

**Visual Cognition** 



ISSN: 1350-6285 (Print) 1464-0716 (Online) Journal homepage: https://www.tandfonline.com/loi/pvis20

# Working memory-driven attention towards a distractor does not interfere with target feature perception

Emma Wu Dowd, Samoni Nag & Julie D. Golomb

**To cite this article:** Emma Wu Dowd, Samoni Nag & Julie D. Golomb (2019): Working memorydriven attention towards a distractor does not interfere with target feature perception, Visual Cognition, DOI: <u>10.1080/13506285.2019.1659895</u>

To link to this article: <u>https://doi.org/10.1080/13506285.2019.1659895</u>

+

View supplementary material 🖸



Published online: 05 Sep 2019.

|--|

Submit your article to this journal  $\square$ 

Article views: 29



🔾 View related articles 🗹

🌔 View Crossmark data 🗹

# Working memory-driven attention towards a distractor does not interfere with target feature perception

Emma Wu Dowd 💿, Samoni Nag and Julie D. Golomb 💿

Department of Psychology, The Ohio State University, Columbus, OH, USA

#### ABSTRACT

The contents of working memory (WM) can influence where we attend – but can it also interfere with what we see? Active maintenance of visual items in WM biases attention towards WM-matching objects, and also enhances early perceptual processing of WM-matching items (e.g., more accurate perceptual discrimination). Here, we asked whether a WM-matching distractor *interferes* with perceptual processing of a target's features. In a dual-task paradigm, participants maintained a shape in WM across an intervening visual search task, during which they had to reproduce the colour of a designated target item using a continuous-report technique. Importantly, the WM shape could match the target item, a distractor item, or no item in the search array. When the WM shape matched a distractor, we found no evidence of systematic perceptual interference (i.e., swapping or mixing with the distractor colour), but observed only general disruptions in target processing (i.e., decreased target accuracy). These results suggest that when visual attention is inadvertently drawn to a WM-matching distractor, any resultant automatic perceptual processing may be too transient or weak to significantly interfere with perceptual processing of the target's features.

ARTICLE HISTORY

Received 21 August 2018 Accepted 16 August 2019

Routledge

Taylor & Francis Group

Check for updates

**KEYWORDS** Working memory; attentional bias; feature perception

Visual attention can be biased towards objects in the environment that match the contents of working memory (WM), even when information in WM is unrelated to or at the expense of current task goals (e.g., Downing, 2000; Olivers, Meijer, & Theeuwes, 2006; Soto, Heinke, Humphreys, & Blanco, 2005). Such WMdriven attentional effects have been demonstrated in numerous studies (see Kiyonaga & Egner, 2013; Olivers, Peters, Houtkamp, & Roelfsema, 2011; Soto, Hodsoll, Rotshtein, & Humphreys, 2008; Woodman, Carlisle, & Reinhart, 2013 for reviews), typically using a dual-task paradigm in which participants remember an item (e.g., a coloured shape) while performing an intervening but unrelated visual search task during the delay period between WM cue and subsequent probe. Critically, the WM item can reappear in the search display, either coinciding with the location of a target (i.e., "valid") or with the location of a distractor (i.e., "invalid"), or fail to reappear in the display at all (i.e., "neutral"). The canonical finding is that search response times are faster when the WM item validly cues the target location, and slower when the WM

item invalidly cues a distractor location (Soto et al., 2008). Even when WM contents can never match the target, such that WM-driven attention is strategically detrimental for search, response times are nevertheless slowed when WM-matching distractors are present (Carlisle & Woodman, 2011; Dowd, Kiyonaga, Beck, & Egner, 2015a; Han, 2015b; Kiyonaga, Egner, & Soto, 2012; Olivers et al., 2006; Soto et al., 2005; but see Dowd, Kiyonaga, Egner, & Mitroff, 2015b; Han & Kim, 2009; Woodman & Luck, 2007). Oculomotor responses also reflect these validity effects, such that more first saccades land on or are directed towards WM-matching items in the search display (e.g., Silvis, Belopolsky, Murris, & Donk, 2015), even when such trajectories are detrimental for search (e.g., Olivers et al., 2006; Soto et al., 2005). These findings are generally interpreted as visual attention being biased towards - or even captured by - WMmatching items.

More recently, WM-driven attention has been shown to enhance perceptual processing at WMmatching locations (Han, 2015a; Pan, Cheng, & Luo,

CONTACT Emma Wu Dowd 😒 emma.wu.dowd@austin.utexas.edu 🗈 Department of Psychology, The Ohio State University, 1835 Neil Ave, Columbus, OH 43210, USA

Supplemental data for this article can be accessed https://doi.org/10.1080/13506285.2019.1659895

 $<sup>\</sup>ensuremath{\mathbb{C}}$  2019 Informa UK Limited, trading as Taylor & Francis Group

2012; Pan, Luo, & Cheng, 2016; Soto, Wriglesworth, Bahrami-Balani, & Humphreys, 2010). Rather than comparing speeded response times, these studies emphasize the accuracy or sensitivity of target discrimination (e.g., which direction is the target line tilted?) in search displays that were presented very briefly (27 ms to 164 ms). Perceptual discrimination (in terms of both percentage correct and signal detection measures) was significantly improved for targets appearing at WM-matching locations (Han, 2015a; Soto et al., 2010). Importantly, these enhancements were found even in the absence of explicit awareness (Pan et al., 2012) or in the absence of external noise (i.e., distractors or masks; Pan et al., 2016), underscoring the "perceptual enhancement" theory that WMdriven attention automatically boosts *early* perceptual processing (compared to later selection or decision criteria only; Soto et al., 2010; but see Cosman & Vecera, 2011). Similarly, in a dual-task WM and saccade orienting paradigm, eye movements to a WM-matching saccade target were faster and more spatially accurate than to a neutral target, suggesting that initial perceptual salience of the saccade target is modulated by the contents of WM (Hollingworth, Matsukura, & Luck, 2013).

Evidence for the perceptual enhancement theory of WM-driven attention has been primarily limited to whether target processing is enhanced - i.e., validly cueing a target location enhances target discrimination. But what if the contents of WM match a distractor location? A perceptual enhancement account should predict enhanced perception of the WMmatching distractor - which might consequently interfere with perception of the target. However, in previous studies, invalidly cueing a distractor location did not adversely impact perceptual processing of the target, at least in terms of accuracy or signal detection sensitivity (Han, 2015a; Pan et al., 2016). This asymmetry stands in marked contrast to WM effects on response times, which are both significantly speeded by WM-matching targets and significantly slowed by WM-matching distractors (e.g., Kiyonaga et al., 2012). To reconcile these effects and gain further insight into WM-driven effects on perception, the current study approaches perceptual processing from a new perspective: probabilistic modeling of continuous feature perception.

We adapted a recent paradigm for measuring distortions in feature perception (Golomb, L'Heureux, & Kanwisher, 2014), in which participants were cued to shift or split spatial attention between two locations, before reporting the colour of an item at a target location. Rather than use a binary discrimination task, colour responses were selected from a continuous colour wheel, encouraging finer-grained estimates of perceptual encoding (Wilken & Ma, 2004). By applying probabilistic mixture models to the distribution of colour responses, Golomb et al. (2014) found that explicitly cueing attention to a nontarget (distractor) location interfered with perception of the target's features, such that distractor features were sometimes erroneously reported instead of the target's (e.g., "swapping errors") or caused more subtle distortions of the target feature (e.g., reporting a blend of target and distractor colours; "mixing errors"). Here, we ask whether target feature perception is likewise disrupted by WM-driven attention to a distractor location.

The current dual-task paradigm combined WM and a continuous feature report: Participants remembered a specific shape (e.g., a triangle) across an intervening visual search, in which they reported the colour of a target item (by selecting from a colour wheel). The contents of WM (shapes) were orthogonal and thus irrelevant to the perceptual colour-report task. However, all items in the search array had distinct shapes, such that the WM shape could match the target item (valid), a distractor item (invalid), or no search item at all (neutral). We expected visual attention to be biased towards WM-matching shapes, and in line with previous studies, for WM-driven attention to enhance feature perception of targets appearing at a WM-matching location (i.e., on valid trials). Our critical guestion was, how does a WM-matching distractor interfere with the processing of a target feature? One hypothesis is that WM-driven attention to a non-target location automatically encodes and enhances the feature of that distractor, resulting in perceptual interference of the target – which could manifest as erroneous misreports of the WM-matching distractor feature (swapping errors) or blending between target and distractor features (mixing errors). Alternatively, WM-driven attention away from the target location might simply disrupt target processing more generally, resulting in impaired encoding of the target feature (i.e., impaired accuracy or precision) without specific perceptual interference.

#### **Experiment 1**

#### Method

#### **Participants**

Twenty-three participants (ages 18–22; 6 male) were recruited from the undergraduate Research Experience Program pool at The Ohio State University. Four additional participants were excluded, 3 for poor WM task performance (<70% WM accuracy) and 1 for not successfully performing the colourreport task (>50% probability of guessing on neutral trials: pU from the probabilistic model described below). All participants reported normal or corrected-to-normal visual acuity and colour vision, received course credit, and provided informed consent in accordance with The Ohio State University institutional review board.

#### Stimuli & procedure

All stimuli were presented against a black background on a 21-inch flatscreen CRT monitor with a refresh rate of 85 Hz and screen resolution of  $1280 \times 1024$  pixels, using Matlab and the Psychophysics Toolbox (Brainard, 1997). The monitor was colour-calibrated with a Minolta CS-100 colorimeter. Subjects were positioned with a chinrest approximately 60 cm from the monitor in a dimly lit room, and eye position was monitored with an EyeLink 1000 eye-tracking system.

Figure 1A depicts an example trial sequence. Each trial began with the presentation of a central fixation  $0.48^{\circ} \times 0.48^{\circ}$ ). cross (white, Participants were instructed to fixate centrally at the start of each trial, but they were not instructed to maintain central fixation throughout the whole trial. Once participants had accurately fixated for 500 ms (i.e., maintain eye position within a 2°-radius of central fixation), as determined by real-time eye-tracking, the WM cue item was presented at the centre of the screen under the word "Remember!" for 1000 ms. The WM cue was a single white shape (circle, square, diamond, pentagon, hexagon, or triangle; 4°×4° area), presented at a random rotation (steps of 45°). Participants were instructed to remember the shape for a subsequent test. The WM cue was followed by a 1000-ms delay period in which only the fixation cross was visible on the screen, and then by the search array.

The search array consisted of three items (each  $4^{\circ} \times 4^{\circ}$  area) of distinct shape (drawn from the same set of

shapes and rotations as for WM task) and distinct colour, positioned at the 12, 4, and 8 o'clock positions on an imaginary clock face centred on fixation (7° eccentricity). One of the three shapes was denoted as the target (T) by a thick white outline (0.12° stroke width); the location of the target item was randomly determined on each trial. The colour of the target was chosen randomly on each trial from 360 possible colours, which were evenly distributed along a 360° circle in CIE L\*a\*b\* coordinates with constant luminance ( $L^* = 70$ , centre at  $a^* = 20$ ,  $b^* = 38$ , and radius 60; Zhang & Luck, 2008). The colours of the remaining two distractor shapes (N1 and N2) were set to be equidistant in opposite directions along the colour wheel (120° clockwise or counterclockwise deviation from the target colour). Participants were instructed to search for the target and encode its colour; they were explicitly told that shape was irrelevant for the colour-report task. The search array was presented for 150 ms, followed by 250 ms of masks  $(4^\circ \times 4^\circ$ squares coloured with a random colour value at each pixel location, covering each of the search locations).

Participants then reported the colour of the target item: A large multicoloured wheel (10° diameter; 1° width) was presented at a random rotation, at the centre of the screen. A white line appeared at a random location along the colour wheel to indicate the starting response value, and participants were instructed to adjust the position of the white line to match the colour of the search target. The line's position was adjusted using an input dial (PowerMate USB Multimedia controller, Griffin Technology, USA). To submit their response, participants clicked down on the dial when they were done adjusting. Accuracy was stressed, and there was a time limit of 5 s. Then participants were shown visual feedback for 1000 ms: The reported colour response was shown as a disc (2° diameter) in the centre of the screen, and the actual target colour was displayed as a disc (2° diameter) in the target location.

Then a WM probe item appeared in the centre of the screen under the words "Same or different?" The probe item was a single white shape  $(4^{\circ} \times 4^{\circ} \text{ area})$  either identical to the initial WM cue item (in both shape and rotation) or a different shape (presented at a random rotation) that had not previously appeared in the trial. Participants were instructed to indicate whether the probe was "same" or "different"

# A) Experiment 1



# B) Experiment 2



# C) Experiment 3 (stimuli only)



# D) Target-Distractor Similarity





90° target-distractor similarity



30° target-distract similarity

**Figure 1.** Example trial sequences for (A) Experiment 1 and (B) Experiment 2, and the object stimuli used in (C) Experiment 3. Participants were to remember a shape across an intervening visual search, in which they reproduced the colour (continuous-report) of a target item denoted by a white outline. Critically, the WM shape could reappear in the search array at the target location ("Valid"), at an adjacent distractor location ("Invalid"), or not at all ("Neutral"). Experiment 1 only included Neutral and Invalid trials. Target-distractor similarity was 120° in Experiment 1, or 90° and 30° in Experiment 2, as illustrated with placeholders in (D); the white outline denotes the T item in these panels. The WM probe was either the same or a different shape from the initial WM cue, and auditory feedback indicated whether the WM response was correct or not. For each sequence, the final panel (dotted outline) denotes which array position corresponds to T, N1, N2, and N3 (Experiment 2 only). On Neutral trials, N1 and N2 are arbitrarily assigned. Experiment 3 used the same design as in Experiment 2, but with different stimuli.

as the initial WM cue, by using designated keypresses on the keyboard. "Same" and "different" WM probes occurred equally often, in a randomized order. Accuracy was stressed, and there was a time limit of 2 s. After the WM response, participants were given auditory feedback, with a high beep (800 Hz, 100 ms) to indicate a correct WM response and a low beep (400 Hz, 100 ms) to indicate an incorrect WM response.

#### Design

Trials were classified by validity (i.e., neutral, invalid), depending on the relationship between the WM cue (i.e., the shape to be held in memory) and the search array. To emphasize attentional capture by WM, only neutral and invalid trials were ever presented in Experiment 1, randomly intermixed within blocks with equal frequency (50% neutral, 50% invalid). In invalid trials, the WM cue shape matched the shape of a distractor, whose colour was 120° from the target colour (± direction varied across trials). This critical distractor was always labeled "N1"; thus, distractor "N2" was a control item with a different shape whose colour was set 120° from the target colour in the opposite direction. In neutral trials, the WM cue shape never reappeared within the search array, and "N1" and "N2" labels were randomly assigned to the two distractor items. Thus, in Experiment 1 the WM cue shape could never match the shape of the search target, de-incentivizing any strategic orienting towards the WM-match (e.g., Olivers, 2009; Soto et al., 2005). Each participant completed as many blocks as possible in a 60-minute experimental session, resulting in 5-9 blocks of 36 trials of intermixed neutral and invalid trials, or 90–162 trials of each validity condition. The experiment began with practice trials for the WM task alone (6 trials), the search task alone (10 trials), and the dual-task WM and search (16 neutral trials).

#### Probabilistic mixture modeling

On each trial, colour report error during the search task was calculated as the angular deviation between the reported colour and the true colour of the target item ( $\theta$  = colour error, range  $-180^{\circ}$  to  $180^{\circ}$ ). Overall target colour report accuracy was calculated as the root mean squared error (RMSE; see Fan & Turk-Browne, 2013). For modeling purposes, the correct target (T) colour was represented as 0°, and critical N1 and control N2 distractors were aligned at

+120° and -120°, respectively. To quantify feature perception of the target, we applied a probabilistic mixture model (Bays, Catalao, & Husain, 2009; Golomb et al., 2014; Zhang & Luck, 2008) that combines three von Mises probability density functions with a uniform guessing component:

$$p(\theta) = (1 - \beta - \delta - \lambda)\phi_{\mu,\sigma} + \beta\phi_{2\pi}, \sigma + \delta\phi_{-\frac{2\pi}{3},\sigma} + \lambda\left(\frac{1}{2\pi}\right),$$

where  $\theta$  is the colour response error in radians,  $\phi_{\mu,\sigma}$  reflects a Von Mises distribution with flexible mean  $\mu$  and flexible standard deviation  $\sigma$  (i.e., flexibly centred on the target value),  $\phi_2$  reflects a Von  $\frac{1}{3}\pi,\sigma$ 

Mises distribution with a fixed mean at  $\frac{2}{3}\pi$  (i.e., centred on the critical N1 value, with the same  $\sigma$  as the target distribution), and  $\phi_{-\frac{2}{3}}\pi,\sigma$  reflects a Von

Mises distribution with a fixed mean at  $-\frac{2}{3}\pi$  (i.e., centred on the control N2 value, with the same  $\sigma$  as the target distribution).  $\beta$  is the probability of misreporting the N1 colour value,  $\delta$  is the probability of misreporting the N2 colour value, and  $\lambda$  is the proportion of trials in which the participant responds at random. The mixture model was fit using a differential evolution Markov chain Monte Carlo algorithm (DE-MCMC; Ter Braak, 2006), as implemented through custom R scripts. Parameter estimates were obtained separately for each individual subject and each validity condition, then group means were evaluated with paired *t*-tests or within-subject ANOVAs. Cohen's *d* (effect size) was calculated as the difference in means divided by the pooled standard deviation.

Conventional MCMC sampling performance is greatly influenced by the choice of the proposal distribution (e.g., Gaussian), which generates candidate samples probabilistically (see van Ravenzwaaij, Cassey, & Brown, 2018 for review). However, the DE-MCMC algorithm overcomes this limitation by using a system of interacting Markov chains to adaptively create new candidate samples, boosting its performance and efficiency (Turner, Sederberg, Brown, & Steyvers, 2013). The DE-MCMC method also achieved better convergence for 30°-similarity trials (Experiments 2 and 3). For the current DE-MCMC sampler, we sampled  $\gamma$  randomly on each step from a uniform distribution between 0.5 and 1 and set b = 0.001 for the small-variance random noise. We ran 24 interacting chains and obtained 500 samples after a burn-in period of 500 samples for each chain and approximated the maximum-likelihood estimates of each parameter for the above mixture model by calculating the mean of each posterior distribution.

#### Results

#### Shape WM task

Performance on the WM task (M = 87.6%, SD = 7.2%) was above chance (50%), t(22) = 58.8, p < .001, d = 12.3, but was not significantly different between invalid and neutral trials, t(22) = 0.15, p = .881, d = 0.03. For all subsequent analyses reported below, all trials were included, regardless of how participants responded to the WM probe. Data were also analyzed including only trials in which participants responded correctly to the WM probe, and this resulted in the same patterns of results (data not shown).

#### Colour-report task

Our primary focus of this experiment was on the continuous colour-report data from the intervening search task, and how these colour reports might be influenced by WM-matching distractors. Mean RMSE (i.e., overall error) on the continuous colour-report task was 36.3°

 $(SD = 16.4^{\circ})$  and was significantly greater for invalid than neutral trials, t(22) = 2.25, p = .035, d = 0.47 (Table 1; Figure 2). To quantify whether the types of errors differed across these two conditions, colour responses were fit with a probabilistic mixture model to obtain average estimates of the probabilities of reporting the target (pT), each of the distractors (pN1, pN2), or random guessing (pU); the models also included parameters for the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of the target distribution (Table 2). Comparing these parameter estimates across validity condition showed significant differences for general target processing: Participants were more likely to report the target in neutral trials compared to invalid trials, pT: t(22) = 3.02, p = .006, d = 0.63, and were more likely to guess in invalid than neutral trials, pU: t(22) = 3.99, p < .001, d = 0.83(Figure 3A). Precision of the target report was not significantly different between conditions,  $\sigma$ : t(22) = 1.54, p = .137, d = 0.32. Critically, we found no evidence of specific perceptual interference from a distractor feature: There was no increased misreporting of the colour of the WM-matching N1 distractor – in a  $2 \times$ 2 repeated-measures ANOVA across factors of validity (invalid, neutral) and non-target (critical N1, control N2), neither of the main effects nor the interaction effect was significant, ps > 0.27 (Table 2; Figure 3F). There was also no evidence of mixing of target and distractor colours; the  $\mu$  parameter did not differ

Table 1. Overall colour-report error (RMSE) for Experiments 1, 2 & 3.

Target-distractor similarity	Valid	Neutral	Invalid	Statistical tests
120° (Expt 1)	_	34.7 (17.2)	37.6 (16.2)	<i>t</i> (22) = 2.25, <i>p</i> = .035, <i>d</i> = 0.47 *
90° (Expt 2)	40.6 (14.1)	44.3 (16.4)	44.3 (17.5)	$F(2, 46) = 2.88, p = .067, \eta^2 = .11$
30° (Expt 2)	44.8 (15.5)	49.8 (14.3)	47.3 (17.3)	F(2, 46) = 3.64, $p = .034$ , $\eta^2 = .14 *$ Post-hoc tests Valid-Neutral: $t(23) = 3.57$ , $p = .002$ , $d = 0.73 **$ Valid-Invalid: $t(23) = 1.36$ , $p = .188$ , $d = 0.28$ Neutral-Invalid: $t(23) = 1.31$ , $p = .204$ , $d = 0.27$
90° (Expt 3)	49.3 (17.1)	52.6 (14.2)	56.5 (18.8)	F(2, 38) = 4.87, $p = .013$ , $\eta^2 = .20$ * Post-hoc tests Valid-Neutral: $t(19) = 1.54$ , $p = .141$ , $d = 0.34$ Valid-Invalid: $t(19) = 2.85$ , $p = .010$ , $d = 0.64$ ** Neutral-Invalid: $t(19) = 1.75$ , $p = .097$ , $d = 0.39$
30° (Expt 3)	47.5 (15.5)	48.7 (14.4)	52.5 (17.4)	F(2, 38) = 3.41, $p = .043$ , $\eta^2 = .15 *$ Post-hoc tests Valid-Neutral: $t(19) = 0.60$ , $p = .555$ , $d = 0.13$ Valid-Invalid: $t(19) = 2.70$ , $p = .014$ , $d = 0.60 **$ Neutral-Invalid: $t(19) = 1.82$ , $p = .085$ , $d = 0.41$ Omnibus $3 \times 2$ ANOVA Validity: F(2, 38) = 6.55, $p = .004$ , $\eta^2 = .26 *$ Similarity: F(1, 19) = 4.10, $p = .057$ , $\eta^2 = .18$ Interaction: F(2, 38) = 0.42, $p = .600$ , $\eta^2 = .02$

\*Denotes statistical significance at p < .05.

\*\*Denotes statistical significance at p < .017 (Bonferroni-corrected for multiple post-hoc comparisons).



**Figure 2.** Overall colour error, as measured by root mean squared error (RMSE; Fan & Turk-Browne, 2013). RMSE was significantly different across validity condition in Experiment 2 (for 30°-similarity trials), and in Experiment 3 (for both 90°-similarity and 30°-similarity trials). The corresponding mean values and statistical comparisons are reported in Table 1. Error bars denote standard error.

from zero for neutral, t(22) = 1.11, p = .278, d = 0.23, or invalid trials, t(22) = 0.22, p = .827, d = 0.05, nor between conditions, t(22) = 0.92, p = .369, d = 0.19.

#### Discussion

Experiment 1 suggests that WM-matching distractors can capture attention, but this WM-driven capture seems to disrupt target processing only in a general manner, increasing the probability of reporting at random. In Experiment 1, there was no evidence for any specific perceptual interference, given the lack of significant swapping errors or feature distortions (cf. Golomb et al., 2014). This pattern suggests that effects of WM-driven capture on distractor processing, at least in this paradigm, might be too transient or weak to consequently disrupt perception of the target's features. One alternative explanation is that the colour-report task in Experiment 1 was too easy: Overall probability of reporting the target colour (*pT*) was generally high (M = 86.5%, SD = 11.2%), and the combined probability of misreporting either nontarget colour was generally low (pN1 + pN2, M = 5.8%)SD = 4.2%), perhaps masking any significant differences in swapping or mixing effects that might be present in a harder task. This was addressed in Experiment 2.

### **Experiments 2 & 3**

The primary goal of Experiments 2 and 3 was to magnify any potential effects of WM-driven attention

on distractor processing. For Experiment 2, we made four major changes to the paradigm: (1) We increased the difficulty of the colour-report task by increasing the number of search items, reducing the search item size, and reducing the search presentation duration. The larger set size increases inter-item competition, which has been argued to produce stronger attentional biases (e.g., Hickey, Olivers, Meeter, & Theeuwes, 2011). (2) Attentional bias to WM-matching items was boosted by including both valid (WM-matching target) and invalid (WM-matching distractor) conditions; although it was probabilistically inefficient for participants to attend to the WM-matching shape in order to find the colour-report target (only 33% valid trials), the inclusion of *any* valid trials could still encourage an overt strategy to attend to WM-matching items (Woodman & Luck, 2007). Thus, the potential absence of target feature disruptions

Та	b	le 2.	. Overal	l mode	parameters	for	Experiment	1
----	---	-------	----------	--------	------------	-----	------------	---

Tuble 2.		er purumeters	
Parameter	Neutral	Invalid	Statistical tests
рT	.879 (.11)	.851 (.12)	<i>t</i> (22) = 3.02, <i>p</i> = .006, <i>d</i> = 0.63 *
рU	.066 (.07)	.089 (.08)	t(22) = 3.99, p < .001, d = 0.83 *
pN1	.030 (.02)	.032 (.03)	t(22) = 0.32, p = .752, d = 0.07
pN2	.026 (.02)	.028 (.03)	t(22) = 0.53, p = .602, d = 0.11
			Omnibus $2 \times 2$ ANOVA
			Validity: F(1, 22) = 0.27, p = .610,
			$\eta^2 = .01$
			Non-Target: $F(1, 22) = 1.25$ ,
			$p = .276, \eta^2 = .05$
			Interaction: $F(1, 22) = 0.003$ ,
			$p = .958, \eta^2 < .01$
σ	16.7 (3.1)	16.2 (2.8)	<i>t</i> (22) = 1.54, <i>p</i> = .137, <i>d</i> = 0.32
μ	0.389 (1.68)	-0.102 (2.21)	<i>t</i> (22) = 0.92, <i>p</i> = .369, <i>d</i> = 0.19

\*Denotes statistical significance at p < .05.



**Figure 3.** Mixture model parameter estimates for Experiment 1 (A), Experiment 2 90°-similarity (B), Experiment 2 30°-similarity (C), Experiment 3 90°-similarity (D), and Experiment 3 30°-similarity (E). Plots on the left show the probability of each type of error (e.g., pT), and plots on the right show estimates for the standard deviation of the target distribution ( $\sigma$ ) and the mean of the target distribution ( $\mu$ , or mu). The plots in (F) are zoomed in on the N1 and N2 swapping errors from (A–E). None of the experiments revealed a significant increase in swapping errors for the WM-matching N1 distractor in Invalid trials. The corresponding mean estimates and statistical comparisons are reported in Tables 2–4. Error bars denote standard error.

90° Target-distractor similarity					
Parameter	Valid	Neutral	Invalid	Statistical tests	
рТ	.831 (.10)	.799 (.12)	.793 (.14)	$F(2, 46) = 5.09, p = .010, \eta^2 = .18 *$	
				Post-hoc tests	
				Valid-Neutral: t(23) = 2.48, p = .021, d = 0.51	
				Valid-Invalid: t(23) = 3.28, p = .003, d = 0.67 **	
				Neutral-Invalid: $t(23) = 0.42$ , $p = .679$ , $d = 0.09$	
рU	.064 (.05)	.080 (.06)	.087 (.08)	$F(2, 46) = 3.83, p = .029, \eta^2 = .14 *$	
				Post-hoc tests	
				Valid-Neutral: $t(23) = 2.07$ , $p = .050$ , $d = 0.42$	
				Valid-Invalid: $t(23) = 3.15$ , $p = .005$ , $d = 0.64$ **	
		/>		Neutral-Invalid: $t(23) = 0.70$ , $p = .493$ , $d = 0.14$	
pN1	.037 (.02)	.043 (.03)	.048 (.03)	$F(2, 46) = 1.82, p = .174, \eta^2 = .07$	
pN2	.040 (.03)	.044 (.03)	.038 (.02)	$F(2, 46) = 0.69, p = .505, \eta^2 = .03$	
				Validity $\Gamma(2, 46) = 1.47$ n = 240 n <sup>2</sup> = 06	
				Valually: $F(2, 40) = 1.47$ , $p = .240$ , $T_1 = .00$ Non Target: $F(1, 22) = 0.61$ , $n = .445$ , $n^2 = .03$	
				Non-farget: $P(1, 25) = 0.01$ , $p = .445$ , $T_1 = .05$ Interaction: $E(2, 46) = 1.15$ , $n = .324$ , $n^2 = .05$	
nN3	028 ( 02)	033 ( 02)	033 (03)	$F(2, 46) = 0.97 \text{ n} = 387 \text{ n}^2 = 0.4$	
σ	23 5 (3 5)	22 8 (4 4)	22 7 (4 1)	F(2, 46) = 0.57, p = .567, n = .567	
U	-0.50(2.7)	0.02 (3.2)	0.51 (3.2)	F(2, 46) = 0.52, p = 5.55, n = 5.02 $F(2, 46) = 0.62, n = 5.04, n^2 = 0.3$	
٣	0.50 (2.7)	0.02 (5.2)	0.51 (5.2)	$(2, 10) = 0.02, p = .511, \eta = .05$	
		30° Tar	get-distractor si	milarity	
Parameter	Valid	Neutral	Invalid	Statistical tests	
рT	.485 (.15)	.438 (.14)	.496 (.14)	$F(2, 46) = 2.74, p = .075, \eta^2 = .11$	
рU	.096 (.07)	.124 (.09)	.117 (.08)	$F(2, 46) = 1.60, p = .213, \eta^2 = .07$	
pN1	.198 (.10)	.176 (.07)	.152 (.07)	$F(2, 46) = 3.07, p = .056, \eta^2 = .12$	
pN2	.188 (.07)	.219 (.09)	.195 (.08)	$F(2, 46) = 1.68, p = .198, \eta^2 = .07$	
				Omnibus 3 × 2 ANOVA	
				Validity: $F(2, 46) = 2.01, p = .145, \eta^2 = .08$	
				Non-Target: $F(1, 23) = 3.51$ , $p = .074$ , $\eta^2 = .13$	
				Interaction: $F(2, 46) = 2.78, p = .073, \eta^2 = .11$	
pN3	.033 (.03)	.044 (.03)	.040 (.04)	$F(2, 46) = 2.55, p = .089, \eta^2 = .10$	
σ	16./ (3./)	16.7 (4.2)	16.2 (4.5)	$F(2, 46) = 0.13, p = .880, \eta^2 = .01$	
μ	-3.32 (8./)	-1.13 (18.3)	-0.42 (4.0)	$F(2, 40) = 0.39, p = .080, \eta = .02$	

Table 3. Overall model parameters for Experiment 2.

\*Denotes statistical significance at p < .05.

\*\*Denotes statistical significance at p < .017 (Bonferroni-corrected for multiple post-hoc comparisons).

under these stronger conditions of WM-driven attention would be even more striking. (3) The introduction of valid trials, however, makes it possible that effects of WM-driven attention might operate postperceptually to reduce uncertainty about the location of the target (Cosman & Vecera, 2011), which would not predict any perceptual interference from the processing of a WM-matching distractor. To eliminate uncertainty about the target location during the colour-report response, we displayed a spatial post-cue along with the colour wheel on every trial (Han, 2015a). (4) We manipulated target-distractor similarity by including trials in which the distance between the target colour and adjacent distractor colours along the colour wheel was either 90° or 30°, which have previously been demonstrated to produce mixing errors (perceptual blending of the target and distractor colours) or repulsion errors (perceptual distortion away from the distractor colour), respectively (Golomb, 2015). Experiment 3 attempted to boost the effects of WM-driven attention even further. Instead of geometric shapes, we used images of real-world objects, such that the semantic associations of the WM stimuli could further guide visual attention (e.g., Moores, Laiti, & Chelazzi, 2003; Soto & Humphreys, 2007).

#### Method

#### **Participants**

For Experiment 2, 24 new participants (ages 18–25; 12 male) were recruited from the undergraduate Research Experience Program pool at The Ohio State University. Three additional participants were excluded, 2 for poor WM task performance (<70% WM accuracy) and 1 for completing too few trials in a 60-minute session.<sup>1</sup> Experiment 3 recruited 20 new participants (ages 18–22; 9 male); 8 additional participants were excluded, 1 for poor WM task performance (<70% WM accuracy), 4 for completing too few trials in a 60-minute session, and 3 for not successfully performing the colour-report task (> 50% probability of guessing on neutral trials). All participants reported normal or

90° Target-distractor similarity						
Parameter	Valid	Neutral	Invalid	Statistical tests		
рТ	.759 (.13)	.729 (.10)	.681 (.16)	F(2, 38) = 7.98, $p = .001$ , $\eta^2 = .30$ * Post-hoc tests Valid-Neutral: $t(19) = 1.66$ , $p = .114$ , $d = 0.37$ Valid-Invalid: $t(19) = 4.35$ , $p < .001$ , $d = 0.97$ ** Neutral-Invalid: $t(19) = 2.12$ , $p = 0.47$ , $d = 0.47$		
рU	.108 (.09)	.127 (.06)	.154 (.10)	F(2, 38) = 4.13, $p = .024$ , $\eta^2 = .18 *$ Post-hoc tests Valid-Neutral: $t(19) = 1.15$ , $p = .265$ , $d = 0.26$ Valid-Invalid: $t(19) = 4.13$ , $p < .001$ , $d = 0.92 **$ Neutral-Invalid: $t(19) = 1.39$ , $p = .182$ , $d = 0.31$		
pN1	.047 (.03)	.059 (.03)	.070 (.06)	$F(2, 38) = 2.13, p = .133, \eta^2 = .10$		
pN2	.045 (.03)	.049 (.02)	.053 (.03)	F(2, 38) = 0.39, $p = .683$ , $\eta^2 = .02$ Omnibus 3 × 2 ANOVA Validity: F(2, 38) = 4.80, $p = .014$ , $\eta^2 = .20$ * Non-Target: F(1, 19) = 4.88, $p = .040$ , $\eta^2 = .20$ * Interaction: F(2, 38) = 0.38, $p = .688$ , $\eta^2 = .02$		
pN3	.040 (.03)	.036 (.02)	.043 (.03)	$F(2, 38) = 0.80, p = .456, \eta^2 = .04$		
σ	21.4 (7.2)	20.3 (4.9)	18.7 (4.7)	F(2, 38) = 4.56, $p = .017$ , $\eta^2 = .19$ * Post-hoc tests Valid-Neutral: $t(19) = 1.13$ , $p = .274$ , $d = 0.25$ Valid-Invalid: $t(19) = 2.77$ , $p = .012$ , $d = 0.62$ ** Neutral-Invalid: $t(19) = 2.16$ , $p = .044$ , $d = 0.48$		
μ	0.47 (2.7)	0.16 (3.3)	0.09 (3.2)	$F(2, 38) = 0.08, p = .925, \eta^2 < .01$		
		30° Ta	rget-distractor sin	milarity		
Parameter	Valid	Neutral	Invalid	Statistical tests		
рТ	.501 (.14)	.502 (.16)	.436 (.17)	F(2, 38) = 4.69, $p = .015$ , $\eta^2 = .20$ * Post-hoc tests Valid-Neutral: $t(19) = 0.02$ , $p = .986$ , $d < 0.01$ Valid-Invalid: $t(19) = 2.80$ , $p = .011$ , $d = 0.63$ ** Neutral-Invalid: $t(19) = 2.23$ , $p = .038$ , $d = 0.50$		
рU	.126 (.08)	.138 (.08)	.160 (.13)	$F(2, 38) = 1.62, p = .212, n^2 = .08$		
pN1	.169 (.06)	.156 (.08)	.189 (.08)	$F(2, 38) = 1.04, p = .365, \eta^2 = .05$		
pN2	.173 (.07)	.164 (.05)	.176 (.07)	F(2, 38) = 0.22, $p = .801$ , $\eta^2 = .01$ Omnibus 3 × 2 ANOVA Validity: F(2, 38) = 1.16, $p = .324$ , $\eta^2 = .06$ Non-Target: F(1, 19) = 0.01, $p = .985$ , $\eta^2 < .01$ Interaction: F(2, 38) = 0.27, $p = .765$ , $\eta^2 = .01$		
pN3	.031 (.02)	.041 (.03)	.038 (.02)	$F(2, 38) = 2.30, p = .114, \eta^2 = .11$		
σ	17.1 (4.3)	16.3 (5.1)	16.6 (3.9)	$F(2, 38) = 0.25, p = .778, \eta_2^2 = .01$		
μ	-2.87 (12.7)	8.71 (22.8)	-2.78 (11.7)	$F(2, 38) = 2.61, p = .087, \eta^2 = .12$		

Table 4. Overall model parameters for Experiment 3.

\*Denotes statistical significance at p < .05.

\*\*Denotes statistical significance at p < .017 (Bonferroni-corrected for multiple post-hoc comparisons).

corrected-to-normal visual acuity and colour vision, received course credit, and provided informed consent.

#### Stimuli & procedure

The apparatus was identical to that used in Experiment 1, but eye-tracking data were no longer collected due to the shortened presentation of the search array. The experimental task was also identical to that of Experiment 1, with the following exceptions (Figure 1B): In Experiment 2, each search array consisted of four items (each  $2.5^{\circ} \times 2.5^{\circ}$  area) of distinct shape and distinct colour, positioned at the 12, 3, 6, and 9 o'clock positions on an imaginary clock face centred on fixation (4° eccentricity). The target item was denoted by a thick white outline (0.08° stroke width), and the colour of the target was chosen randomly on each trial; the colours of the two adjacent items were set to be equidistant in opposite directions, deviating either 30° or 90° (clockwise or counterclockwise), and the fourth diagonal item was set 180° away in colour space (Figure 1D). The fourth item was included to increase the difficulty of colour-report task; for the 90° target-distractor similarity trials, the fourth colour also served to balance the array, so that the target colour was not predictable (see Golomb, 2015). The search array was presented for only 50 ms, followed by a black screen for 100 ms (to prevent immediate backward masking of the search display), and then by 200 ms of masks. The continuous colour-report task was presented as a large colour wheel (14° diameter; 1.4° width) at a random rotation. A post cue (white X,  $0.12^{\circ} \times 0.24^{\circ}$ ) indicating the target location also appeared with the onset of the colour wheel to ensure that uncertainty regarding the target location was eliminated (see Cosman & Vecera, 2011). Instead of adjusting a response line with a dial, participants clicked on the appropriate part of the wheel with the mouse to input their response. Participants still received visual feedback for the colour task, but it was displayed as smaller discs (1.25° diameter) for only 500 ms.

In Experiment 3, the apparatus and experimental task were identical to those of Experiment 2, except that the stimulus set comprised 39 images of real-world objects chosen from Brady, Konkle, Alvarez, and Oliva (2008). Each object was transformed into a binary image, resulting in a largely filled-in object shape (approximately  $3.5^{\circ} \times 3.5^{\circ}$ ) that could be assigned a single colour value. We selected objects whose outlines would be generally recognizable and visually distinctive (Figure 1C). Because each object's outline was distinct, the target item was now denoted by a white X ( $0.24^{\circ} \times 0.24^{\circ}$ ) overlaid on the centre of the object. The search array was presented for 75 ms, followed by a black screen for 100 ms, and then by 250 ms of masks.

#### Design

Trials were classified by factors of validity and targetdistractor similarity (i.e., colour deviation). Unlike Experiment 1, Experiments 2 and 3 included valid trials, in which the WM cue matched the target object within the search array. Neutral, valid, and invalid trials were randomly intermixed within blocks with equal frequency (33% neutral, 33% valid, 33% invalid). Target-distractor similarity was manipulated such that the colours of the two distractors adjacent to the target deviated by either 30° or 90° (clockwise, counterclockwise) from the target colour. Each participant completed 12 blocks of 36 trials of intermixed validity (neutral, valid, invalid) and target-distractor similarity (30°, 90°), resulting in 72 trials of each validity-similarity condition.

#### Probabilistic mixture modeling

The data were analyzed in the same way as in Experiment 1, but with minor adjustments to the probabilistic mixture models. As before, the correct T colour was represented as 0°, but the critical N1 and control N2 distractors were aligned at either  $+30^{\circ}$  and  $-30^{\circ}$ , or  $+90^{\circ}$  and  $-90^{\circ}$ , depending on target-distractor

similarity. The additional diagonal N3 distractor was always represented as 180°. The two mixture models used in Experiment 2 combined four Von Mises probability density functions with a uniform guessing component:

$$egin{aligned} p( heta) &= (1 - eta - \delta - \lambda) \phi_{\mu,\sigma} + eta \phi_{+\mathsf{M},\sigma} + \delta \phi_{+\mathsf{M},\sigma} \ &+ \xi \phi_{\pi,\sigma} + \lambda igg(rac{1}{2 \pi}igg), \end{aligned}$$

where all notations are identical to the model in Experiment 1, except that M represents target-distractor similarity of  $\frac{\pi}{6}$  (30°) or of  $\frac{\pi}{2}$  (90°). Thus,  $\phi_{+M,\sigma}$  reflects a Von Mises distribution with a fixed mean centred on the critical N1 distractor, and  $\phi_{-M,\sigma}$  is centred on the control N2 distractor. The models also include centred on the diagonal N3 distractor, where  $\xi$  is the probability of misreporting the N3 colour value. Parameter estimates were obtained with DE-MCMC separately for each individual subject, each validity condition, and each target-distractor similarity condition, then group means were evaluated with within-subject ANOVAs; post-hoc pairwise comparisons were evaluated with a Bonferroni-corrected significance threshold of p = .017.

#### **Results: Experiment 2**

#### Shape WM task

Performance on the WM task (M = 85.7%, SD = 7.8%) was above chance (50%), t(23) = 54.0, p < .001, d = 11.0, and was significantly different across validity conditions, F(2, 46) = 10.8, p < .001,  $\eta^2 = .32$ . Post-hoc tests showed that WM performance on valid trials was significantly better compared to neutral trials, t(23) = 4.64, p < .001, d = 0.95, and to invalid trials, t(23) = 3.10, p = .005, d = 0.63. These memory benefits likely resulted from an updating of the memory representation due to reprocessing the valid target item. WM accuracy was not significantly different between invalid and neutral trials, t(23) = 1.13, p = .270, d = 0.23, as in Experiment 1.

#### Colour-report task (90°)

For trials in which target-distractor similarity was 90°, mean RMSE on the continuous colour-report task was 43.4° ( $SD = 15.3^{\circ}$ ). Compared to the mean RMSE for neutral trials in Experiment 1, a two-sample *t*-test revealed that RMSE for 90°-similarity neutral trials in Experiment 2 was marginally greater, t(45) = 1.96, p = .056, d = 0.57, suggesting that the modified task was more difficult. A repeated-measures ANOVA revealed that colour-report error was marginally different across validity conditions, F(2, 46) = 2.88, p = .067,  $\eta^2 = .111$ , with mean RMSE on valid trials numerically smaller than on neutral or invalid trials (Figure 2).

To quantify the types of errors, each model parameter estimate was compared across validity (Table 3). As in Experiment 1, general target processing was impacted by validity condition, as indicated by significant differences across both *pT*, *F*(2, 46) = 5.09, *p* = .010,  $\eta^2$  = .18, and *pU*, *F*(2, 46) = 3.83, *p* = .029,  $\eta^2$  = .14 (Figure 3B). As evaluated at Bonferroni-adjusted *p*-value of .017, post-hoc tests showed that participants were significantly more likely to report the target in valid compared to invalid trials, *t*(23) = 3.28, *p* = .003, *d* = 0.67, and were significantly less likely to guess in valid trials compared to invalid trials, *t*(23) = 3.15, *p* = .005, *d* = 0.64. The precision of target report was not significantly different across validity condition, *F*(2, 46) = 0.52, *p* = .599,  $\eta^2$  = .02.

Overall rates of misreporting the colour of any nontarget (i.e., N1, N2, or N3; *M* = 11.5%, *SD* = 5.7%) were significantly greater than those of Experiment 1, t (45) = 3.90, p < .001, d = 1.14. Yet we again found no evidence of specific perceptual interference from a distractor feature: A 3 × 2 repeated-measures ANOVA across factors of validity (valid, neutral, invalid) and non-target (critical N1, control N2) revealed that neither of the main effects nor the interaction effect was significant,  $p_{\rm S} > 0.24$  (Figure 3F). There was also no increase in feature mixing when the target colour and critical N1 colour were 90° apart (cf. Golomb, 2015), as the mean of the target distribution for invalid trials did not differ from zero, t(23) = 0.78, p = .446, d = 0.16, or change across validity condition, F  $(2, 46) = 0.62, p = .544, \eta^2 = .03.$ 

#### Colour-report task (30°)

For trials in which target-distractor similarity was 30°, mean RMSE on the continuous colour-report task was 47.6° ( $SD = 14.8^{\circ}$ ), which was significantly larger than RMSE for 90°-similarity trials, t(23) = 3.57, p = .002, d = 0.73. A  $3 \times 2$  repeated-measures ANOVA across factors of validity (valid, neutral, invalid) and target-distractor similarity (90°, 30°) revealed significant main effects for both validity, F(2, 46) = 5.49, p = .007,  $\eta^2$  = .19, and target-distractor similarity, *F*(1, 23) = 12.3, *p* = .002,  $\eta^2$  = .35, but no significant interaction effect, *F*(2, 46) = 0.51, *p* = .607,  $\eta^2$  = .02, indicating that validity effects on colour-report error were robust across both 90°- and 30°-similarity conditions, and were not modulated by target-distractor similarity distance (Table 1; Figure 2).

However, target-distractor similarity did have an impact on critical N1 swapping errors, as indicated by a significant three-way interaction across factors of validity, non-target, and similarity, F(2, 46) = 3.8, p = .029,  $n^2$  = .143 (Figure 3F), although this three-way interaction is not easily interpretable. Breaking it down by similarity, a  $3 \times 2$  repeated-measures ANOVA across factors of validity (valid, neutral, invalid) and non-target (critical N1, control N2) revealed non-significant interactions for both similarity conditions: There was no significant interaction for the 90°-similarity trials (reported above: ps > 0.24), and a weak but non-significant interaction effect for the 30°-similarity trials, F(2, 46) = 2.78, p = .073,  $\eta^2$ =.11; if anything, the results went numerically in the opposite direction of predicted N1 swapping errors for these 30°-similarity trials. Moreover, there was no evidence of feature distortion - either mixing or repulsion - when the target colour and critical N1 colour were 30° apart (cf. Golomb, 2015), as the mean of the target distribution for invalid trials did not differ from zero, t(23) = 0.51, p = .613, d = .11, or change across validity condition, F(2, 46) = 0.39, p = .680,  $n^2$ =.02. Comparisons of pT, pU, and the precision of the target report ( $\sigma$ ) were also not significantly different across validity condition for the 30°-similarity trials (Figure 3C; Table 3).

#### **Results: Experiment 3**

#### **Object WM task**

As in the previous two experiments, performance on the WM task (M = 90.3%, SD = 6.3%) was above chance (50%), t(19) = 64.5, p < .001, d = 14.4, and was significantly different across validity conditions, F(2, 38) = 6.0, p = .005,  $\eta^2 = .24$ .

#### Colour-report task (90°)

For trials in which target-distractor similarity was 90°, mean colour RMSE was  $53.2^{\circ}$  (*SD* = 15.8°). Mean RMSE for 90°-similarity neutral trials in Experiment 3 was greater than RMSE for neutral trials in Experiment 1, t(41) = 3.69, p < .001, d = 1.13, suggesting that this version of the task was indeed more difficult. Colour-report error was also significantly different across validity conditions, F(2, 38) = 4.87, p = .013,  $\eta^2 = .20$ , with mean RMSE on valid trials smaller than invalid trials (Table 1; Figure 2).

In terms of model parameter estimates (Table 4), we again found that general target processing was impacted by validity condition, as indicated by significant differences across both *pT*, *F*(2, 38) = 7.98, *p* = .001,  $\eta^2$  = .30, and *pU*, *F*(2, 38) = 4.13, *p* = .024,  $\eta^2$  = .18 (Figure 3D). Unlike in previous experiments, the precision of target report was significantly different across validity condition, *F*(2, 38) = 4.56, *p* = .017,  $\eta^2$  = .19, such that the standard deviation of valid trials was significantly greater than that of invalid trials, *t*(19) = 2.77, *p* = .012, *d* = 0.62.

Overall rates of misreporting the colour of any nontarget (i.e., N1, N2, or N3; *M* = 14.8%, *SD* = 6.1%) were again greater than those of Experiment 1, and similar (or marginally greater) than those of Experiment 2, t(42) = 1.82, p = .0756, d = 0.55. Yet we again found no evidence of specific perceptual interference from a distractor feature: A  $3 \times 2$  repeated-measures ANOVA across factors of validity (valid, neutral, invalid) and non-target (critical N1, control N2) did not find a significant interaction effect, F(2, 38) =0.38, p = .688,  $\eta^2 = .02$ . There was also no increase in feature mixing when the target colour and critical N1 colour were 90° apart (cf. Golomb, 2015), as the mean of the target distribution for invalid trials did not differ from zero, t(19) = 0.13, p = .900, d = 0.03, or change across validity condition, F(2, 38) = 0.08, p  $= .925, n^2 < .01.$ 

#### Colour-report task (30°)

For trials in which target-distractor similarity was 30°, mean colour RMSE was 49.9° ( $SD = 14.8^{\circ}$ ). A 3 × 2 repeated-measures ANOVA across factors of validity (valid, neutral, invalid) and target-distractor similarity (90°, 30°) found only a significant main effect for validity, F(2, 38) = 6.6, p = .004,  $\eta^2 = .26$ , but no significant interaction effect, F(2, 38) = 0.42, p = .660,  $\eta^2 = .02$ , indicating that validity effects on colour-report error were not modulated by target-distractor similarity distance (Figure 2; Table 1).

Unlike in Experiment 2, in Experiment 3 target-distractor similarity did *not* have a significant impact on critical N1 swapping errors, as evidenced by the lack of a three-way interaction across factors of validity, non-target, and similarity, F(2, 38) = 0.11, p = .899,  $\eta^2$ < .01 (Figure 3F). For the 30°-similarity trials, only general target processing was impacted by validity condition, with significant differences across pT, F(2,38) = 4.69, p = .015,  $\eta^2 = .20$ , There was no evidence of feature swapping or feature distortion in the presence of a WM-matching distractor, and comparisons of pU and the precision of the target report ( $\sigma$ ) were also not significantly different across validity condition (Figure 3E; Table 4).

#### Discussion

The primary goal of Experiments 2 and 3 was to increase difficulty and magnify any WM-driven effects on distractor processing. The adapted paradigms were indeed more difficult, with greater overall colour error (RMSE) and greater overall likelihood of misreporting one of the non-target colours. Critically, we still found no evidence for specific perceptual errors. In other words, the WM-matching distractor did not seem to interfere with the target's feature information. On invalid trials, there was neither an increase in swapping errors for the WMmatching N1 distractor specifically, nor any evidence for feature distortion between the target and N1 features.

The secondary goal of Experiments 2 and 3 was to manipulate target-distractor similarity (90°, 30°), to see if different feature distortion effects might be found if mixing versus repulsion errors were expected (Golomb, 2015). We did not find evidence of mixing errors in 90°-trials or repulsion errors in 30°-trials; there were no significant feature distortions in either case. It should be noted that the general effects of WM-driven attention on target perception (i.e.,  $pT_{r}$ *pU*) were significant in both experiments for the 90°similarity trials, but only significant in Experiment 3 for 30°-similarity trials. Nonetheless, when WM-driven attention effects were present, the effects were limited to these generic target performance measures; we found no systematic perceptual interference in the presence of a WM-matching distractor for any targetdistractor similarities.

A major difference between Experiment 1 and Experiments 2 and 3 was the addition of a post-cue during the colour-report task. While the intention of the post-cue was to rule out the uncertainty

hypothesis of WM-driven attention (Cosman & Vecera, 2011), participants could have potentially changed their search strategy by encoding all colours and locations, then strategically using the post-cue to retrieve the correct response (though this would have been very difficult to execute with such a brief stimulus duration). However, the presence of validity effects in general target processing for 90°-trials in Experiments 2 and 3 indicates that the contents of WM still had an impact on that initial encoding phase. Thus, even if a participant encoded all four colours and their locations, the data suggest that the WM-matching item still received greater attention than a non-matching item. Therefore, the addition of a post-cue should not impact our interpretations of these results.

One notable limitation of including multiple targetdistractor distances in Experiments 2 and 3 is the reduction in the number of trials that each participant completed. Because parameter estimates were obtained by fitting separate probabilistic models for each individual subject and each validity-similarity condition, the data for these experiments reflect only 72 trials per model (see Supplemental Table S1). In other words, it is possible that the current grouplevel results are based on underpowered models at the individual-level. To address this, we ran another experiment that was identical to Experiment 2, except it included only 90°-similarity trials. Although this resulted in 144 trials per model, that experiment - surprisingly - found no evidence for the general indicators of WM-driven attention, even for basic colour error RMSE (see Supplemental Experiment 6). In the interest of transparency, we report this experiment in the Supplement, along with two other experiments in our "file drawer" similar to Experiment 1 that also failed to produce reliable generic WM-driven attention effects (Supplemental Experiments 4-5). These data suggest that WM-driven attention in general may not be as robust or reliable in this paradigm, as discussed more below. However, the fact that Experiments 2 and 3 produced such similar patterns of data despite the smaller number of trials per subject - and the fact that a very similar pattern was found in Experiment 1 – suggests that the lack of specific perceptual interference is unlikely due to a small number of trials, but more likely reflects a property of WM-driven attention.

#### **General discussion**

When visual attention is biased towards WM-matching items, WM-driven attention has been shown to enhance early perceptual processing of that information (e.g., more accurate and more sensitive perceptual discrimination) - but only when the WMmatching items are task-relevant (i.e., targets; e.g., Han, 2015a; Pan et al., 2016). But what happens when WM-matching items are task-irrelevant (i.e., distractors)? Here, we examined whether WM-driven attentional capture by a WM-matching distractor interferes with perceptual processing of a target. When WM-driven effects on attention were present, we found no evidence of specific perceptual interference - participants were no more likely to report the incorrect WM-matching feature (swapping errors), nor to blend target and distractor features (feature distortion errors). Instead, we observed only general disruptions in target processing, such that participants were less likely to report the target colour at all and more likely to guess randomly. These results suggest an account in which visual attention may be drawn to a WM-matching distractor, but any automatic perceptual processing that results from that attentional focus is too transient or weak to significantly interfere with perceptual processing of the target's features.

Previous studies that have examined the impact of WM-driven attention on perceptual processing found an asymmetry between task-beneficial enhancements of target accuracy (Han, 2015a; Pan et al., 2012, 2016; Soto et al., 2010) and the lack of task-detrimental interference from attending to a distractor (Han, 2015a; Pan et al., 2016). These studies, however, utilized paradigms in which the attentional task required a binary response (e.g., whether the target is oriented left or right; Cosman & Vecera, 2011; Han, 2015a; Pan et al., 2012, 2016; Soto et al., 2010) - limiting measures of perception to discrimination accuracy or sensitivity (i.e., A'). In contrast, the current study used a continuous-report colour task (initially described by Wilken & Ma, 2004), which could allow for more nuanced interpretations of target feature perception via probabilistic mixture modeling (e.g., Bays et al., 2009; Golomb et al., 2014; Zhang & Luck, 2008). Nevertheless, the current modeling results are consistent with previous studies in that invalidly cueing a WM-matching distractor location does not adversely impact perception of the target's features.

The direction of this perceptual asymmetry stands in contrast to typical WM-driven attentional effects, in which response times are both speeded by WMmatching targets and slowed by WM-matching distractors (e.g., Soto et al., 2005; Kiyonaga et al., 2012). In fact, several recent studies have demonstrated that invalid costs (i.e., slower responses) are numerically greater than valid benefits (i.e., faster responses; Dowd, Pearson, & Egner, 2017; Kiyonaga & Egner, 2014; Soto, Rotshtein, & Kanai, 2014), although costs and benefit effect sizes were not specifically compared in those studies. Only one previous study has explicitly reported greater benefits in WM-driven effects on response times (Carlisle & Woodman, 2011). The differences in response time patterns are likely because each of these studies used different variants of dualtask WM and attentional paradigms, with distinct stimuli, timing, instruction manipulations, and attentional loads. Regardless of direction, the asymmetry in WM-driven effects may reflect distinct mechanisms of target enhancement versus distractor suppression (e.g., Eriksen & Hoffman, 1974; Noonan et al., 2016). For instance, one study found that not only were WM-driven benefits uncorrelated with WM-driven costs (on response times) within a sample of 44 participants, but benefits versus costs were also associated with two separate neural regions within left posterior parietal cortex (Soto et al., 2014). Thus, the current results may reflect a strategic boosting of target perception while also minimizing distractor interference (see also Carlisle & Woodman, 2011).

Note that in the present data, we generally find "enhanced target perception" in terms of target accuracy (i.e., pT) and not in terms of precision ( $\sigma$ ): In Experiment 2, target reports for valid trials were not any more precise than for neutral or invalid trials, while in Experiment 3, target reports for invalid trials were, counterintuitively, more precise. One interpretation is that these data support only a general attentional effect, in which WM-driven attention serves as a spatial gatekeeper - and not a perceptual magnifier - for target processing. Another possibility is that the effects of WM-driven attention in the current paradigm are too weak to reliably disrupt feature perception. For instance, the WM feature (shape) was orthogonal to the critical continuous-report feature (colour), potentially reducing WM-driven effects on a colour-report task. In the current experiments, we used orthogonal features because we were specifically interested in the phenomenon of automatic perceptual enhancement that might arise from incidental WM-driven attention (see Soto et al., 2008). However, according to a theory of WM in which multiple items can be maintained with various levels of "activation" or "priority" (e.g., Olivers et al., 2011), maintaining a feature that is not directly relevant for search (i.e., shape) may have a weaker impact on WM-driven spatial attention - and subsequently, WM-driven processing. In contrast, other studies have demonstrated that maintaining a directly-relevant feature in WM distorts perception of that same feature (e.g., Kang, Hong, Blake, & Woodman, 2011; Olkkonen & Allred, 2014; Scocchia, Cicchini, & Triesch, 2013). While our current data do not directly speak to this, one possible hypothesis is that only when maintaining search-relevant feature information in WM would subsequent WM-driven feature processing be distorted; otherwise, merely holding information in WM influences only the spatial guidance of attention.

The lack of WM-driven perceptual interference in Experiments 1–3 stands in contrast to a recent study of exogenously-driven attention: Using a similar paradigm, Chen, Leber, and Golomb (in press) compared colour reports when attention was exogenously captured by a salient cue, either at the location of the target (i.e., valid trials) or at the location of a distractor (invalid trials). Exogenously-driven attention did enhance target perception: valid trials had the best precision and lowest guessing rates. More importantly, attentional capture by an exogenously-cued distractor also interfered with target perception, with large swapping errors (misreporting the cued distractor colour) and subtle but significant repulsion errors (reporting a colour shifted away from the cued distractor). In conjunction with the current results, these data support a framework in which WM-driven attention and exogenously-driven attention rely on distinct mechanisms (e.g., Corbetta & Shulman, 2002).

More broadly, the basic effect of WM-driven attention in our dual-task paradigm was not as robust or reliable as expected, given the prior literature (e.g., Soto et al., 2008). The standard measure by which WM-driven attention is typically assessed is through speeded search response times; however, we speculate that our paradigm's inclusion of a continuousreport task may have diluted our ability to assess WM-driven attentional effects, and perhaps that is why we did not find differences in colour-report error or general target processing across validity conditions in Supplemental Experiments 4–6 (see Supplement for more discussion). However, the data in Experiments 1–3 that do reveal significant validity effects did not find any significant feature swapping or distortion, suggesting that even when a WM-matching distractor reliably affects general target processing, it does not alter perception of the target's features. Thus, WM-driven attention in our paradigm may be too transient or weak to interfere with feature processing and feature binding in the same way as other attentional manipulations (e.g., goaldriven shifting/splitting [Dowd & Golomb, 2019; Golomb et al., 2014] or stimulus-driven attentional capture [Chen et al., in press]).

Overall, we found that participants' reports of a target feature may be impaired but not perceptually disrupted by a WM-matching distractor. Combined with previous work (Han, 2015a; Pan et al., 2012, 2016; Soto et al., 2010), the current results suggest that WM-driven attention may strategically enhance early perceptual processing of task-relevant information, without equivalent costs for task-irrelevant information.

#### Note

 For probabilistic mixture modeling, there is no standard minimum number of trials that ensures reliable parameter estimates. While we utilized DE-MCMC algorithms to boost fitting performance, we also set a minimum number of 72 trials per validity-similarity condition in order for a participant's data to be included. Participants who completed too few trials within the fixed time limit (60 minutes) were excluded from analysis (see Supplemental Table S1).

#### Acknowledgements

We thank Anisha Babu, Jack Carey, Julius de Castro, Matthew Heard, Christopher Jones, Madison Jupina, Xingyao Li, Emma Megla, Robert Murcko, Makaela Nartker, Deniz Ozkardas, Maurryce Starks, and Shangchao Sun for assistance with data collection and Brandon Turner for assistance with DE-MCMC analyses.

#### **Disclosure statement**

No potential conflict of interest was reported by the authors.

#### Funding

This work was supported by the National Institutes of Health under grants R01-EY025648 to J. D. Golomb and F32-EY028011 to E. W. Dowd and by the Alfred P. Sloan Foundation to J. D. Golomb.

## ORCID

*Emma Wu Dowd* b http://orcid.org/0000-0003-0219-8527 *Julie D. Golomb* http://orcid.org/0000-0003-3489-0702

#### References

- Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9(10), 1–11. doi:10.1167/9.10.7
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, 105(38), 14325–14329. doi:10.1073/pnas.0803390105
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. doi:10.1163/156856897x00357
- Carlisle, N. B., & Woodman, G. F. (2011). Automatic and strategic effects in the guidance of attention by working memory representations. *Acta Psychologica*, *137*(2), 217–225. doi:10. 1016/j.actpsy.2010.06.012
- Chen, J., Leber, A. B., & Golomb, J. D. (in press). Attentional capture alters feature perception. *Journal of Experimental Psychology: Human Perception & Performance*.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. doi:10.1038/nrn755
- Cosman, J. D., & Vecera, S. P. (2011). The contents of visual working memory reduce uncertainty during visual search. *Attention, Perception, & Psychophysics, 73*(4), 996–1002. doi:10.3758/s13414-011-0093-y
- Dowd, E. W., & Golomb, J. D. (2019). Object-feature binding survives dynamic shifts of spatial attention. *Psychological Science*, *30*(3), 343–361. doi:10.1177/0956797618818481
- Dowd, E. W., Kiyonaga, A., Beck, J. M., & Egner, T. (2015a). Quality and accessibility of visual working memory during cognitive control of attentional guidance: A Bayesian model comparison approach. *Visual Cognition*, 23(3), 337–356. doi:10.1080/ 13506285.2014.1003631
- Dowd, E. W., Kiyonaga, A., Egner, T., & Mitroff, S. R. (2015b). Attentional guidance by working memory differs by paradigm: An individual-differences approach. Attention, Perception, & Psychophysics, 77(3), 704–712. doi:10.3758/ s13414-015-0847-z
- Dowd, E. W., Pearson, J. M., & Egner, T. (2017). Decoding working memory content from attentional biases. *Psychonomic Bulletin & Review*, 24(4), 1252–1260. doi:10. 3758/s13423-016-1204-5

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science*, 11 (6), 467–473. doi:10.1111/1467-9280.00290

Eriksen, C. W., & Hoffman, J. E. (1974). Selective attention: Noise suppression or signal enhancement? *Bulletin of the Psychonomic Society*, 4(6), 587–589. doi:10.3758/BF03334301

Fan, J. E., & Turk-Browne, N. B. (2013). Internal attention to features in visual short-term memory guides object learning. *Cognition*, 129(2), 292–308. doi:10.1016/j.cognition.2013.06. 007

Golomb, J. D. (2015). Divided spatial attention and featuremixing errors. *Attention, Perception, & Psychophysics, 77*(8), 2562–2569. doi:10.3758/s13414-015-0951-0

Golomb, J. D., L'Heureux, Z. E., & Kanwisher, N. (2014). Featurebinding errors after eye movements and shifts of attention. *Psychological Science*, 25(5), 1067–1078. doi:10.1177/ 0956797614522068

Han, S. W. (2015a). Working memory contents enhance perception under stimulus-driven competition. *Memory & Cognition*, 43(3), 432–440. doi:10.3758/s13421-014-0460-8

- Han, S. W. (2015b). Working memory contents revive the neglected, but suppress the inhibited. *Cognition*, 145(C), 116– 121. doi:10.1016/j.cognition.2015.08.012
- Han, S. W., & Kim, M.-S. (2009). Do the contents of working memory capture attention? Yes, but cognitive control matters. *Journal of Experimental Psychology: Human Perception and Performance*, 35(5), 1292–1302. doi:10.1037/ a0016452

Hickey, C., Olivers, C., Meeter, M., & Theeuwes, J. (2011). Feature priming and the capture of visual attention: Linking two ambiguity resolution hypotheses. *Brain Research*, *1370*(C), 175–184. doi:10.1016/j.brainres.2010.11.025

Hollingworth, A., Matsukura, M., & Luck, S. J. (2013). Visual working memory modulates rapid eye movements to simple onset targets. *Psychological Science*, 24(5), 790–796. doi:10.1177/0956797612459767

Kang, M.-S., Hong, S. W., Blake, R., & Woodman, G. F. (2011). Visual working memory contaminates perception. *Psychonomic Bulletin & Review*, 18(5), 860–869. doi:10.3758/ s13423-011-0126-5

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, 20(2), 228–242. doi:10.3758/s13423-012-0359-y

Kiyonaga, A., & Egner, T. (2014). Resource-sharing between internal maintenance and external selection modulates attentional capture by working memory content. *Frontiers in Human Neuroscience*, 8(670), 1–9. doi:10.3389/fnhum. 2014.00670/abstract

Kiyonaga, A., Egner, T., & Soto, D. (2012). Cognitive control over working memory biases of selection. *Psychonomic Bulletin & Review*, 19(4), 639–646.

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience*, *6*(2), 182–189. doi:10.1038/nn996

Noonan, M. P., Adamian, N., Pike, A., Printzlau, F., Crittenden, B. M., & Stokes, M. G. (2016). Distinct mechanisms for distractor suppression and target facilitation. *The Journal of Neuroscience*, *36*(6), 1797–1807. doi:10.1523/JNEUROSCI. 2133-15.2016

- Olivers, C. N. L. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1275–1291. doi:10.1037/ a0013896
- Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1243–1265. doi:10.1037/0096-1523.32.5.1243
- 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. doi:10.1016/j.tics.2011.05. 004
- Olkkonen, M., & Allred, S. R. (2014). Short-term memory affects color perception in context. *PloS ONE*, *9*(1), e86488. doi:10. 1371/journal.pone.0086488
- Pan, Y., Cheng, Q.-P., & Luo, Q.-Y. (2012). Working memory can enhance unconscious visual perception. *Psychonomic Bulletin & Review*, 19(3), 477–482. doi:10.3758/s13423-012-0219-9
- Pan, Y., Luo, Q., & Cheng, M. (2016). Working memory-driven attention improves spatial resolution: Support for perceptual enhancement. Attention, Perception, & Psychophysics, 1–8. doi:10.3758/s13414-016-1138-z
- Scocchia, L., Cicchini, G. M., & Triesch, J. (2013). What's "up"? Working memory contents can bias orientation processing. *Vision Research*, *78*, 46–55. doi:10.1016/j.visres.2012.12. 003
- Silvis, J. D., Belopolsky, A. V., Murris, J. W. I., & Donk, M. (2015). The effects of feature-based priming and visual working memory on oculomotor capture. *PloS ONE*, *10*(11), e0142696. doi:10.1371/journal.pone.0142696
- Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(2), 248–261. doi:10. 1037/0096-1523.31.2.248
- Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, *12*(9), 342–348. doi:10.1016/j. tics.2008.05.007
- Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 730–737. doi:10.1037/0096-1523.33.3730
- Soto, D., Rotshtein, P., & Kanai, R. (2014). Parietal structure and function explain human variation in working memory biases of visual attention. *NeuroImage*, 89(C), 289–296. doi:10.1016/ j.neuroimage.2013.11.036
- Soto, D., Wriglesworth, A., Bahrami-Balani, A., & Humphreys, G.
  W. (2010). Working memory enhances visual perception: Evidence from signal detection analysis. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 36 (2), 441–456. doi:10.1037/a0018686

- Ter Braak, C. J. F. (2006). A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: Easy Bayesian computing for real parameter spaces. *Statistics and Computing*, *16*(3), 239–249. doi:10.1007/s11222-006-8769-1
- Turner, B. M., Sederberg, P. B., Brown, S. D., & Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods*, *18*(3), 368– 384. doi:10.1037/a0032222
- van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic Bulletin & Review*, *25*(1), 143–154. doi:10.3758/ s13423-016-1015-8

- Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, 4(12), 1120–1135. doi:10.1167/4.12.11
- Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. G. (2013). Where do we store the memory representations that guide attention? *Journal of Vision*, *13*(3), 1–1. doi:10.1167/13.3.1
- Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance*, 33(2), 363– 377. doi:10.1037/0096-1523.33.2.363
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233–235. doi:10.1038/nature06860