

A comparison of mental imagery and perceptual cueing across domains of attention

Ming-Ray Liao¹ · James D. Grindell¹ · Brian A. Anderson¹

Accepted: 10 June 2023 / Published online: 22 June 2023 © The Psychonomic Society, Inc. 2023

Abstract

Mental imagery and perceptual cues can influence subsequent visual search performance, but examination of this influence has been limited to low-level features like colors and shapes. The present study investigated how the two types of cues influence low-level visual search, visual search with realistic objects, and executive attention. On each trial, participants were either presented with a colored square or tasked with using mental imagery to generate a colored square that could match the target (valid trial) or distractor (invalid trial) in the search array that followed (Experiments 1 and 3). In a separate experiment, the colored square displayed or generated was replaced with a realistic object in a specific category that could appear as a target or distractor in the search array (Experiment 2). Although the displayed object was in the same category as an item in the search display, they were never a perfect match (e.g., jam drop cookie instead of chocolate chip). Our findings revealed that the facilitation of performance on valid trials compared with invalid trials was greater for perceptual cues than imagery cues for low-level features (Experiment 1), whereas the influence of these two types of cues was comparable in the context of realistic objects (Experiment 2) The influence of mental imagery appears not to extend to the resolution of conflict generated by color-word Stroop stimuli (Experiment 3). The present findings extend our understanding of how mental imagery influences the allocation of attention.

Keywords Imagery · Visual search · Stroop

Introduction

When searching for a specific object in a complex visual environment, our search is facilitated if we have recently seen the object (see Kristjánsson & Ásgeirsson, 2019, for a review). Even in the complete absence of the object, however, we can still conjure up a mental image of the object to guide our attention. Mental imagery refers to the perceptual representation of sensory information in the absence of external stimuli (Pearson et al., 2015), and has been shown to have overlapping neural representations with perception

Ming-Ray Liao m4liao@tamu.edu

> James D. Grindell james.grindell@gmail.com

Brian A. Anderson brian.anderson@tamu.edu

¹ Department of Psychological and Brain Sciences, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235, USA (Kosslyn et al., 2001; Kosslyn & Thompson, 2003; but see also Pylyshyn, 2003). Many studies have demonstrated effects of perceptual cues on visual search in which cues that were similar to or matched the upcoming target (i.e., valid cues) facilitated target detection (e.g., Gunseli et al., 2016; Soto et al., 2005; Theeuwes et al., 2006; Wolfe et al., 2004), and the same has been more recently shown with mental imagery cues (Cochrane et al., 2018, 2019; Cochrane, Townsend, et al., 2021a; Cochrane, Wang, et al., 2021b; Moriya, 2018; Reinhart et al., 2015; Wantz et al., 2015), but how the two might differ in guiding different types of attention remains unclear.

In a particular study (Cochrane et al., 2019, Experiments 5a and 5b), participants generated color imagery prior to a color singleton search task and demonstrated faster responses when the imagined color and the color of the search target were matching (see also Chang et al., 2013, and Pearson et al., 2008, for analogous findings in the context of binocular rivalry). Even after articulatory suppression was implemented in Experiments 6a and 6b, the participants still exhibited an imagery effect on visual search.

Using the same paradigm in a separate study, Cochrane, Townsend, et al. (2021a) found a significant N2 posterior contralateral (N2pc) component only on congruent trials. The N2pc is an electroencephalogram (EEG) index for attentional allocation (Eimer, 2014; Hickey et al., 2009; Luck & Hillyard, 1994), that is also present when a perceptual cue is congruent with the search target (e.g., Woodman et al., 2009). Generally, studies comparing the effects of cueing from perception and mental imagery on attention employ a between-subjects design (e.g., Cochrane et al., 2019; Cochrane, Wang, et al., 2021b) or a block design (Cochrane et al., 2023). In both instances, the validity effect was greater for mental imagery compared with perceptual cues. This was also evident in studies using binocular rivalry which showed numerically larger effects for imagery (Chang et al., 2013; Pearson et al., 2008).

Many papers have demonstrated that visual mental imagery and visual perception exhibit similar patterns of activations throughout the ventral visual stream (see Dijkstra et al., 2019, for a review, but see also Pylyshyn, 2003). To further demonstrate that the neural overlap between visual mental imagery and perception is not due to factors like attention and semantic representation, Naselaris et al. (2015) used voxel-wise modeling and decoding to confirm that lowlevel visual features are encoded during mental imagery of complex scenes. The similarities in neural activity between mental imagery and perception also extend to frontoparietal areas (Dijkstra et al., 2019). The parietal cortex is important for spatial and feature-based attention (Corbetta & Shulman, 2002; Corbetta et al., 2008; Serences et al., 2005; Yantis et al., 2002), while the role of the frontal cortex is more related to representing task-related information rather than stimulus property (Bugatus et al., 2017; Hebart et al., 2018; Lee et al., 2013). The N2pc component elicited by both mental imagery and perception cues during visual search has been localized to a circuit that includes both parietal and occipito-temporal areas (Hopf et al., 2000). However, there seems to be more overlap between mental imagery and perception in higher-level, object-selective regions of the brain than in low-level, retinotopic areas (Lee et al., 2012). It is likely that both mental imagery and perceptual cues activate what amounts to a target template somewhere in the parietal lobe (Corbetta et al., 1995) that can influence the allocation of attention (Cochrane et al., 2018; Moriya, 2018; Vickery et al., 2005; Wolfe et al., 2003, 2004), but the stimuli in these studies have been limited to simple objects.

Although real-world objects can be broken down into conjunctions of simple features, the whole appears to be greater than the sum of its parts. For example, working memory capacity is greater for real-world objects than objects defined solely by variation in a small range of feature dimensions, possibly due to representations for lowlevel features being supplemented with activations from higher-level visual areas (Brady & Störmer, 2021). It is also the case that response times are slower for objects with larger inferred sizes (Collegio et al., 2019), consistent with a myriad of additional top-down effects that real-world objects can exert on attention. The differences in neural representations for simple and complex visual features extend to the realm of visual imagery, as activations differ depending on the type of stimuli being imagined (Dijkstra et al., 2019; Ragni et al., 2021; Slinn et al., 2023). Since visual working memory and mental imagery are strongly related (Keogh & Pearson, 2011; Tong, 2013), it is likely that the nature of attentional guidance by mental imagery and perceptual cues is to some degree dependent on the type of stimuli in question.

In the cueing literature, perceptual cues that were exact matches to targets were found to be more effective at guiding attention than more abstract cues such as words and categories (Wolfe et al., 2004) and top-down knowledge about the target identity (Theeuwes et al., 2006). To the degree that mental imagery is similarly less precise and more abstract, we might expect weaker effects of mental imagery relative to perceptual cues in the guidance of attention to a target, particularly when the cue and target are more closely matched with respect to specific features. On the other hand, we might expect perceptual cues for image categories to have a more comparable effect to cues generated via mental imagery.

The goal of the present study was to determine how attentional guidance from mental imagery and perceptual cues vary across different visual attention tasks. The imagery and perceptual cueing manipulations were based on those employed in prior studies (e.g., Cochrane et al., 2018, 2019, 2022; Cochrane, Townsend, et al., 2021a; Cochrane, Wang, et al., 2021b) to investigate differences in attentional cueing on low-level visual search (Experiment 1), visual search with real-world objects (Experiment 2), and executive attention using color-word Stroop stimuli (Experiment 3). Given findings in the cueing literature concerning perceptual and more abstract cues (Theeuwes et al., 2006; Wolfe et al., 2004), we predicted an advantage of perceptual cues over mental imagery cues in Experiment 1. On the other hand, since abstract cues and similarly-matched cues produced less of an advantage on visual search (Theeuwes et al., 2006; Wolfe et al., 2004), and given the high degree of neural overlap between mental imagery and perception in higher-level visual regions (Lee et al., 2012), we predicted that visual search for objects would be the domain in which performance between imagery and perception is most similar. Since imagery involves the recruitment of perceptual representation (Kosslyn & Thompson, 2003; Kosslyn et al., 2001), we expected the task least grounded in perceptionthe executive attention task in Experiment 3-to be where the effects of imagery on attention would be the least pronounced. In other words, in Experiment 3 we expected to

observe a robust validity effect following perceptual cues, but little to no validity effect following imagery cues.

Experiment 1

Method

Participants

One hundred and twenty-two participants were recruited from the Texas A&M University community. The sample size would yield power $(1 - \beta) > 0.80$ to detect an effect of d = 0.32 (computed using G*Power 3.1), which was the effect size from experiment 5a in Cochrane et al. (2019), and was chosen to be $n \ge 100$ to account for increased noise in the data due to the online method of experiment delivery. Two participants did not complete the study and data for 18 participants were omitted due to low accuracy (<50%). The final sample included 102 participants (66 females and 36 males; 55.9% non-Hispanic Whites) with a mean age of 18.5 (SD = 0.93). All reported normal or corrected-to normal visual acuity, normal color vision, and provided informed consent. Participants were compensated with course credit. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

Apparatus Participants completed the experiment online using their web browser. Stimulus presentation was controlled by jsPsych (de Leeuw, 2015) hosted on a physical lab server with JATOS (Lange et al., 2015) and distributed to members of the Texas A&M human subject pool through Sona systems. For more detail on experiment code, see Liao et al. (2021).

Stimuli, design, and procedure Each trial consisted of a fixation cross $(80 \times 80 \text{ pixels})$ at the center of the screen for 400-60 ms, followed by a 200-ms blank, a cueing sequence, another 200ms blank, a singleton search display, and a blank intertrial interval for 500-700 ms (Fig. 1). There was also a feedback display (TOO SLOW or INCORRECT) for 1,200 ms that was inserted before the intertrial interval depending on participant response, only when participants made an error or did not make a response before the trial timed out. The perceptual cueing sequence was comprised of a square $(400 \times 400 \text{ pixels})$ in either red, green, blue, or yellow (RGB: [255,0,0], [0,220,0], [0,127,255], [240,240,0], respectively) presented for 1,500 ms. The imagery cueing sequence was comprised of a letter (24 pixels) representing the colors red, green, blue or yellow for 500ms (R, G, B, and Y, respectively), followed by a 500-ms blank and then by an empty square (400×400) for 1,500 ms. Participants



Fig. 1 Time course of trial events for Experiments 1 and 3. Each trial began with a fixation cross for a random duration between 400 and 600 ms, followed by the cueing sequence. The imagery cueing sequence begins with a letter cue for 500 ms, followed by a 500-ms blank, followed by an empty square for 1,500 ms. The perception cueing sequence consists of a colored square for 1,500 ms. A 200-ms blank is displayed after the cueing sequence, followed by the search display for 1,200-ms or until a response has been registered. If participants were too slow or incorrect, there would be a feedback display presented for 1,200 ms. Each trial ends with a blank screen for a random duration between 500 and 700 ms. Since red is the cued color in this example, it is a valid trial for Experiment 1 and invalid trial for both color-ink (relevant dimension) and color-word (irrelevant dimension) for Experiment 3. (Color figure online)

were instructed to mentally visualize the cued color inside the empty square.

The search display was comprised of either three circles $(150 \times 150 \text{ pixels})$ and a diamond $(125 \times 125 \text{ pixels})$, or three diamonds and a circle for 1,200 ms or until response. The shapes were presented in cardinal locations on an imaginary circle with a radius of 270 pixels. The shapes also contained a small black square $(25 \times 25 \text{ pixels})$ on either the left or right side. Before the experiment began, participants completed a series of practice trials that introduced components of the trial sequence successively. Participants were first introduced to the perceptual and imagery cueing sequence, before rating the vividness of what they just saw or imagined. The participants were then instructed to search for the odd shape out and report the location of the black dot within this shape on the keyboard with the 'Z' and 'M' key for left and right, respectively. The participants then practiced singleton search, then singleton search following imagery cueing and perceptual cueing. The practice for each component began with two trials without time limit followed by two trials with time limit. Participants also had the opportunity to repeat practice trials.

Once the experiment began, participants were given a 60 second break after 128 and 256 search trials. In total, there were 384 search trials and 16 vividness probe trials. Half of the search (192 trials) and probe (8) trials involved imagery cues, and the other half involved perceptual cues. Within the 192 search trials for each cue type, each color served as the cue equally often (48 trials each). The target was equally likely to be in each of the four locations and each of the two shapes, and cue validity was set to 25%. Cue validity is defined with respect to the probability with which the cue predicts the target identity, with valid trials occurring when the color of the cue and target match. Consequently, the validity effect is defined as the difference in performance between invalid and valid cue trials. The order of cue type and visual search array combinations was randomly intermixed throughout each block of the experiment, as opposed to being grouped by blocks of imagery and perception trials as in some prior experiments (e.g., Cochrane et al., 2023). One probe trial was presented randomly within each block of 24 search trials. A probe trial amounted to a trial in which the probe question appeared in lieu of the search display. On probe trials, participants were given 30 seconds to indicate on a 0–100 scale how vividly they either imagined or saw the color.

Data analysis Data for Experiment 1 were binned by cue type (perception vs. imagery) and validity (valid vs. invalid), and response times (RTs) were trimmed if they were faster than 150ms (anticipatory) or exceeded three standard deviations of the conditional mean. As participants varied in the extent to which they prioritized speed over accuracy (see supplementary tables for mean and standard deviations of RT and accuracy across conditions), we combined RT and accuracy into a single measure: inverse efficiency (IE;

Townsend & Ashby, 1978, 1983). This measure allowed us to capture the full range of influences on performance. So, for a given participant, the *IE* for the condition consisted of the conditional mean *RT* of the correct responses after trimming divided by the *PC*, proportion of correct responses for that condition:

$$IE = RT/PC.$$
 (1)

Conditional IEs were analyzed with a 2x2 analysis of variance (ANOVA) and pairwise comparisons. Vividness ratings for imagery and perceptual cues were also analyzed with pairwise comparisons. Data wrangling and analyses were conducted with JavaScript and R.

Results

A 2 \times 2 ANOVA with cue type (perceptual vs. imagery) and validity (valid vs. invalid) as factors revealed a main effect of validity, F(1, 101) = 57.6, p < .001, $\eta_p^2 = .3$, and a significant interaction effect, F(1, 101) = 4.7, p = .032, $\eta_n^2 = .045$ (Fig. 2A). The main effect of cue type was not significant, $F(1, 101) = 1.8, p = .186, \eta_p^2 = .017$. Both imagery and perceptual cueing produced significant validity effects, ps < .001, but valid trials following perceptual cues were faster compared with those following imagery cues, t(101) = 2.6, p = .012, d = .26. The interaction effect observed was thus driven by larger validity effects for perceptual over imagery cues. Vividness ratings following perceptual cues were significantly greater than ratings following imagery cues, t(101)= 10.4, p < .001, d = 1.03, but did not correlate with validity effects for either cue types, r(100) = .02, p = .860, and r(100) = .04, p = .694, respectively (Fig. 2B). Vividness ratings ranged from 2.14-100 for imagery cues and 1.25-100 for perceptual cues. Removing participants with vividness scores less than 10 did not change our results.



Fig. 2 Mean (A) inverse efficiencies (RT/accuracy) as a function of validity and cue type and (B) vividness ratings for each cue type for Experiment 1. Error bars represent standard error of the mean. (Color figure online)

Discussion

We replicated prior findings that mental imagery can guide attention during a subsequent visual search in similar fashion to perceptual cueing (Cochrane et al., 2018, 2019; Cochrane, Townsend, et al., 2021a; Cochrane, Wang, et al., 2021b; Moriya, 2018; Reinhart et al., 2015; Wantz et al., 2015). In contrast to prior findings (Chang et al., 2013; Cochrane et al., 2018; Pearson et al., 2008; Reinhart et al., 2015), however, we found that cueing effects are stronger following a perceptual cue than with imagery. This is broadly consistent with prior studies demonstrating an advantage of exact match cues on attention (Theeuwes et al., 2006; Wolfe et al., 2004). Cochrane et al. (2018) employed a between-subjects design, so it is possible that mental imagery takes longer to map onto attentional priority or to learn to effectively engage, but with repeated practice, ends up being more robust. Even though participants had 192 imagery cueing trials in our experiment, having to alternate with perceptual cueing trials may have interfered with the robustness of mental imagery that might have otherwise been evident without the inclusion of perceptual cues. With that potential caveat, from Experiment 1, we see an advantage of perceptual cueing over imagery cueing.

Experiment 2

Although several studies have compared attentional cueing between mental imagery and perception, they have been limited to the manipulation of low-level features (most notably with respect to colors; Cochrane et al., 2018; Reinhart et al., 2015). There is significant overlap in the recruitment of higher-level visual areas when comparing activations from mental imagery and perception using real-world objects (Dijkstra et al., 2019; Lee et al., 2012), but it remains unclear if corresponding attentional guidance is similar. Experiment 2 was thus conducted to investigate whether results for realworld objects mirrors findings for low-level features.

A caveat with respect to the interpretation of the significant interaction observed in Experiment 1 is that when participants are presented with a perceptual cue, they are primed to that exact shade of color, but there is no way of knowing what exact shade participants are visualizing on a given trial, which may differ at least slightly from the corresponding color presented during the search array. This is supported by prior findings that have demonstrated an advantage for exact-match cues in visual search (Theeuwes et al., 2006; Wolfe et al., 2004). With Experiment 2, we introduce some variability with respect to the perceptual cue by having the cue and search stimuli never be identical (e.g., jam drop cookie as opposed to chocolate chip) to better equate the two cue conditions. Given the greater neural overlap in higher-level visual areas (Dijkstra et al., 2019; Lee et al., 2012), in addition to the fact that perceptual cues and visual search stimuli were no longer exactly matched (see Theeuwes et al., 2006; Wolfe et al., 2004), we hypothesized little or no difference in the magnitude of validity effect attributable to perceptual and imagery cues.

Method

Participants

One hundred and five total participants were recruited from the Texas A&M University community, informed by the same sample size considerations. Data from one participant were omitted due to low accuracy (<50%). The final sample included 104 participants (61 females, and 41 males; 40.4% non-Hispanic whites) with a mean age of 19.6 (SD = 2.3). All reported normal or corrected-to normal visual acuity, normal color vision, and provided informed consent. Participants were compensated with course credit. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

Apparatus

The apparatus was identical that in Experiment 1.

Stimuli, design, and procedure

The sequence of displays was similar to that of Experiment 1, but the background color was changed to white and the duration of the imagery cueing display and search array were lengthened to accommodate the increase in task difficulty (Fig. 3). The perceptual cueing sequence was comprised of an image $(300 \times 300 \text{ pixels})$ displayed at the center of the screen for 1,500 ms. The imagery cueing sequence was comprised of the category name displayed at the center of the screen for 1,000 ms, followed by a 500-ms blank, and ending with an empty square (300x300 pixels) for 1800ms. Participants were instructed to look directly at the perceptual cue or to imagine a specific image in the cued category within the empty square.

There were 64 categories of objects, each of which comprised nine exemplars to be used as noncues and six exemplars to be used as perceptual cues, for a total of 960 objects. Object images were taken from the Konkle et al. (2010) stimulus set. The search display comprised four objects (300 \times 300 pixels) with three randomly chosen from at least two different categories, and the previously cued category served as either distractor or target. Participants were required to locate and report the orientation of the sideways 'T' (300 \times 100 pixels in 4-pixel width) amongst upright and upside



Fig. 3 Time course of trial events for Experiment 2. Each trial began with a fixation cross for a random duration between 400 and 600 ms, followed by the cueing sequence. The imagery cueing sequence begins with the category presented at the center of the screen for 1,000 ms, followed by a 500-ms blank, followed by an empty square for 1,500 ms. The perception cueing sequence consists of an object from the category that appears in the following search trial, presented for 1,500 ms. A 200-ms blank is displayed after the cueing sequence, followed by the search display for 1,200 ms, or until a response has been registered. If participants were too slow or incorrect, there would be a feedback display presented for 1,200 ms. Each trial ends with a blank screen for a random duration between 500 and 700 ms. (Color figure online)

down 'T's that have been overlaid on top of the objects. Participants were given 2,000 ms to respond with the 'Z' and 'M' key for left and rightward targets, and the feedback display from Experiment 1 was presented for 1200ms following incorrect or omitted responses.

Practice trials were similar to that of Experiment 1 with components of the trial sequence presented successively. There were 384 total search trials (half perceptual and half imagery cues) with breaks as in Experiment 1. Each category was cued three times per cue type, and served as the target 25% of the time. The target direction and target location were fully counterbalanced for each cue type. There were 96 vividness probe trials that were randomly inserted within each four-trial block, and participants had 30 seconds to respond.

Data analysis

Data analysis was conducted in the same manner as Experiment 1.

Results

A 2 × 2 ANOVA with cue type (perceptual vs. imagery) and validity (valid vs. invalid) as factors revealed a main effect of validity, F(1, 103) = 73.8, p < .001, $\eta_p^2 = .42$, and a main effect of cue type, F(1, 103) = 6.9, p = .010, η_p^2 = .06, with more efficient performance for imagery trials than perception trials, but no interaction effect, F(1, 103)= 3.1, p = .082, $\eta_p^2 = .029$ (Fig. 4A). Vividness ratings following perceptual cues were significantly greater than ratings following imagery cues, t(103) = 4.6, p < .001, d= .45, but did not correlate with validity effects for either cue types, r(102) = .06, p = .566, and r(102) = -.06, p =



Fig. 4 Mean (A) inverse efficiencies as a function of validity and cue type and (B) vividness ratings for each cue type for Experiment 2. Error bars represent standard error of the mean. (Color figure online)

.522, respectively (Fig. 4B). Vividness ratings ranged from 0-97.4 for perceptual cues and 2.72–100 for imagery cues. Removing participants with vividness scores less than 10 did not change our results.

Discussion

Some studies have suggested mental imagery to be weaker than perception (e.g., Pearson et al., 2015), which was reflected in weaker attentional guidance from mental imagery in Experiment 1. In Experiment 2, however, we see no interaction between validity and cue type; if anything, the cueing effect tended to be large for imagery cues. This is likely due to the fact that the perceptual cue was never a direct match to the stimulus in the search display (Theeuwes et al., 2006; Wolfe et al., 2004), mirroring the imprecise nature of mental imagery. With the correspondence between the cue and the search target more closely equated across cue conditions, we do not see a benefit of perceptual over imagery cues using real-world objects. It was unsurprising that we observed robust validity effects in general, as studies on mental imagery often have participants visualize complex objects rather than simple colors and shapes (Dijkstra et al., 2019; Kosslyn et al., 2001; Kosslyn & Thompson, 2003; Lee et al., 2012; Naselaris et al., 2015; Ragni et al., 2021). As stated previously, neuroimaging studies have demonstrated more substantial overlap between mental imagery and perception in higher level regions along the ventral visual stream (Dijkstra et al., 2019; Lee et al., 2012), which could also help explain the apparently comparable cueing effects we observed.

In both Experiments 1 and 2, we did not observe a correlation between the validity effect and vividness ratings for either type of cue. Prior studies reporting significant correlations in the case of mental imagery often do not probe the vividness of perception (e.g., Cochrane et al., 2021b), which in the present study might have influenced vividness ratings for mental imagery based on relative comparisons to perceptual cues.

Experiment 3

Our findings from Experiments 1 and 2 support and extend our understanding of how mental imagery guides attention across more than just simple features. Experiments 1 and 2, however, both still involved perceptually grounded visual search tasks that predominantly taxed the visual perceptual system of the brain (Peelen & Kastner, 2011; Serences et al., 2005). Attention is much broader than the biasing of perceptual representations, and it remains unclear if mental imagery also guides more executive attention. To continue our exploration into the extent of mental imagery effects on attention, we conducted a final experiment using the colorword Stroop task.

We limited the stimuli to incongruent Stroop stimuli to focus our investigation on the modulatory effects of congruency with respect to the perceptual or imagery cue, both for the relevant (color-ink) and irrelevant (color-word) dimension. We removed vividness ratings for this experiment since vividness ratings were not related to the strength of cueing effects in either Experiment 1 or 2 and excluding participants with low vividness ratings did not change the outcome of our analyses.

Method

Participants

One hundred and twenty-four participants were recruited from the Texas A&M University community, again informed by the same sample size considerations. Data from 19 participants were excluded from analyses due to low proportion of correct responses (<50%). The final sample included 105 participants (47 females, and 57 males; 54.3% non-Hispanic Whites) with a mean age of 18.6 (SD = 0.72). All reported normal or corrected-to normal visual acuity, normal color vision, and provided informed consent. Participants were compensated with course credit. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

Apparatus

The apparatus was identical to that of Experiment 1 and 2.

Stimuli, design, and procedure

The cueing procedure for Experiment 3 was identical to that of Experiment 1, as was the order and timing of trial events (see Fig. 1). The stimulus display was comprised of only incongruent color-word Stroop stimuli (red, green, blue, and yellow in lowercase; 24 pixels), and participants were required to use the keyboard to respond to the ink color of the word while ignoring the meaning of the word. Participants received practice with the color-key pairing before the start of the experiment: "z" = red, "x" = green, "n" = blue, "m" = yellow. They were first presented with colored squares (400 pixels \times 400 pixels) along with the matching key for each color, before receiving four practice trials responding to single colored-squares with the appropriate key. Afterwards, participants were given four untimed and four timed practice trials with incongruent color-word Stroop stimuli. Participants then completed practice for the cueing sequence and the entire trial sequence like in Experiments 1 and 2. There were 384 total trials that were equally counterbalanced by cue type, cue color, ink color and (incongruent) color-word, such that there were 96 total trials per block with 60 second breaks after each block. Unlike the prior two experiments, this experiment did not have any vividness probes.

Data analysis

Data for Experiment 3 were analyzed the same way except separately for color-ink (task-relevant dimension—i.e., the dimension that participants were supposed to attend and respond to) and color-word (irrelevant dimension, i.e., the dimension that participants were supposed to ignore). Valid and invalid trials for both color-ink and color-word trials were defined in the same way, except valid for color-ink was when the cue color matched the ink color of the stimuli and valid for color-word was when the cue color matched the color indicated by the word.

Results

Color-ink (relevant dimension)

A 2 × 2 ANOVA on the ink color of the word with cue type (perceptual vs. imagery) and validity (valid vs. invalid) as factors revealed a significant main effect of cue type, F(1, 104) = 30.0, p < .001, $\eta_p^2 = .224$, in which participants were less efficient following imagery cues (Fig. 5A). The main effect of validity, F(1, 104) = 2.2, p = .140, $\eta_p^2 = .021$, and the interaction effect, F(1, 104) = 1.72, p = .193, $\eta_p^2 =$.016 were not significant. In order to investigate the nonsignificant main effect of validity, which was expected but not observed, we conducted a post hoc pairwise comparison of the validity effects following perceptual cues and imagery cues. Trials following perceptual cues exhibited a significant validity effect, t(105) = 2.23, p = .028, d = .22, but trials following imagery cues did not, t(105) = .39, p = .07, d = .04.

Color-word (irrelevant dimension)

A 2 × 2 ANOVA on the meaning of the word (color-word) revealed a significant main effect of cue type, F(1, 104) = 22.91, p < .001, $\eta_p^2 = .180$, in which participants were less efficient following imagery cues (Fig. 5B). The main effect of validity, F(1, 104) = .3, p = .610, $\eta_p^2 = .002$, and the interaction effect, F(1, 104) = .9, p = .339, $\eta_p^2 = .009$ were not significant. Our post hoc comparisons revealed that neither trials following perceptual cues nor imagery cues exhibited a validity effect, t(105) = -1.18, p = .240, d = .12, and t(105) = .26, p = .795, d = .03, respectively.

Discussion

In Experiment 3, we extended our comparison between perceptual and imagery cueing on attention from visual search to executive attention with the color-word Stroop task. We observed significant validity effects with respect to the ink color of the word for trials following perceptual cues, but not for trials following imagery cues. On the other hand, neither perceptual nor imagery cues significantly influenced word processing.

Had participants used a strategy of phonologically rehearsing the words for the cued colors, especially with respect to imagery cues, we would have expected more robust effects of the cues on the processing of the words (irrelevant dimension). It seems that the influence of both imagery and perceptual cues is restricted largely to information processing within the visual system in the form of guidance during visual search and does not extend to the resolution of conflict generated by visually presented words.



Fig. 5 Mean inverse efficiencies as a function of validity and cue type for (A) color-ink (relevant dimension) and (B) color-word (irrelevant dimension). Error bars represent standard error of the mean. (Color figure online)

Interestingly, in Experiment 3, performance was worse following imagery cues than perceptual cues, which was not the case in either of the prior two experiments. When the task was more cognitively demanding, we see a cost to engaging in mental imagery that suggests that mental imagery is itself to some degree cognitively demanding, resulting in something akin to a greater switch cost compared with the passive viewing of perceptual cues.

General discussion

In the present study, we investigated attentional cueing from mental imagery and perception across a set of attention tasks. As in several prior studies (Cochrane et al., 2018, 2019; Cochrane, Townsend, et al., 2021a; Cochrane, Wang, et al., 2021b; Moriya, 2018; Reinhart et al., 2015; Wantz et al., 2015), we observed robust cueing effects for both mental imagery and perceptual cues for low-level color-based visual search and extend these findings to visual search with real-world objects. The influence of mental imagery on the allocation of attention extends beyond the prioritization of basic visual features and may therefore have broad implications for subsequent information processing in complex realworld situations. Through Experiment 3, we also highlight a boundary condition for the influence of both perceptual and imagery cueing on the control of attention, which does not appear to extend beyond the attentional selection of visual stimuli to the resolution of conflict generated by color-word Stroop stimuli.

There have been differing reports concerning the strength of mental imagery cues compared with perceptual cues (e.g., Cochrane et al., 2018, 2023), and our findings suggest that the imprecise nature of mental imagery could help explain this. When the perceptual cues were not an exact match to a visually presented stimulus in the case of complex objects, as is necessarily the case with mental imagery of an exemplar of a category, effects of perceptual and imagery cues were comparable. In contrast, an advantage for perceptual cues was evident when they were an exact feature match in the context of color-defined stimuli (Wolfe et al., 2004). It might also be the case that participants are more adept at generating a mental image of a meaningful object as opposed to a more abstract color-square, producing more robust consequences for imagery on attention in the case of the former. At least when it comes to cueing within-category, we do not see an advantage for actually presenting an image of an exemplar of the category compared with asking participants to imagine an exemplar. Neuroimaging studies have identified greater amounts of overlap in higher-order visual pathways between mental imagery and perception (Lee et al., 2012), which provides a potential mechanistic explanation for the difference in the relative strength of the validity effect between perceptual and imagery cues across Experiments 1 and 2.

It is important to note that several prior comparisons of attentional guidance between perception and mental imagery employed either a between-subjects approach (e.g., Cochrane, Wang, et al., 2021b), or a blocked manipulation (e.g., Cochrane et al., 2023). In these instances, imagery had a stronger effect than perception on attention (e.g., Cochrane, Wang, et al., 2021b), and continued to influence attention in no-imagery blocks such that the effect of perception was a combination of imagery and perception (Cochrane et al., 2023; see also Cochrane et al., 2018, 2019; Cochrane, Townsend, et al., 2021a; Reinhart et al., 2015). It is likely that after mental imagery is deployed repeatedly and consecutively, the imaged representation strengthens and persists, eventually overtaking perception. For all three of our experiments, we employed trial-by-trial cueing where the order of imagery and perception cues were randomly distributed. Perhaps switching between imagery and perception prevented the imagery-related representation from gaining momentum, resulting in a perception benefit.

Executive attention, on the other hand, did not seem to be impacted by mental imagery cues in either the task-relevant or irrelevant dimension in Experiment 3, suggesting that the effects of mental imagery on attention are to some degree taskdependent (see also Cochrane et al., 2022, for similar findings with the attentional blink task). When the task involved the guidance of attention in the context of visual search (Experiments 1 and 2), robust cueing effects were evident, but when the task involved the resolution of conflict between task-relevant and task-irrelevant perceptual input, significant cueing effects were in general not observed. In this respect, the influence of perceptual and imagery cues were again similar, although there was limited evidence for a small effect of perceptual cues on the processing of task-relevant information. From Experiment 3, we conclude that the influence of perceptual cues and imagery on the allocation of attention is particularly robust in the context of explicitly visual information processing such as attentional guidance in the context of visual search (e.g., Cochrane et al., 2018, 2019; Cochrane, Townsend, et al., 2021a; Cochrane, Wang, et al., 2021b; Moriya, 2018; Reinhart et al., 2015) and some forms of perceptual biases (Chang et al., 2013; Pearson et al., 2008); this is broadly consistent with the idea that mental imagery recruits perceptual representations (Dijkstra et al., 2019; Kosslyn et al., 2001; Kosslyn & Thompson, 2003; Lee et al., 2012; Naselaris et al., 2015; Ragni et al., 2021), which may influence attentional guidance in a similar fashion to stimulus representations held in working memory (Harrison & Tong, 2009; Olivers et al., 2006, Serences et al., 2009; Soto et al., 2005, 2008).

Since our stimuli in Experiment 3 only included incongruent color-word Stroop stimuli, we cannot rule out the possibility that either mental imagery cueing or perceptual cues would have a more robust effect on neutral or congruent trials, when there is no conflict between a task-relevant and task-irrelevant dimension. The elevated task demands brought about by the need for conflict resolution may have overshadowed more subtle effects of cueing that might have been measurable in a task more comparable in difficulty to that of Experiments 1 and 2. Nevertheless, we do not see evidence for an influence of perceptual cues and imagery on the resolution of conflict generated by taskirrelevant perceptual input.

Finally, we found some evidence that engaging in mental imagery might engage cognitive processes in a manner capable of producing a switch cost or otherwise impair performance on a subsequent task. Such a cost was only evident in Experiment 3, however. This might be because the Stroop task was more cognitively demanding than the visual search tasks of Experiments 1 and 2, although given that it was only observed once across the three experiments, we hesitate to draw any strong conclusions about any performance costs associated with engaging in mental imagery. However, Cochrane, Wang, et al. (2021b) did show that compared with a control group, participants required more time to initiate an eye movement following mental imagery, indicating a potential disengagement cost that may have required more cognitive resources to overcome. The degree of cognitive demand associated with mental imagery reflects a potentially interesting and fruitful topic for future research relating mental imagery to attention and cognition.

In conclusion, our study offers novel insights into the scope with which mental imagery influences the allocation of attention. We extend the influence of mental imagery to attention to real-world objects, and further show that mental imagery for simple colors might underestimate the relative strength of mental imagery's effect compared with perceptual cues. At the same time, the influence of mental imagery on the allocation of attention seems to be the most pronounced for distinctly visual attention in the context of a visual search task.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13414-023-02747-9.

Funding This research was supported by the NIH [R01-DA406410] to B.A.A.

Data availability All anonymized study data, experiment code, data wrangling, and analyses code are free available on the Open Science Framework (https://osf.io/jfb37/?view_only=24394219a32c41f78823 fc5e31703c33).

Declarations

Conflicts of interest The authors declare no conflicts of interest.

References

- Brady, T. F., & Störmer, V. S. (2021). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(7), 942–958. https://doi.org/10.1037/xlm0001014
- Bugatus, L., Weiner, K. S., & Grill-Spector, K. (2017). Task alters category representations in prefrontal but not high-level visual cortex. *NeuroImage*, 155, 437–449. https://doi.org/10.1016/j. neuroimage.2017.03.062
- Chang, S., Lewis, D. E., & Pearson, J. (2013). The functional effects of color perception and color imagery. *Journal of Vision*, 13(10), 4–4. https://doi.org/10.1167/13.10.4
- Cochrane, B. A., Nwabuike, A. A., Thomson, D. R., & Milliken, B. (2018). An imagery-induced reversal of intertrial priming in visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(4), 572–587. https://doi.org/ 10.1037/xlm0000470
- Cochrane, B. A., Pratt, J., & Milliken, B. (2023). Top-down then automatic: Instructions can continue to influence visual search when no longer actively implemented. *Attention, Perception, & Psychophysics*, 85(1), 76–87. https://doi.org/10.3758/ s13414-022-02558-4
- Cochrane, B. A., Sclodnick, B., MacLellan, E., & Milliken, B. (2022). It hurts more than it helps: Cuing T1 with imagery can impair T2 identification in an attentional blink task. *Attention*, *Perception*, & *Psychophysics*, 84(7), 2141–2154. https://doi.org/ 10.3758/s13414-022-02552-w
- Cochrane, B. A., Siddhpuria, S., & Milliken, B. (2019). Cueing color imagery: A critical analysis of imagery-perception congruency effects. *Journal of Experimental Psychology: Learning, Mem*ory, and Cognition, 45(8), 1410–1421. https://doi.org/10.1037/ xlm0000653
- Cochrane, B. A., Townsend, B., Lee, R. C., Legere, J. K., Milliken, B., & Shedden, J. M. (2021a). Visual imagery influences attentional guidance during visual search: Behavioral and electrophysiological evidence. Attention, Perception, & Psychophysics, 83, 58–66. https:// doi.org/10.3758/s13414-020-02208-7
- Cochrane, B. A., Wang, C., Pratt, J., Milliken, B., & Sun, H. J. (2021b). Comparing imagery and perception: Using eye movements to dissociate mechanisms in search. Attention, Perception, & Psychophysics, 83, 2879–2890. https://doi.org/ 10.3758/s13414-021-02336-8
- Collegio, A. J., Nah, J. C., Scotti, P. S., & Shomstein, S. (2019). Attention scales according to inferred real-world object size. *Nature Human Behaviour*, 3(1), 40–47. https://doi.org/10.1038/ s41562-018-0485-2
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, 58(3), 306–324. https://doi.org/10.1016/j.neuron. 2008.04.017
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neurosci*ence, 3(3), 201–215. https://doi.org/10.1038/nrn755
- Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, 270(5237), 802–805. https://doi.org/10.1126/science.270.5237.80
- de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47, 1–12. https://doi.org/10.3758/s13428-014-0458-y
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. (2019). Shared neural mechanisms of visual perception and imagery. *Trends in*

Deringer

Cognitive Sciences, 23(5), 423–434. https://doi.org/10.1016/j. tics.2019.02.004

- Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005
- Gunseli, E., Olivers, C. N. L., & Meeter, M. (2016). Task-irrelevant memories rapidly gain attentional control with learning. *Journal* of Experimental Psychology: Human Perception and Performance, 42(3), 354–362. https://doi.org/10.1037/xhp0000134
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635. https://doi.org/10.1038/nature07832
- Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. *Elife*, 7, e32816. https://doi.org/10.7554/ eLife.32816
- Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. *Journal of Cognitive Neuroscience*, 21, 760–775. https:// doi.org/10.1162/jocn.2009.21039
- Hopf, J. M., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., & Heinze, H. J. (2000). Neural sources of focused attention in visual search. *Cerebral cortex*, 10(12), 1233–1241. https://doi.org/10.1093/cercor/10.12.1233
- Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. PLOS ONE, 6(12), e29221. https://doi.org/10.1371/ journal.pone.0029221
- Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual longterm memory for real-world objects. *Journal of experimental Psychology:* general, 139(3), 558–578. https://doi.org/10.1037/a0019165
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature reviews neuroscience*, 2(9), 635– 642. https://doi.org/10.1038/35090055
- Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, 129, 723–746. https://doi.org/10.1037/0033-2909. 129.5.723
- Kristjánsson, Á., & Ásgeirsson, Á. G. (2019). Attentional priming: recent insights and current controversies. *Current Opinion in Psychology*, 29, 71–75. https://doi.org/10.1016/j.copsyc.2018. 11.013
- Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLOS ONE*, 10(6), -e0130834. https://doi.org/10.1371/journal.pone. 0130834
- Lee, S. H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, 59(4), 4064–4073. https://doi.org/10.1016/j.neuroimage.2011.10.055
- Lee, S. H., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nature Neuroscience*, 16(8), 997–999. https://doi.org/10.1038/ nn.3452
- Liao, M.-R., Grindell, J., & Anderson, B. A. (2021). jsPsych solutions for vision science. *PsyArXiv Preprint*. https://doi.org/10.31234/ osf.io/gqf2m
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31, 291–308. https://doi.org/10.1111/j.1469-8986.1994. tb02218.x
- Moriya, J. (2018). Visual mental imagery influences attentional guidance in a visual-search task. Attention, Perception, & Psychophysics, 80(5). https://doi.org/10.3758/s13414-018-1520-0

- Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage*, 105, 215–228. https://doi.org/10.1016/j.neuroimage.2014.10.018
- 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. https:// doi.org/10.1037/0096-1523.32.5.1243
- Pearson, J., Clifford, C. W., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, 18(13), 982–986. https://doi.org/10.1016/j.cub.2008.05.048
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, 19(10), 590–602. https://doi.org/10. 1016/j.tics.2015.08.003
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of* the National Academy of Sciences, 108(29), 12125–12130. https:// doi.org/10.1073/pnas.1101042108
- Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Sciences*, 7, 113–118. https://doi.org/10.1016/S1364-6613(03)00003-2
- Ragni, F., Lingnau, A., & Turella, L. (2021). Decoding category and familiarity information during visual imagery. *NeuroImage*, 241, 118428. https://doi.org/10.1016/j.neuroimage.2021.118428
- Reinhart, R. M. G., McClenahan, L. J., & Woodman, G. F. (2015). Visualizing trumps vision in training attention. *Psychological Science*, 26, 1114–1122. https://doi.org/10.1177/0956797615577619
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulusspecific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207–214. https://doi.org/10.1111/2Fj. 1467-9280.2009.02276.x
- Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science*, 16(2), 114–122. https://doi.org/10.1111/j.0956-7976.2005.00791.x
- Slinn, C., Nikodemova, Z., Rosinski, A., & Dijkstra, N. (2023). Vividness of visual imagery predicts performance on a visual working memory task when an imagery strategy is encouraged. *PsyArXiv Preprint*. https://doi.org/10.31234/osf.io/34wsv
- 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. https://doi.org/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. https://doi.org/10.1016/j. tics.2008.05.007
- Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition*, 14(4/8), 466–489. https://doi.org/10. 1080/13506280500195110
- Tong, F. (2013). Imagery and visual working memory: one and the same? Trends in Cognitive Sciences, 17(10), 489–490. https://doi. org/10.1016/j.tics.2013.08.005
- Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3., pp. 200–239). Erlbaum.
- Townsend, J. T., & Ashby, F. G. (1983). Stochastic modeling of elementary psychological processes. Cambridge University Press.
- Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, 5, 81–92. https://doi. org/10.1167/5.1.8

- Wantz, A. L., Borst, G., Mast, F. W., & Lobmaier, J. S. (2015). Colors in mind: A novel paradigm to investigate pure color imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 1152–1161. https://doi.org/10.1037/xlm00 00079
- Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 483–502. https://doi.org/10.1037/0096-1523.29.2.483
- Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of topdown guidance in visual search. *Vision Research*, 44, 1411–1426. https://doi.org/10.1016/j.visres.2003.11.024
- Woodman, G. F., Arita, J. T., & Luck, S. J. (2009). A cuing study of the N2pc component: An index of attentional deployment to objects

rather than spatial locations. *Brain Research*, *1297*, 101–111. https://doi.org/10.1016/j.brainres.2009.08.011

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature Neuroscience*, 5(10), 995–1002. https://doi.org/10.1038/nn921

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.