrically influenced by saliency. In a second experiment, by assigning the same tilt to targets within a trial, we showed that the performance is not only influenced by the saliency difference between targets, but also by their *absolute* saliency. The second experiment also replicated results of the first. Our third experiment used typical displays (three squares) and demonstrated that – in contrast to the real world – saliency is virtually maxed out for relevant objects in typical laboratory studies of VWM, likely yielding a pronounced underestimation of this major influence on VWM. A fourth experiment examined the influence of saliency across 7 different encoding times and showed that encoding time has an influence on the effect of saliency, but also that, even after encoding for 2 seconds this effect remains present, showing that our manipulation of saliency are difficult to overcome. The fifth experiment increased the relevance of less salient targets by probing them more often. After 2 seconds (but not 350ms) of encoding time, this manipulation was effective and salient targets were not recalled more accurately demonstrating that even when participants deliberately tried to override the strong effect of saliency, it still took them a considerable amount of time. Even then, they were unable to revert the effect in line with task goals, but merely compensated for it.

Methods

- Participants had to remember the color of the targets, with a **continuous recall paradigm**
- **Saliency** was determined by the targets' tilt (more tilted targets were more salient; Exp. 1, 2, 4 & 5) or their eccentricity (targets closer to fixation were more salient; Exp. 3)
- **Recall error** was measured in **absolute angular distance** between the color chosen by the participant and the actual color of the target
- Sample sizes were determined by **sequential Bayes factor** testing

Experimental design



Experiment 1

- 33 irrelevant vertical bars
- 3 target bars with each a different tilt (12°, 28° & 45°)
- 350ms presentation time

Experiment 2

- Two types of displays, randomly presented
- *Mixed* displays (left) → identical to Exp. 1 displays
- *Same* displays (right) → all targets share the same tilt
- 350ms presentation time

Experiment 3

- 3 target squares
- One square at each eccentricity (2°, 4° & 6°)
- 350ms presentation time

Experiment 4

- Displays identical to Exp. 1
- Presentation time: 14ms, 50ms, 100ms, 350ms, 500ms, 1000ms or 2000ms

Experiment 5

- Displays identical to Exp. 1
- Presentation time: 350ms, 2000ms or 3000ms
- 12° target probed on 3/6 of trials
- 28° probed on 2/6
- 45° probed on 1/6

Paired samples t tests table

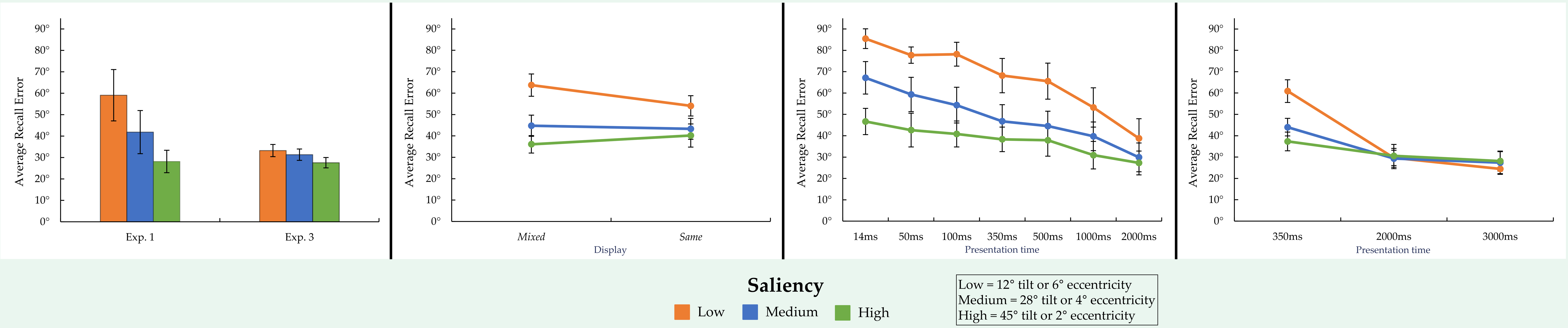
Comparison	t	df	d_z	BF	Favors
Experiment 1					
12° > 28°	6.56***	9	2.07 [0.93, 3.19]	551.51	H_+
28° > 45°	4.66***	9	1.47 [0.54, 2.37]	70.60	H_+
Experiment 2 *Mixed* displays					
12° > 28°	10.57***	30	1.90 [1.30, 2.49]	1.44e+9	H_+
28° > 45°	5.83***	30	1.05 [0.60, 1.48]	1.68e+4	H_+
Experiment 2 *Same* displays					
12° > 28°	7.79***	30	1.40 [0.90, 1.89]	2.39e+6	H_+
28° > 45°	3.10**	30	0.56 [0.17, 0.93]	18.85	H_+
Experiment 2 *Mixed* vs. *Same* displays					
12° – *Mixed* > *Same*	6.02***	30	1.08 [0.63, 1.52]	2.69e+4	H_+
28° – *Mixed* ≠ *Same*	1.57	30	0.28 [−0.08, 0.64]	1.75	H_0
45° – *Mixed* < *Same*	2.88**	30	0.52 [0.13, 0.89]	11.56	H_-
Experiment 3					
6° > 4°	2.56**	33	0.44 [0.08, 0.79]	6.04	H_+
4° > 2°	4.78***	33	0.82 [0.43, 1.21]	1308.45	H_+

Note. *$p < .05$, **$p < .01$, ***$p < .001$

Comparison	t	df	d_z	BF	Favors
Experiment 4					
14ms – 12° > 28°	4.89***	15	1.22 [0.56, 1.87]	316.33	H_+
14ms – 28° > 45°	11.37***	15	2.84 [1.71, 3.95]	2.59e+6	H_+
50ms – 12° > 28°	5.54***	15	1.38 [0.68, 2.07]	935.97	H_+
50ms – 28° > 45°	7.66***	15	1.91 [1.07, 2.74]	2.49e+4	H_+
100ms – 12° > 28°	7.57***	15	1.89 [1.05, 2.71]	2.20e+4	H_+
100ms – 28° > 45°	5.95***	15	1.49 [0.76, 2.20]	1851.30	H_+
350ms – 12° > 28°	7.21***	15	1.80 [0.99, 2.60]	1.30e+4	H_+
350ms – 28° > 45°	3.02**	15	0.76 [0.19, 1.30]	12.36	H_+
500ms – 12° > 28°	5.93***	15	1.48 [0.75, 2.19]	1774.43	H_+
500ms – 28° > 45°	2.45*	15	0.61 [0.07, 1.14]	4.82	H_+
1000ms – 12° > 28°	4.68***	15	1.17 [0.52, 1.80]	218.08	H_+
1000ms – 28° > 45°	3.29**	15	0.82 [0.24, 1.38]	19.39	H_+
2000ms – 12° > 28°	3.50**	15	0.87 [0.28, 1.44]	28.09	H_+
2000ms – 28° > 45°	1.69	15	0.42 [−0.10, 0.93]	1.53	H_+
Experiment 5					
350ms – 12° > 28°	10.32***	35	1.72 [1.20, 2.23]	4.72e+9	H_+
350ms – 28° > 45°	4.23***	35	0.70 [0.33, 1.07]	323.81	H_+
350ms – 12° > 45°	4.23***	35	0.70 [0.33, 1.07]	1.77e+8	H_+
2000ms – 12° > 28°	0.15	35	0.02 [−0.30, 0.35]	4.97	H_0
2000ms – 28° > 45°	−0.96	35	−0.16 [−0.49, 0.17]	10.14	H_0
2000ms – 12° > 45°	−0.36	35	−0.06 [−0.39, 0.27]	7.21	H_0
3000ms – 12° > 28°	−1.31	35	−0.22 [−0.55, 0.11]	11.91	H_0
3000ms – 28° > 45°	−0.49	35	−0.08 [−0.41, 0.25]	7.81	H_0
3000ms – 12° > 45°	−1.79	35	−0.30 [−0.63, 0.04]	14.35	H_0

Note. *$p < .05$, **$p < .01$, ***$p < .001$

Experiments 1 (N = 10) and 3 (N = 34) · Experiment 2 (N = 31) · Experiment 4 (N = 16) · Experiment 5 (N = 36)



Results → Conclusion

- In Experiment 1, more tilted targets were recalled more precisely → VWM performance improves with target saliency
- Results of Experiment 1 were replicated with the *Mixed* displays of Experiment 2
- Performance was better for 12° targets and worse for 45° targets in *Same* compared to *Mixed* displays → VWM performance for a specific target depends on its *relative* (i.e. competition with other targets) saliency
- For the *Same* displays, performance was lower in displays with targets that were less salient → *Absolute* (i.e. regardless of other targets) saliency also enhances VWM performance
- In Experiment 3, less eccentric squares were recalled more precisely → Eccentricity impacts VWM performance
- In Experiment 4, targets were recalled more precisely with more encoding time → Encoding time influences VWM performance
- Performance for the 45° target at 14ms of encoding time was better than performance for 12° targets at all but the longest encoding time → Saliency can enhance performance almost instantaneously
- 28° targets were recalled more precisely than 12° targets at 2000ms of encoding time → The effect of saliency was not fully overridden by the effect of encoding time
- In Experiment 5, the effect of saliency was still present at 350ms → Top-down control did not have enough time to counteract saliency
- After 2000ms, the effect of saliency was not observed → The time needed for top-down control to deploy is far longer than what would be predicted by the visual search literature
- At 3000ms, more relevant targets (i.e. less salient) were not recalled more precisely than less relevant ones → This manipulation was not strong enough to fully override saliency

This work is licensed under a "Creative Commons Attribution 4.0 International" license.